Culture in the Making:

Jointly structuring shared spaces of meaning and action in infant-caregiver-object interactions over the first year of life

Nicole Rossmanith

This thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy at the University of Portsmouth.

September 2017
Abstract

How do infants grow in and into culture? How do they become competent participants in networks of meaning-making including people and artefacts? Typically research addressing these questions starts looking from the end of the first year, when infants’ early dyadic social interactions are supposed to turn “triadic”, that is, are extended to include objects and aspects of the world, only then giving rise to cultural learning, symbol use, co-operative participation. In the face of mismatches with everyday experience and counter-evidence from recent empirical studies, we revisit several research programs dealing explicitly with the development of infant-caregiver-object interactions to arrive at a critical appreciation of how the concept of triadic interaction and the core narrative developed. On this basis, and drawing from embodiment, situatedness, and dynamical systems, we construct our own approach for exploring the development of jointly practicing social object activities, which we frame in terms of attention- and action-coordination.

We conducted a naturalistic longitudinal study visiting 16 infants in their homes once a month from 3-12 months and documenting infant-caregiver-object interactions. Adapting techniques from interaction and conversation analysis, and using macro- and micro-analysis of video recordings, we 1) explore and analyse the development of book sharing as a model activity over the first year of life, we 2) develop concepts and methods to characterize and analyse different modes of engagement, and patterns of coordination, infants and caregivers employ in a wide variety of ecological contexts, and 3) introduce the notion of jointly moving through affect-imbued action arcs together. Finally, we 4) sketch a tentative developmental trajectory of participation in social object activities, reconceptualising the shift from “dyadic” to “triadic” interactions as “jointly structuring shared spaces of meaning-and-action” of increasing complexity. In particular, we propose the mapping of complex action structures on familiar affect-imbued action arcs as a bridge towards activities such as collaborative participation, symbol and conventional object use, cultural learning and co-creation.
### Table of Contents

Abstract ................................................................................................................................. 2

Declaration ............................................................................................................................. 7

List of Figures .......................................................................................................................... 8

Acknowledgements ................................................................................................................ 10

Dissemination ........................................................................................................................ 12

  Publications .......................................................................................................................... 12

  Conference Presentations ................................................................................................... 12

1 General Introduction: The development of triadic infant-caregiver-object interactions – encountering a narrative .............................................................................................................. 14

  1.1 Investigating growing in and into culture ............................................................................ 14

  1.2 Encountering the narrative: A first approximation ................................................................. 17

  1.3 “Triadic interaction” as “joint attention” ........................................................................... 19

  1.4 Counter-evidence ............................................................................................................. 20

  1.5 Aims of the thesis and chapter overview .......................................................................... 21

Part I: A critical theoretical analysis of approaches investigating infant-caregiver-object interactions .................................................................................................................................. 27

2 The development of triadic infant-caregiver-object interactions revisited .......................................................... 27

  2.1 Aims of this chapter ........................................................................................................... 27

  A comparison of different accounts for the emergence of triadic interactions .............. 28

  2.2 Accounts I: General cognitive complexity and differentiation (Piaget, Bates, Sugarman-Bell) ......................................................................................................................... 28

  2.3 Accounts II: Social relating and co-ordination of social and object realms (Trevarthen & Hubley, Adamson & Bakeman) ................................................................. 58

  2.4 Accounts III: Cognitive Accounts (Tomasello and colleagues) ..................................... 79

  2.5 Accounts IV: Cultural Accounts (Vygotsky, Rodriguez, Zukow-Goldring) ................ 99
2.6 Accounts V: Current and interactive accounts ........................................127

2.7 Context box: On the separation of (manual) object - and (language-mediated) people interactions; parallels and interconnections between infant research, human- and world-conceptions, and socio-political developments .................................................................................136

2.8 Conclusions and Outlook .............................................................................144

Part II An Empirical investigation into infant-caregiver-object interactions over the first year of life: a naturalistic, longitudinal study ..............................150

3 Jointly structuring triadic spaces of meaning and action: book sharing from 3 months on ...........................................................................................................150

3.1 INTRODUCTION ...............................................................................................151

3.2 MATERIALS AND METHODS ........................................................................154

3.3 RESULTS AND DISCUSSION ...........................................................................157

3.4 ECOLOGIES IN TRANSFORMATION: SKETCHING A DEVELOPMENTAL TRAJECTORY OF BOOK SHARING OVER THE FIRST YEAR ..............................179

3.5 CONCLUSIONS, GENERAL DISCUSSION, AND OUTLOOK ..........................192

3.6 ACKNOWLEDGMENTS ....................................................................................203

3.7 GLOSSARY .......................................................................................................203

4 Varieties of triads: different patterns of joint attention-and-action-coordination in different ecological infant-caregiver-object activity contexts and their development over the first year ..........................205

4.1 Introduction .......................................................................................................205

4.2 Developing a conceptual framework and methodology: 1) A macro-level perspective ..............................................................................................................208

4.3 Micro Analysis: Bottom lift in nappy changing as an example for infants’ active participation in caregiving tasks ........................................................................238

Part III: Integration of empirical and theoretical parts .....................................249

5 Structure and Openness in the Development of Self in Infancy ......................249

5.1 Patterns, Structure, and Openness ....................................................................250
5.2 Methodological Challenges and Theoretical Stances on Self in Infancy
251

5.3 The Emergence of Self in Infant Engagement..............................................252

5.4 Infant Selves in Early Dialogues .................................................................254

5.5 Selves Enacted through Joint Participation in Shared Cultural Routines
259

5.6 Jointly Modifying and Negotiating Shared Routines.................................261

5.7 Opening a Conversation with Spiritual Traditions .......................................263

5.8 Structures Enhancing and Hindering Engagement.......................................267

5.9 Structure and Openness .............................................................................269

5.10 Acknowledgments ......................................................................................269

6 The development of participation in infant-caregiver-object practices:
changes in attention-and-action co-ordination.............................................270

6.1 Setting the stage..........................................................................................270

A tentative trajectory of the development of participation in terms of
coordination of attention and action over the first year of life ......................273

6.2 3-4 months: infants participate in co-ordinated affect imbued arc-shaped
social object routines.....................................................................................273

6.3 4-6 months: a range of social object routines was fluently established,
and their shapes, i.e. space-time-intensity contours, playfully varied............277

6.4 Intermission: Initiating and re-eliciting established routines from 6
months on ........................................................................................................283

6.5 6-9 months: combinatorial variation of (object mediated) social play
routines versus solitary object play.................................................................284

6.6 9-12 months: co-composing complex social action structures including
manual object-relative object acts. .................................................................287

Conclusions and Outlook ..............................................................................290

6.7 From engaging in simple (dyadic) interchanges to participating in,
negotiating and co-creating complex (triadic) interactions: the role of jointly
enacting affect-imbued action arcs.................................................................290
6.8 Shifting the perspective from triadic interaction to jointly structuring shared multi-modal sense-and-action(land)scapes

7 General Discussion

8 Bibliography
Declaration

This work is supported by the Marie-Curie Initial Training Network, "TESIS: Towards an Embodied Science of InterSubjectivity" (FP7-PEOPLE-2010-ITN, 264828). Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word Count: 86846
List of Figures

Figure 1.1: Illustration of the developmental narrative ............................................. 16

Figure 1.2: Schematic depiction of dyadic interactions ............................................. 18

Figure 1.3: Momentary snapshot showing a 4 month old immersed in a rich cultural context involving people and objects ................................................................. 18

Figure 2.1: Collaborative activity after Tomasello et al. 2005 ................................. 86

Figure 2.2: Sign mediation after Vygotsky ................................................................. 102

Figure 3.1: Social book sharing interactions (red) and solitary book exploration episodes (green) for all infants from 3–12 months ......................................................... 159

Figure 3.2: Three examples of book sharing with books specifically designed for young infants .......................................................... 163

Figure 3.3: Mother multimodally presenting a book ................................................. 165

Figure 3.4: ELAN analysis detail showing pitch (red) and intensity (green) curves. The mother is reading a picture book about animal actions to her 4-month-old son enacting the essence of “leaping” and “jumping” (a rising motion) through a rising intonation contour (highlighted) .................................................. 170

Figure 3.5: ELAN analysis detail showing pitch (red) and intensity (green) curves. Mother enacting and embodying a “BI::g YA:::wn and a STREtch” vocally and through acting on the 5-month-old infant’s body (highlighted). Upper case letters (A–D) map upper row stills to ELAN time line ................. 170

Figure 3.6: ELAN analysis detail showing pitch (red) and intensity (green) curves. The mother is building up an action arc through surprise exclamations of increasing intensity and pitch before releasing the tension through a quick page turn ........................................................................................................ 173

Figure 3.7: ELAN analysis detail showing pitch (red) and intensity (green) curves. This book sharing interaction at 4 months unfolds as smooth interplay between the actions of caregiver and infant ............................................ 177

Figure 3.8: ELAN analysis detail of book sharing interaction with 6-month-old infant sitting at a 90° angle on the mother’s lap. Infant and mother looking at book together .......................................................... 182
Figure 3.9: ELAN analysis detail of book sharing interaction with 6-month-old infant sitting at a 90° angle on mother’s lap. Mother using extensive voice and hand acting to illustrate animals and animal sounds..........................185

Figure 3.10: Book sharing interaction with 6-month-old infant sitting at a 90° angle on mother's lap including extensive voice and hand acting..................188

Figure 3.11: Advanced book sharing interactions..................................................191

Figure 3.12: Ecologies in transformation. The table gives an overview of book sharing as it changes over the first year.................................................................197

Figure 4.1: Sources and targets of action...............................................................210

Figure 4.2: Connection between nodes...............................................................210

Figure 4.3: Modes of engagement........................................................................211

Figure 4.4: Ecological Activity Contexts...............................................................213

Figure 4.5: Illustrations of caregiving activities in (from top): feeding, bathing, soothing, and nappy change.................................................................226

Figure 4.6: Nappy changing activity, details see text.................................239

Figure 4.7: The mother shapes nappy change into an action arc...................246

Figure 6.1: The development of participation. Details see text..........................272

Figure 6.2: Shaping a peek-a-boo activity into a series of affect-imbued shared action arcs.................................................................296

Figure 6.3: Understanding complex actions through mapping them on affect-imbued shared action arcs.................................................................300
Acknowledgements

First of all I send out a “thank you!” to all the participating babies and families, without whom this thesis wouldn’t exist. Thank you, for opening up your homes and sharing this most precious time with me!

Related to that, many thanks to Tina Reichelt and Elisabeth Zimmermann, for their interest and openness to get involved and their kind, enthusiastic support with the data collection, making the audacious plan of including and hopping between two countries actually work!

This thesis grew out of the encounter of my long standing interest how a shared world of meaning can develop out of shared object interactions, Alan Costall’s insistence on the importance of materiality conspicuously absent in depictions of early infant-mother interaction, and Vasu Reddy’s 2nd person approach, in particular her account of the development of (affective) attentional engagements.

Since this thesis is part of a longer journey exploring how meaning is created, it also owes a lot to a great number of curious and persistent human beings who set out to explore these issues long before and more thoroughly than me: such as Francisco Varela, Terrence Deacon, David Dusenbery, just to name a few, and first and foremost Karl Bühler. So much so, that it will take more time, distance and reflection to exactly figure out and pay tribute to their impact to my thinking, since my ways of thinking partially consist of theirs…

I am grateful to Alan Costall, who is a never-ceasing well of information and insight, and to whom I owe one of the most productive supervisory “perturbations”, as he remarked, while we were talking about Bühler: “What do we actually need those sign-notions for anyway?" - Albeit throwing my thinking slightly off balance back then, the question kept simmering in the back of my head and finally forced me to think through how to “translate” signs into forms of co-ordination. So I hope - even if it didn’t always become immediately obvious, since I always need some time to digest - it becomes visible between the lines of the thesis, that our long discussions did have some impact!

A big thank you also to Beatriz López and Alessandra Fasulo, for your taking a fresh look at data in our sharing sessions, and your open doors for support. Many thanks for fruitful discussions to Hanne De Jaegher & Ezequiel Di Paolo (who also gave me a warm welcome during my secondment in San Sebastian!), as well as to my other TESIS colleagues, in particular to my TESIS sibling in Portsmouth: Vale, I miss you and our conversations a lot! … and to Markus Peschl, who helped me to disentangle my thoughts via skype… and most recently to my two examiners Eszter Somogyi and Joanna Račzaszek-Leonardi who created a very engaging viva with thought-provoking
and helpful remarks from experimental and language perspectives, respectively. Joanna’s comment “There seems to be a lot of semiotic thought implicit in this thesis…,” closed the circle Alan had opened up and is taking me on the next adventure to audaciously tackle language together!

A deep-felt and smiling thanks to my old friends in Vienna, Tina, Alex, Salka & Roswitha who despite my being far away and spending far to little time with them are there for me as if I’d never been away...

My deepest gratitude and love to my family, in particular my dad, for his love and support during these PhD years, also when it meant being too far away for too long (even if it is in the places I got to know and love during our long camping travels with our motor van). My grateful thoughts also go to those to whom we had to say good-bye during the PhD: my granny and my father in law.

A deep thanks going far beyond these few lines to my first supervisor, Vasu Reddy - I know I'm not the easiest person to supervise, in particular once I’m hooked on a problem which I feel needs to be solved. Thanks for bearing with me on this journey, thanks for your trust and engagement, for giving me freedom, and support. And thank you for reminding and making me more aware of what it means to hold up and stand up for authentic and affective engagement in science – we need it urgently! – as well as for reassuring me in my own struggles to live this engagement in science.

And finally a loving thanks to my partner Andreas for sharing our research and our lives, across different continents, thinking styles, and temperaments; for engaging in midwifery for my thoughts during this thesis – bearing my struggles to get my idea across in conversation with more or less (im ;-):patience, silly and wise remarks, when not – at times and to my frustration – quickly running off with my thoughts in completely different directions – Thanks for your love, support and being there, and I'm looking forward to travel towards new horizons together!
Dissemination

Publications


Conference Presentations

Mutual understanding and sharing an object world in the first year of life, Poster presented at the *IMPRS workshop: Perspectives on the ontogeny of mutual understanding*, October 1-3, 2015, Max Planck Institute for Psycholinguistics, Nijmegen, NL

Action anticipation and co-ordination in infant-caregiver-object interactions over the first year of life. Talk given at the *Loch Lomond Symposium on Action Anticipation*, September 2-3, 2015, Ross Priory, Loch Lomond, Scotland

Jointly structuring multimodal spaces of attention & action in the first year of life. Poster presented at the *social cognition meeting/workshop*, August 27-29, 2015, Bangor, UK

The many faces of “jointness” in the development of infant-caregiver-object interactions over the first year of life. Talk given at the *Joint Action Meeting (JAM)*, July 1-4, 2015, CEU, Budapest, Hungary

The development of participation in cultural practices involving everyday objects during the first year of life. Talk given at *Revisiting Participation - Language and Bodies in Interaction*, June 24-27, 2015, University of Basel, Switzerland


Karl Bühler meets Enactive Cognitive Science. Towards a Science of social co-regulation and making sense. Talk given at the *IAS Centre for Life, Mind, and Society, Department of Logic and Philosophy of Science*, University of the Basque Country, Donostia/SanSebastian, Spain, October 6, 2014

Multimodal shared spaces of attention and action in early infant-caregiver book sharing interactions. Talk given as part of the symposium: “Coordinating bodies: How hands structure learning opportunities in natural social contexts over the first two years of life” at the *XIX Biennial International Conference on Infant Studies*, July 3-5, 2014, Berlin, Germany
Structuring shared spaces of meaning and action together: booksharing in early infancy. 3rd: TESIS Summer School: Meeting Minds. Ways of Engaging and Making Sense, June 2014, Midhurst, UK

Shared spaces of meaning: early triadic infant-caregiver-object interactions. Talk given as part of the TESIS symposium “Interacting with objects in typical and atypical development: participation, anticipation, and imitation” at the 16th European Conference on Developmental Psychology, September 3-7, 2013, Lausanne


Growing in & into culture: early triadic interactions between infants, caregivers & objects and the creation of "joined spaces of meaning. Talk given at the TESIS midterm-review meeting, February 2013, Heidelberg.


1 General Introduction: The development of triadic infant-caregiver-object interactions – encountering a narrative.

*For the psychological theoretician this [...] triangle will be as interesting and cause as many difficulties as the three-body-problem [is and does] for the physicist.*

*Karl Bühler (1927) about 3-pole interaction.*

1.1 Investigating growing in and into culture

How do we arrive at a shared world? People jointly act in, communicate about, transform and co-create their world. They do so by smoothly navigating and building complex networks of meaning-making involving persons, objects, and symbols. How do we do this?

How, in particular, do children grow in and into culture? How do they become competent participants in cultural practices, in networks of meaning-making including people and artefacts? Language and tool use, as well as more recently co-operation as well as teaching and social learning are regarded as essential characteristics of human culture, as means to jointly create as well as to transmit and further develop them across many generations, thus giving rise to a cumulative cultural process (Tomasello, Kruger, & Ratner, 1993).

Researchers interested in the development of cultural and social learning typically focus on a time period starting from the end of the infant's first year of life: at this point the eyes of parents as well as psychologists are caught by a range of behaviours becoming abundant all at the same time: infants engage in giving and taking objects together with their caregivers, waving hello and goodbye, imitating acts on objects, following instructions, labelling objects, and often seeking visual contact and checking back with their parents while doing something. Those activities appear to make such a powerful impression on observers and are experienced as such a striking change by people interacting with infants that this period is regarded as a major shift in development, a “transition” from one phase into another, e.g. into intersubjectivity (Stern, 1985) or from “primary” to “secondary intersubjectivity” (Trevarthen & Hubley, 1978) and is sometimes even referred to as a “revolution” as in Tomasello’s “9 month
(social-cognitive) revolution” (Tomasello, 1995) or a “quantum leap” (Stern, 1985).

What is going on here? What exactly is changing? How is this change coming about? Looking for something that those activities have in common, their “greatest common factor”, that which connects those activities and at the same time lets them stand apart from previous interactions, what is often put forth is the “triadic character” of these interactions, sometimes also referred to as a “referential triangle” (e.g. Tomasello, 1999; Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979). This triangle is supposed to connect the activities of infant and caregiver to each other and to material culture such as objects, activities, and ideas. It is striking that otherwise very different accounts of child development share this emphasis on triadic interaction. Taking a closer look, “triadic interaction” is often part of a larger developmental narrative, which can be loosely summarized as follows (see Figure 1.1 for an illustration): at around 9 months, two lines of development, considered to be separate before that time – viz. dyadic infant-caregiver communication (predominant in the first months) and infant-object interaction (gaining predominance from 6 months) – come together in the emergence of triadic interactions as the infant becomes able to co-ordinate engagement with people and engagement with objects (Bakeman & Adamson, 1984). And only this ability in turn gives rise to conventional object understanding and use, language use and symbolic activities in general, as well as boosts forms of cultural learning such as learning by imitation or learning by instruction (Tomasello et al., 1993; Trevarthen & Hubley, 1978).
Figure 1.1: Illustration of the developmental narrative sketching a sudden onset of triadic interactions at the end of the first year around 9 months, when two lines of development considered separate until this point – dyadic infant-caregiver communication and dyadic infant-object interaction – come together, as the infant is for the first time credited with being able to co-ordinate engagement with people and objects (supposedly based on a newly emerging capacity of joint attention). And only this in turn gives rise to conventional object meaning and use, labelling, language use, symbol play, as well as cultural learning and co-operation.
1.2 Encountering the narrative: A first approximation.

Short versions of this narrative can readily be found in introductions and opening sections, interdisciplinary or popular accounts as well as textbooks, in short, places where – as pointed out by Ludwik Fleck (1936) – (probably simplified but) securely established parts of a field’s knowledge can be found:

Six-month-old infants interact dyadically with objects, grasping and manipulating them [...]. If people are around when they are manipulating objects, they mostly ignore them and they interact dyadically with other people, expressing emotions back-and-forth in a turn-taking sequence [...]. If objects are around when they are interacting with people, they mostly ignore them [...]. But at around 9-12 months of age a new set of behaviors begins to emerge that are not dyadic, like these early behaviors, but are triadic in the sense that they involve a coordination of their interactions with objects and people, resulting in a referential triangle of child, adult, and the object or event to which they share attention. (Tomasello, 1999)

Does this mean that infants before the age of nine months literally live in, or perhaps rather enact, two different realms – a lonely object world and an object-less social realm? Do objects only start to play a role in social interactions at the end of the first year as seems to be implied in this perhaps overly stringent version of the narrative?

Graphical illustrations often give a similar impression: while capturing aspects of proto-conversations (see section 2.3.1) depictions of early dyadic social interactions often suggest that early interactions are exclusively dyadic (see Figure 1.2), abstracting away the material and cultural context from the interaction.
**Figure 1.2**: Schematic depiction of dyadic interactions, often held to be representative for social interactions in general in early infancy (figure 5.1 in Moore, 2006, whose account of early infancy is, however, much more nuanced)

**Figure 1.3**: Momentary snapshot showing a 4 month old immersed in a rich cultural context involving people and objects, and following the motions of complex interchanges, as the mother, while holding and engaging with her infant at the end of nappy change, manages two more interaction tasks as they arise: answering her mobile phone and at the same time tending to her older daughter’s runny nose, with a doll and a nappy waste bag nearby telling of the ongoing background activities.
In everyday life, however, they rather look like figure 1.3, showing a 4 month-old infant immersed in a rich and complex culture-suffused everyday-life situation including other people as well as objects.

Clearly, infants are from the beginning immersed in a world of people and objects embedded in culturally shaped everyday life practices and routines, at least potentially giving infants plenty of opportunity for learning about culture, raising the questions of, first, how these routines are organised and practiced together, second, what infants take away from this immersion and joint practice, and, third, what enables infants to participate and take advantage of being immersed in and part of culture from early on.

1.3 “Triadic interaction” as “joint attention”

A central characteristic of triadic interaction is a shared focus of engagement. This is often operationalized in terms of (visual) “joint attention”, defined as two people, typically mother and infant, attending to the same object and knowing that they are both attending to the same object (e.g. Tomasello, 1995). This position takes its starting point mainly from research on gaze checking and gaze following (Carpenter, Nagell, Tomasello, Butterworth, & Moore, 1998; Scaife & Bruner, 1975) and accordingly narrows down the focus from Bakeman and Adamson’s (1984) relatively general phrasing “joint engagement” with people and objects to (visual) “attention” as the relevant aspect to look at (Moore & Dunham, 1995; or further differentiated variants thereof such as “joint perception”: Tomasello, Carpenter, Call, Behne, & Moll, 2005).

Moreover, it introduces the notion of “knowledge”. Behaviours that were repeatedly evoked in experimental studies, such as acts of gaze following are condensed to a “capacity” based on specific “knowledge”. By making this move, the approach leaves Adamson and Bakeman’s more descriptive stance and enters the space of theoretical claims, making the researcher turn away from looking at the process and instead letting him/her engage in the activity of asking for and giving explanations: What kind of “knowledge” is the capacity of joint attention based on?
At some point the tentatively constructed hypothetical bridge suggesting to explain joint attention by “understanding someone as an intentional agent” is crossed in the opposite direction turning “joint attention” itself into a criterion for “social understanding” (Tomasello, 1995), which is then used to evaluate the (“lack of”) social cognition abilities in autism (e.g. Mundy & Newell, 2007) and even rule out social cognition abilities in other species (e.g. Shepherd & Cappuccio, 2011), thereby implicitly declaring the hypothetical bridge for solid ground.

Conceptualized as a capacity “joint attention” then also makes its appearance as a pre-requisite for more complex forms of social interaction and cultural learning – including social referencing, learning by imitation, learning by instruction, symbol play, joint labelling, etc. - which therefore by definition become possible only after 9 months.

1.4 Counter-evidence

However, in addition to countervailing common-sense impressions from everyday life experience (see above figure 1.2), there are a number of naturalistic studies describing a gradual development of infant-caregiver-object interactions of incrementally increasing complexity over the first year, starting with the seminal studies of Hubley and Trevarthen (Hubley & Trevarthen, 1979; Trevarthen & Hubley, 1978) and Adamson and Bakeman (Adamson & Bakeman, 1984, 1985; Bakeman & Adamson, 1986, see section 2.3.2), as well as qualitative studies from cultural psychology (Moro & Rodríguez, 2004; Zukow-Goldring, 1997, 2012), and more recent ones (e.g. De Barbaro, Johnson, & Deák, 2013; Nomikou & Rohlfing, 2011).

Moreover, there is growing evidence from experimental studies that 1) infants show behaviours associated with cultural learning well before the “nine months revolution”: at six months, for example, infants are able to recognize everyday objects by their word labels above chance (Bergelson & Swingley, 2012). They can also be credited with some understanding of conventional object use, as they show predictive gaze to the goal positions of objects as they are used in everyday life, such as a cup put to the mouth or a cell phone put to the ear (Hunnius & Bekkering, 2010). 2) Moreover, aspects
of joint attention such as gaze following, capacities regarded as catalysts of cognitive development at the end of the first year, are present already as early as three months (Striano & Reid, 2006, 2009).

1.5 Aims of the thesis and chapter overview

This thesis seeks to take a fresh look at early – yet largely overlooked – everyday infant-caregiver-object interactions, both theoretically exploring how they are conceived of by different approaches in developmental psychology, and empirically looking into how they are organized and co-ordinated by caregivers and infants in daily praxis. Investigating how these processes of co-ordination change over the course of the first year we seek to capture and make tractable “culture in the making”.

Chapter overview:

Part I: A critical theoretical analysis of approaches investigating infant-caregiver-object interactions (chapter 2)

Chapter 2 “The development of triadic infant-caregiver-object interactions revisited” constitutes the first, theoretical part of the thesis, and engages in conceptual and historical analysis, critically assessing previous and current research programmes explicitly dealing with the development of infant-caregiver-object interactions. I compare and relate these diverse approaches in terms of their backgrounds, aims, conceptual frameworks, methods of data collection, analysis, as well as modes of explanation, and seek to analyse how these shape the key concepts and accounts advocated by the respective approach. The following approaches, grouped together into distinct types of accounts, are discussed:

I) “cognitive complexity” (Piaget, Bates, and Sugarman-Bell, followed by II) “coordination of separate object and person realms” (Hubley and Trevarthen, and Adamsom and Bakeman), III) cognitive, individual knowledge and capacity based accounts (exemplified by Tomasello & colleagues), IV) cultural accounts (Vygotsky, Rodríguez, and Zukow-Goldring as different examples), and V) other recent and interactive accounts.
Such an approach-analysis not only addresses 1) the current narrative about the beginnings of cultural learning in the emergence of triadic interactions around 9-12 months, its conceptual and empirical foundations, its different versions and historical development, but also, more generally, by locating this narrative within 2) a variety of different accounts of infant development, it directs our attention to their scope and limits, and to what they each have to offer for developing an approach suitable to address, investigate, and better understand the development of increasingly complex ways of co-ordinating social object interactions resulting in the co-creation of culture.

**Part II, An Empirical investigation into infant-caregiver-object interactions over the first year of life: a naturalistic, longitudinal study and examples of analysis of infant-caregiver-object interactions and activity contexts (chapters 3-4).**

For the second, empirical part, informed by and based on the theoretical part, a naturalistic longitudinal study was conducted, following infants and their caregivers through their everyday interactions over the course of the first year of life (from 3-12 months), focusing in particular on – yet under-researched - *infant-caregiver-object interactions and everyday (conventional) routines.*

**Chapters 3 and 4** represent examples of different forms of analysis: first looking at co-ordination in one model-activity “book sharing” in detail, and then extending the analysis to and comparing co-ordination across different activity contexts.

**Material and methods: Chapter 3,** since published as a stand-alone paper (see below), provides a self-contained account of our methodology including study design, data collection, exploration, and analysis (qualitative micro-analysis of multiple action strands combining ELAN and PRAAT) for a detailed investigation of one particular activity routine (see section 3.2). Building on this, **chapter 4** further develops key concepts and generalizes the methodology to a wider set of activities and activity contexts, including developing novel macro-measures to distinguish the structural
characteristics of different ecological activity contexts and to compare forms of co-ordination across these (see section 4.2).


From the 300+ hours of video material of diverse everyday activities - we here select book sharing, due to its prototypically cultural characteristics and its surprisingly early and widespread occurrence, as a first model activity for culturally shaped infant-caregiver-object interactions to investigate and gain first insights into how they are organized. Drawing on enactive cognitive science, conversation and interaction analysis, we use detailed qualitative micro-analyses and investigate how the interaction is co-ordinated as it unfolds, what the three poles – infant, caregiver, and object – each contribute to the multi-modal, multi-strand interaction, and how action-and-attention coordination develops over the course of the first year, as it becomes increasingly triadic.

In chapter 4 “Varieties of joint attention-and-action coordination in early everyday infant-caregiver-object activities” we then seek to extend the analysis from the example of book sharing - testing, generalizing, adapting the concepts and methods developed there - to a wider range of infants’ everyday social object interactions.

 Whereas the standard narrative of the development of triadic interaction rests on only a small number of research settings – mostly proto-conversation for early interactions and (staged) object play, labelling for later ones – and largely focuses on (visual) joint attention as the single crucial form of co-ordination, we here seek to investigate infant-caregiver-object-interactions across different ecological activity contexts, such as e.g. feeding, nappy change, “witnessing” caregiver’s chores, peekaboo, object exploration in baby gym etc. These differ in terms of their settings and activity structures and hence may afford, different forms of (multi-modal, multi-strand) attention-and-action-co-ordination.
As this constitutes a larger ongoing research project exceeding the scope of the current thesis chapter 4 focuses on:

- Developing concepts and a tentative method, in particular marco-analytic measures, allowing us to distinctively capture the structural characteristics of different ecological contexts and compare forms of co-ordination across these contexts.
- Documenting the process and results of macro-analysis in more detail for one examplar ecological activity context: infant-directed caregiving tasks involving objects.
- Conducting a detailed micro-analysis for nappy change as an example of a particularly rich type of infant-directed caregiving task, analysing and illustrating the complex temporal dynamic patterns of multi-modal, multi-strand co-ordination with distinctive functional infant participation, which is already seen in specific joint object routines from 3 months.

**Part III: Integration of empirical and theoretical parts**

The final part of the thesis seeks to integrate what has been gained from conceptual analysis in the theoretical part and what has been gained from empirical observational analysis and to reframe and reconceptualize the problems and challenges of the development of triadic interaction. **Chapters 5 and 6** explore through two examples – the development of understanding “self and other” and the development of participation, respectively – how this approach can be applied longitudinally and what it then can contribute to a better understanding of development.

**Chapter 5 “Structure and Openness in the Development of Self in Infancy”** is a spin-off paper, which takes the reader on an excursion to explore how a second-person approach in combination with the view of infant-caregiver-object interaction developed in this thesis may open up new perspectives on the development of “self” and “other”. It has been published as a stand-alone paper in the *Journal of Consciousness Studies* as
part of a dialogue about the “self” between psychological and philosophical approaches and spiritual traditions (Rossmanith & Reddy, 2016).

Triadic interaction, seen as related to the ability to co-ordinate people and objects, implying, more generally, the ability to take into account the relation between two things, has been ascribed an important role in the development of the self, either as opening the door to intersubjectivity proper, understanding self and other as “minds” with a hidden (albeit shareable) inner life (Stern 1984), or emphasising how over time the triangulation with caregiver and objects allows to reflectively see the self in relation to other people from the perspective of other people, leading to a more distanced, objectified understanding of self (e.g. Davidson, Fuchs).

Here, however – applying and testing the approach developed in this thesis: that triadic interaction is grounded in and constituted by a (developing) network of jointly created shared action structures – we look at self-and-other development in a novel way: and ask: what role do 1) participating in joint object routines, 2) jointly creating action structures and 3) these created action structures themselves – over time forming shared sense-and-action-scapes – play in the development of infants’ experience and understanding of “self”, “other”, and “together- or we-ness”?

In chapter 6 “The development of participation in infant-caregiver-object practices: changes in attention-and-action co-ordination”, we take up concepts and methods developed in the previous chapters and show how they can be used longitudinally to investigate and understand the development of triadic infant-caregiver-object interactions, here focusing on the example of how participation in social object routines develops.

We first revisit common rational and knowledge-based notions of participation, joint action, and co-operation and reframe them into a more embodied and situated version, emphasising co-ordination and jointly orienting in shared spaces of meaning and action. Using examples from our data and building on and extending the results of our analyses in part II, we then sketch a tentative trajectory of how participation in social object activities develops in terms of attention-action-coordination over the first year of life. We discuss the potential role of “jointly enacted affect-imbued
action arcs” later turning into “systemically structured shared sense-and-action-scapes” as key factors enabling infants to understand and participate in complex activities.

Finally in 7 Conclusions, general discussion and outlook, we briefly revisit and summarize the conclusions of the respective chapters and relate them to the literature, and compare and contrast them with the accounts of the development of triadic interaction revisited in chapter 2. We discuss the limits of the current study and – giving a brief outlook – sketch possible next steps to further develop research beyond those limits.

As we conceive of “thinking” as embodied and situated engaging with the world, affecting and being affected by it, and conceive of “research” as also including “conversation with materials” (Schön, 1983) – we consider this thesis as a collection of “thinking and working materials”, having resulted from engagement with empirical research and its products (such as literature and video data). They provide a palpable target for engagement in the process of developing an approach which allows to address, investigate, and better understand the development of joint co-ordination of infant-caregiver-object interactions.
Part I: A critical theoretical analysis of approaches investigating infant-caregiver-object interactions

2 The development of triadic infant-caregiver-object interactions revisited

2.1 Aims of this chapter

In the introduction (see section 1.2-3) we carved out a core narrative about the development of triadic interactions common to many accounts in developmental psychology: two separate lines of development – dyadic infant-caregiver communication and dyadic infant-object interaction – come together around 9 months in triadic interactions only then giving rise to cultural learning, symbol use, and co-operative participation. On closer inspection this framework of separate lines of development is based on a narrow and selective reading of the evidence from the literature, is at odds with some of the evidence from everyday life as well as naturalistic and experimental studies, and in particular when it posits a sudden emergence of a new “capacity”, actually prevents researchers from even taking a closer look at earlier periods and blocks the generation of new hypotheses or insights that would allow us to better understand the processes of development (see sections 1.4).

Therefore the aim of this chapter is to revisit accounts explicitly dealing with the early development of infant-caregiver-object interaction, the emergence of sharing a world, growing in and into culture, and, in particular, the narrative about the emergence of triadic interactions out of previously separated social and object domains, and investigate this narrative’s conceptual and empirical background. To do so, we will discuss in the following section a set of research programs dealing explicitly with the development of infant-caregiver-object interactions. As the approaches are very diverse, coming from different traditions, fields and eras, and as we seek to understand how the approach and methods used influence the results and accounts given (compare Bühler, 1927; Fleck, 1936), we will give
each approach more space than in the usual review section of a paper. We will give a short summary (mostly letting the approaches speak for themselves) and analysis concerning their backgrounds, aims, conceptual frameworks, methods of data collection and analysis, and modes of explanation for each approach to stand and to be understood in its own right. As we conceive of thinking as embodied and situated engaging with the world, and of research as also including “conversation with materials” (Schön, 1983) – we consider these condensed sketches as “thinking and working material”, for ourselves and perhaps also useful for others as they allow to literally put all accounts next to each other on a workspace to compare and work with. In this comparison we try not to (mis-)construe them merely as strawmen to set our own approach apart, nor to pretend that they all fit into one framework right away, contributing different “facts” in favour of or “falsifying” certain claims in one single framework. Rather we seek to let them stand in their own right, compare and relate them to each other to organize them, and then understand them through, their similarities and differences, and explore their respective scope and limits. We do so to explore what they can contribute to developing an approach suitable for investigating the development of infant-caregiver-object practices with regard to attention and action co-ordination.

A comparison of different accounts for the emergence of triadic interactions

2.2 Accounts I: General cognitive complexity and differentiation
(Piaget, Bates, Sugarman-Bell)

In this section we discuss approaches which set out to explain the developments over the first year and, in particular, the changes toward the end of it, in terms of and by linking them to processes of general intelligence becoming increasingly complex and differentiated (including in particular the development and understanding of instrumental means-end actions).
1) The starting point for our exploration of different accounts will be Jean Piaget. Even though Jean Piaget was not himself explicitly concerned with, nor explicitly addressed the development of infant-caregiver-object interactions, his approach was largely formative for developmental psychology, and his infamous “neglect” of early social interactions actually prompted several generations of researchers to explore the field of early infant social interaction. Piaget’s approach has explicitly or implicitly shaped their research frameworks (even as they sought to set themselves apart from it), so an appreciation of Piaget’s approach seems a necessary starting point to make sense of the development of research on infant-caregiver-object interactions.

We then move on to two examples among the pioneering language researchers who sought to investigate preverbal communication in the 1970s, and, in order to do so, drew on and extended Piaget’s framework into the social realm:

2) Elisabeth Bates and colleagues, combining the frameworks of Austin’s speech act theory and Piaget’s sensorimotor intelligence model to investigate the onset of “intentional communication” in form of proto-imperatives and proto-declaratives. And

3) Susan Sugarman-Bell, analysing the development of infants’ coordination of person and object-directed actions within social contexts in terms of action complexity.

2.2.1 Jean Piaget (1936, 1945): Co-ordinating secondary circular reactions in novel contexts as the first truly intentional actions marking the beginnings of intelligence.

2.2.1.1 Background, interest, and theoretical framework

Coming from Biology and inclined towards philosophical questions, Piaget was interested in the continuity between 1) the development of biological form in interaction with and adaptation to the environment and 2) the development of structures of “mind”, in particular complex systems of thought such as logic and mathematics. For our current context we can just
briefly note, that these are frequently seen both as “proto-typical” forms as well as pinnacles of thought, often closely linked to human culture, both as hallmarks and motors.

Given the difficulties in getting at the evolution of mind (which doesn’t fossilize), Piaget turned to ontogeny engaging in a kind of [behavioural] “embryology of intelligence” (Piaget, 1971, 1976; Newson & Newson, 1975)

2.2.1.2 Empirical investigation

After investigating the logical thinking of school children through conversation, Piaget in the 1920s and 1930s turned to explore the earliest forms of intelligence by observing and testing his own three children, seeking to demonstrate how these early forms of intelligence grow out of infants’ interactions with their (object) world (even before the onset of language) from which the infant constructs increasingly complex structures of mind.

2.2.1.3 Resulting account of intelligence emerging from sensorimotor interaction

In his resulting account of early sensorimotor development, *The origin of intelligence in the child* (Piaget, 1936), he distinguishes a sequence of cumulative, increasingly complex phases and/or levels of engaging with the world, resulting in increasingly complex mental structures:

As infants exercise basic reflex actions such as sucking (sensorimotor stage I), they soon seek to sustain the experience and extend the reflex activity beyond its primary function and immediate context for its own sake, e.g. from drinking to sucking the thumb (assimilating new contexts to the action), thus engaging in self-reinforcing activity dubbed “primary circular reactions” (sensorimotor stage II) by Piaget (following Baldwin and Wallon), leading to the development of their first acquired adaptations: habits. Through variation inherent in these extended actions they make new experiences and through seeking to sustain or recreate some of these experiences their action varieties differentiate (accommodating to these new experiences and contexts) into new distinct habits.
For Piaget the next qualitative change (and hence the beginning of **sensorimotor stage III**) occurs as infants from around 3 months extend these self-reinforcing circular reactions from sustaining and recreating interesting bodily experiences to sustaining and recreating interesting effects on objects (thus even further removed from basic bodily need satisfaction) – hence called **secondary circular reactions**. Even though the secondary circular reactions may already provide the infant with a broad repertoire of actions, they still are habits, accidentally discovered and then sustained, and are only adaptively used in specific familiar contexts where they are applied repetitively – thus still distinct from intelligence considered as an “adaptation to a novel situation”.

In contrast, consider the following situation: an infant, who is confronted with a pillow blocking his reach to a desired object – rather than unsuccessfully repeating his reaching schemes toward the object – first pushes down the pillow, and then reaches for the object. Seeking to trace the emergence of intelligence, these first instrumental means-end actions, which he observed in his children around 8-9 months, are of major importance to Piaget and for him constitute a major developmental milestone (labelled **sensorimotor stage IV**), as they mark the first forms of **a) intentional action** and **b) intelligence proper** and are closely linked to the first forms of “representation” and the central concept of “object permanence”, as well as the processes of “objectification” and “spatialization” of causality (Piaget, 1955)

Thus already Piaget speaks of a major qualitative shift occurring in the period around 8-9 months of age – at the same point in time and hence equivalent to what the modern developmental narratives refer to as “the quantum leap”, “the socio-cognitive revolution”, etc. However, as Piaget was not concerned with triadic infant-caregiver-object interaction – indeed was at this point not concerned with social interactions at all, he framed these changes in a different way than “dyadic interaction converging into triadic ones” and valued them, as mentioned above, for different reasons: he conceived of them as **a) the first truly intentional actions and valued them as** **b) the first manifestations of intelligence proper.**
Let us take a closer look: The action described above counts as a demonstration of true **intelligence**, insofar as it constitutes an “**adaptation to a novel situation**” (the object of desire is suddenly blocked), newly combining two **familiar actions** (pushing down pillow and reaching for object) **in a novel way** and subjecting them to an overarching new goal.

Conversely, as Piaget was struggling with existing definitions of “**intention**” which were either too narrow (considered as actions controlled by language related representations) or too vague (considered as actions controlled by any conscious awareness or expectation), he sought to define the notion of intentionality from the organism's perspective: he speaks of an **organism having an intention when it is able to separate out a goal from the action and keep it in mind**, persistently pursuing it even in the face of obstacles.

For Piaget, this ability develops out of secondary circular reactions in the following way: As a specific goal arises for the organism, e.g. the tendency to grasp or move a specific object, but in the context of a new, unfamiliar situation (e.g. the object is blocked by an obstacle), the holistic secondary circular reaction – containing as an undifferentiated unit the arising goal and the grasping action realizing it – cannot proceed as usual, but is broken up. Only now, as the goal cannot be reached immediately, an organism can begin to distinguish and hence perceive desires or goals (as) separate from the means to realize them.

As much as Piaget stresses the perspective of the organism and proposes a story about how the organism may learn to distinguish and hold an intention, taking a closer look, this specific definition of intentionality seems above all a **definition that allows the investigating observer to attribute the existence of a separate goal to another organism without doubt**. As for an observer an action unit is here made visible and defined through behaviour directed at an object as its visible goal, two objects are needed – with one serving as an obstacle – to make an intention visible, resulting in the setup described above: the infant, instead of directly attempting to grasp the desired object, acts on the second object as she would usually do, but not for its own sake as she removes rather than attains the object and does not
confine herself to the first action, but continues to combine it with a second one, now grasping for the first – apparently desired – object, allowing to infer that grasping the first object had been the intention singled out and kept in mind all along.

However, in these first instrumental means-end actions the infant only uses familiar action schemas and directly acts on the objects as he/she usually does: first pushing down the pillow, then grasping the object. It takes another 2-3 months until the infant from around 12 months on, starts systematically varying object schemas for experimentation (stage V: tertiary circular reactions), and through such experimentation finds novel means to specific goals, e.g. using a stick not for banging on an object, but as a means for fetching an out of reach object.

Piaget's characterization of the changes occurring at 8-9 months as the onset of intentionality in the form of instrumental means-end actions (stage IV), as well as his characterization of more complex tool-mediated instrumental actions (stage V) had great impact on what came to be the modern developmental narrative about the development of triadicity. It provided the frame for generations of researchers to work from, either building on Piaget explicitly, such as Bates and Sugarman (see below) as well as Tomasello (section 2.4), or seeking to distance themselves from his framework (e.g. Trevarthen, see section 2.3.1). To arrive at modern versions of the narrative, much has been added to this frame, both methodologically as well as theoretically, including – crucially – sociality, as will be explored in the next sections.

However, for Piaget this was not the end of an infant's journey from sensorimotor interactions to thinking. According to him, it takes yet another couple of months until (from around 16 months) the infant can effectively solve problems using tools in a novel situation without engaging in overt experimentation but by thought and insight alone (e.g. using a nearby stick to fetch an out of reach object without ever having done so before). Thus, with this invention via imagination the infant now not only has exercised the earliest forms of true internal thinking (and has thus achieved Stage VI: The invention of new means through mental
combination), but also the ground has been laid, in principle, to contribute – in combination with language – to culture creation.

More about the development of the foundations of language – symbol use – can be found in Piaget’s “Play, Dreams and Imitation in Childhood” (1962), where he lays out the development of symbol play – along with the development of imitation, another ability considered crucial for culture – again both seen as more or less solitary processes rooted exclusively in the individual, emerging out of the individual children’s interaction with their world. According to Piaget, symbol or pretend play develops when infants – exercising their repertoire – begin to perform various actions on inappropriate objects repeatedly and in increasingly ritualised form – thus not serving any functional adaptation but apparently done merely for joy (e.g. pretending to sleep by using a cloth, piece of coat, or tail of a donkey as a pillow and laughing) and by doing so start to separate actions from their context and using objects as “symbols” standing for something else.

2.2.1.4 Piaget’s individualist account and “neglect” of the social

Thus Piaget has accounted for several abilities often taken to be “pillars” of culture, such as instrumental reasoning, imitation, and symbol use, exclusively in terms of an active individual’s cognitive or mental structures developing from autonomous interactions with the object-world, without ascribing any major role to – or even taking into account – the infant’s social world and interactions.

“The social” for Piaget only seems to play a role for acquiring certain cultural-specific content (such as specific language signs, words, which for Piaget are based on arbitrary social conventions not discoverable for infants on their own, as well as morals), rather than for developing most of the basic mental structures (with the exception of some aspects of logic, Lourenço & Machado, 1996). He even at times seems to consider social interaction as potentially harmful: for research, as it might conceal what an infant is “truly” able to do on his/her own, and for development in general, distorting, or
interfering with natural development (see e.g. his comments when setting apart true imitation from pseudo-imitation in Piaget, 1945, 1962)¹.

Piaget’s “neglect of the social” has famously prompted much criticism: Bruner’s comments from his autobiography, neatly compiled by Adamson, paint the world of the Piagetian infant under the exclusion of sociality in gloomy colours:

To Bruner, Piaget’s child appeared to be “a calm and a lone one” (p. 139). Bruner also saw the world as “a quiet place” where “he is virtually alone” in "a world of objects that he must array in space, time and causal relationships. He begins his journey egocentrically and must impose properties on the world that will eventually be shared with others. But others give him little help. The social reciprocity of infant and mother plays a very small role in Piaget’s account of development. And language gives neither hints nor even a means of unraveling the puzzles of the world to which language applies. (J. S. Bruner, 1983, p. 138; Adamson, 1995)

Trevarthen and Hubley criticise Piaget for reducing infants to their “object motive” and denying them the “social motive” altogether:

Indeed, Piaget presents himself as a responsive but invisible examiner of the developing imagination and reason of the child. Even studying imitation and play, he makes his analysis in terms of a “thinking” that allows the infant no special awareness of humans as persons having a unique potentiality for shared awareness and shared intention. (Trevarthen & Hubley, 1978)

¹[...] nous avons précisément pris grand soin, dès le début de nos observations, d’éliminer dans la mesure du possible l’intervention du dressage. C’est pourquoi, soustraits à certain influences adultes (jeux suggérés, etc.) et en particulier à la manie pédagogique des nurses, nos trois sujets ont présenté un progrès dans l’imitation beaucoup plus lent et plus régulier que se n’est le cas chez les bébé sans cesse déformés par leur entourage. (Piaget, 1945, p. 25, quoting the French original here, since the English 1962 translation seems to gloss over this point.)
Reddy (2008) describes and analyses how Piaget— even when describing affect-laden social interaction— astonishingly manages to describe the situation individually in terms of the infant's sensory interests, blinding out the caregiver, her affective response, and any relationality, and himself remains unengaged as an observer:

Although his observations are beautiful and reveal a powerful empathy for children's intentions in relation to the physical world, he was strangely unengaged with their relations to their social world. Take the following powerful description of Jacqueline at 10 months: OBS 63. At O:10(3) J. put her nose close to her mother’s cheek and then pressed against it, which forced her to breathe much more loudly. This phenomenon at once interested her, but instead of merely repeating it or varying it so as to investigate it, she quickly complicated it for the fun of it; she drew back an inch or two, screwed up her nose, sniffed and breathed out alternately very hard (as if she were blowing her nose), then again thrust her nose against her mother’s cheek, laughing heartily. These actions were repeated at least once a day for more than a month, as a ritual. (Piaget, 1951/1972, p. 94)

Piaget’s observation focuses on the individual actions and sensory interests. They do not include reactions from others or from himself (even if they had occurred they may have been seen as irrelevant to this phenomenon). In this particular instance, however, it is hard to believe that the mother, whose cheek was being rubbed and breathed into, and whose ears were filled with a 10-month-old daughter's hearty laughter, did not react at all. [...] His observation of Jacqueline, although containing some acute perceptions of her intentionality, curiosity and playfulness, nonetheless portrays her as separated from and unengaged with himself (Reddy, 2008, p. 37)

Analogously to the example Reddy picked for social interactions there are similar description of social object interactions, which might well be considered as social object games as described by Trevarthen and Hubley (1978).

OBS 61. At 0 ; 7 (13), after learning to remove an obstacle to gain his objective, T. began to enjoy this kind of exercise. When several times in succession I put my hand or a piece of cardboard between him and the toy
he desired, he reached the stage of momentarily forgetting the toy and pushed aside the obstacle, bursting into laughter (Piaget, 1962, p. 92).

Here one cannot but wonder whether the funny effect is simply due to the object, or whether the interaction is experienced as sharing something funny, playing with intentions and expectations in a social game where both self and other may be felt vividly as the other’s actions are highly attuned to, and hence can be experienced as responses to and anticipations of, the infant’s action and indeed her intention.

Seeking to put Piaget’s apparent neglect of the social into context, Lorenço and Machado point out in their “Defense of Piaget”, that Piaget did not simply neglect the social but that the degree of importance he ascribed to it with regard to his respective questions changed in different phases of his work:

_During the initial and functionalist phase of his studies, Piaget (1923, 1932) considered social interaction the main factor responsible for the transition from egocentric to socialized thinking and gave a purely social explanation of cognitive structures. Later, when he found a sensorimotor intelligence and logic before the emergence of verbal language (Piaget, 1936, 1937), Piaget (1976a) confessed that in his initial phase he had overestimated the role of language and social interaction in the construction of knowledge. He then moved to a strongly structuralist phase and pursued the idea that cognitive structures and operations come from the subject's own coordination and self-regulation of his or her actions (Lourenço & Machado, 1996)^2_  

^2 Another factor pointed out by Lourenço & Machado as a probably contributing to Piaget's neglect of the social is that Piaget wanted to trace the origins of necessary knowledge and got particularly interested in the role played by reflective abstraction as the main source for mathematical knowledge and accounting for the construction of new knowledge. And reflective abstraction for Piaget does not originate from interaction with the world but from the co-ordination of actions themselves – hence it is something the infant supposedly can only construct from within him/herself
Thus (striking a similar chord) Chapman, rather than criticising a neglect of social interaction instead points out a lack of integration:

Although Piaget recognized for both operative [object directed] and communicative [people directed] forms of interaction in various phases of his work, he never integrated those components in a single model. In particular, communicative interaction was addressed in some detail in the 1920s and was relatively neglected in his work on operational thought after 1940 (Chapman, 1991).

In other words, Chapman criticises that Piaget only looks at dyadic interaction and not at triadic interaction, whereas according to Chapman it would be crucial to look at the “epistemic triangle” of triadic interaction in order to understand any form of knowledge construction.

2.2.1.5 The productive impact of Piaget’s “neglect of the social” on following research

Besides wide criticism Piaget’s “neglect of the social” above all sparked a wide variety of research responses:

a) One line of researchers sought to “complement” Piaget by simply adding “social interaction” as one more, yet overlooked research area. They basically maintained Piaget’s theory and as their main task engaged in exploring how Piaget’s theory could be slightly modified and applied to this new area, spelling out how infants exercise their developing general intelligence skills and cognitive structures in social interactions, this specific and making it futile to look at anything (any “content”) coming from the outer world including social partners. However it might well be the establishment of joint action routines, carefully structured and held stable by caregivers, and the joint modification of these jointly created structures, that provides a space for the first reflective abstraction to occur. This possibility will be taken up again and further developed based on evidence from our own naturalistic longitudinal study on everyday infant-caregiver routines in chapters 5 and 6.
not yet investigated part of their lives (Rivièrè & Coll, 1987; Chapman, 1991, Bates and Sugarman see sections 2.2.2. & 2.2.3).

b) Other researchers however insisted that social interactions with people provided such a different context for infants, and that indeed infants showed such different actions in social contexts than in object contexts that any account of infants’ minds based on object interactions remains fundamentally insufficient and we need to assume, investigate and describe multiple different systems through which infants fuel and guide their interactions with different aspects of the world. (e.g. Trevarthen, see section 2.3.1)

c) While yet another, third group goes even further, suggesting that the general focus, unit of analysis, and direction of explanation need to be fundamentally changed and extended beyond the infant, as the processes and abilities in question do not exclusively originate in nor are restricted to the infant(‘s mind), but only emerge collaboratively in interaction with the partners affecting each other in a process of mutual transformation (e.g. Newson & Newson, 1975)

In the further course of this chapter we will again and again encounter researchers whose approach to triadic interaction can at least partly be traced back to these different ways of responding to Piaget, which can hence serve as a rough map to what we are going to encounter in the course of the chapter ahead.

We will meet representatives of the first strategy a) extend Piaget to the social in the remainder of this section: first Elizabeth Bates and colleagues (2.2.2) and Sugarman-Bell (2.2.3), who applied Piaget’s framework to the social realm developing the concept of “social tool use” and – when running into the limits of Piaget’s framework – started to move beyond it. Second, we will meet its later modified echoes in the section about the cognitivist approach of Tomasello and colleagues (2.4), an account in which Piagetian “intention understanding” plays a crucial role and is complemented by a specific form of “social knowledge”.

39
We will encounter the second strategy \textit{b) distinguish 2 separate systems/domains} in the following section when discussing Trevarthen and Hubley (2.3.1), who endow the infant from the beginning with the special social ability of “(primary) intersubjectivity”, over time integrated with object skills resulting in “secondary intersubjectivity”, as well as Adamson and Bakeman (2.3.2) for whom bridging what they consider separate domains of people and object engagement constitutes a major challenge of development to be slowly overcome and finally mastered in “coordinated joint engagement” before infants can start on language learning.

And we will recognize the third strategy \textit{c) “zoom out and reconceptualise social interactions as a systemic process of mutual affecting and transformation”} pursued in various different forms in Section 2.5 on cultural psychology, as well as in Section 2.6 discussing more recent embodiment, situatedness and 2\textsuperscript{nd} person approaches.

\textbf{2.2.2 Bates and colleagues (1975-79): Communicative intentionality as social tool use: proto-imperatives and proto-declaratives}

\textbf{2.2.2.1 Background, interest, and theoretical framework}

The collaborating researchers Elizabeth Bates, Laura Benigni, Inge Bretherton, Luigia Camaioni, and Virginia Volterra were among the pioneers of language researchers who, at the beginning of the 1970s, got increasingly aware of and interested in communication processes prior to language, their development, and – with an eye on identifying potential pre-requisites of language – their interdependence with other aspects of development, such as general cognition and sociality, considered as separate systems (or in the
then emerging Cognitive Science vernacular “software packages”\(^3\)). They chose to explore the onset of “intentional communication”, later defined as “signalling behaviour in which the sender is aware a priori of the effect that a signal will have on his listener, and he persists in that behaviour until the effect is obtained or failure is clearly indicated.” (Bates et al., 1979)

Starting from this concept created a point of entry into the then (theoretically and methodologically) mostly uncharted territory of pre-verbal communication, as the concept lies at an intersection of, and hence can bridge two hitherto separate research frameworks: Jean Piaget’s account of sensorimotor development and James Austin’s speech act theory.

Piaget’s theory (sketched above) presents the earliest forms of general intelligence as developing out of pre-verbal infants’ sensorimotor interactions with the object-world, and the occurrence of means-end actions around 9 months serves as a visible indicator for the emergence of “intentional” actions marking the onset of intelligence proper.

Austin, after struggling with verbal utterances whose meaning largely eludes classic semantic analysis (like the “I do” in a marriage ceremony), proposed that their meaning was better captured by considering them as performing a specific action (hence dubbed “performatives”) and went on to generally addressing language in terms of action, as “performing speech acts” (Austin, 1962). This shift to action had made his theory interesting to a number of infant researchers (J. S. Bruner, 1975; Dore, 1978; Halliday, 1975) as it made pre-verbal, action based communication addressable – in analogy to verbal utterances – as non-verbal speech acts, more particularly, performatives.

\(^3\) “some sort of software package, a ‘program’ that an individual child or adult ‘has’, which permits generation of behaviors that are externally identifiable as ‘linguistic’, ‘cognitive’ or ‘social’” (Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979)
Thus, both theories blended together opened up a new research framework, where early pre-verbal communication could be addressed together with early general intelligence and later emerging language to elucidate development.

Additionally Bates et al. adapted Austin’s distinction between multiple aspects or types of speech acts, which an utterance could embody (often (all) at the same time): locutions (verbal utterances, strings of words with specific references), illocutions (fulfilling specific conventional actions recognized by participants), and perlocutions (effects on the listener). Bates and colleagues sequentialized the different kinds of speech acts and projected them onto development, suggesting the following developmental trajectory for the development of communication:

1) “a perlocutionary stage, in which the child has a systematic effect on his listener without having an intentional, aware control over that effect”;  
2) “an illocutionary stage, in which the child intentionally uses non-verbal signals to convey requests and to direct adult attention to objects and events”; and  
3) “a locutionary stage, in which the child constructs propositions and utters speech sounds within the same performative sequences that he previously expressed nonverbally” (Bates, Camaioni, & Volterra, 1975)

This framework now a) allows stating that there is meaningful communication from birth, though still unintentional, and b) offers a direct point of attack to investigate the onset of “intentional communication”: Bates et al. simply look for the emergence of what they call non-verbal performatives as markers for intentional communication, in particular for the early pre-verbal forms of the “two most general forms of performatives”:  
1) proto-imperatives: using a person as a means to reach some goal – which they quickly narrow down to: “a desired object”, arguably for pragmatic reasons of observability – and 2) proto-declaratives: using an object to attain a person’s attention (e.g. showing, pointing). These they then
sought to relate to the occurrence of markers of social and cognitive development (Bates et al., 1975).

2.2.2.2 Empirical study (Bates et al., 1975)

For a pilot they visited 3 children at home once or twice a month during different yet overlapping age periods (starting from 2, 6, and 12 months) and observed (but only partly videotaped due to high costs) the infants for approximately 2 hours in spontaneous communicative interaction as well as planned interventions for a) communicative behaviour (using their own categories, see below), b) cognitive development (object permanence and means-end tests following Piaget), and c) social development (attachment tests following Bowlby and Spitz).

For their qualitative analysis they transcribed the videos into written text and subsequently analysed the text for the occurrence of predefined key behaviours, which were then sorted into the respective columns: communicative (proto-imperative, proto-declarative), social (smiling, following mother, crying at stranger) and cognitive (object permanence, spatial displacements, causality, and imitation [note: not listed under social, probably following Piaget]). They then compared the onsets of these key behaviours across columns.

2.2.2.3 Results and Interpretation

Only at around 10 months they find the first proto-imperatives and proto-declaratives. For the first months they report infants smiling, crying and parents “responding to satisfy their needs” but state to “find no evidence for intentional communication before 9 months”. Rather – following psychoanalytic theory – they interpret infants’ early communicative acts as “merely a built-in reaction to a particular internal state”. “Prior to that time [9 months], communication certainly takes place. The infant cries, or reaches toward his goal, and the adult interprets the child’s desires and intervenes to meet them (1979, p. 34). This adult response to the infant’s signals establishes a circular means-end relationship which is the first step in the development of communicative intentions. But there is no evidence that the child himself is aware a priori of the signal value of his smiles and cries.” When Bates and
colleagues make the claim that communication prior to 9 months is automatic and inner-state-induced – probably in order to further conceptually carve out the difference to later intentional communication – this dichotomous conception is apparently shaped by another classic figure of thought dealing with “intentionality” of “goal directedness” in a different context: Aristotle’s distinction between “efficient cause/causality” vs. “final cause/teleology”, which constitute two of his “four causes”. But while for Aristotle these were two out of four different causes for one phenomenon, in other words, two different kinds or aspects of explanation, here they become reified into two different kinds of causation, supposed to distinguish two different phenomena: pre-intentional from intentional communication: “Prior to 9 months communication is efficiently caused but not finally caused” (Bates et al., 1975).

Even though they acknowledge that parents do indeed “treat the infant’s expressions from early on as deliberate products of a conscious agent”, and even though they consider many of already early observed behaviours towards objects as well as towards people as showing signs of intentionality before 9 months, e.g. persistently using a variety of actions to grasp the mother’s face (suggesting a less strict use of the term than Piaget’s) – they still decidedly rule out any "communicative intentionality” before 9 months. Why is that and what makes them so certain? A possible answer is offered by a sentence found in their earlier paper, which gives valuable insight into the scientific process, frankly expressing the struggle to shape a new phenomenon into something accountable for (in familiar terms):

However if they [infants < 9m] were indeed capable of intentional communication at this stage, we would be unable to explain their apparent inability to look toward adults and invoke their help in later attempts to reach objects (Bates et al., 1975).

One puzzling situation reported by Bates and colleagues can serve as an example for the kind of interactions referred to in the above sentence: a roughly 9 month old infant “in an effort to attain a box from her mother’s arms” is “pulling at the arms, pushing her whole body against the floor, and
approaching the box from several angles”, without however ever looking up at her mother’s face.

Whereas the (lack of) communicative actions exhibited by the infant in the example above will clearly seem striking to an adult observer, the way the above argument is posed takes the form of an argument from the impossibility to think otherwise or to imagine alternative explanations, implicitly making and holding on to a number of tacit assumptions about what actually might be framed as a different problem, than the one assumed.

Bates and colleagues go on to account for this situation and the assumed inability to communicate intentionally by concluding (in retrospect) that roughly for the first 9 months “schemes for interacting with adults and schemes for interacting with non-social objects are apparently kept separate.” (Bates et al., 1979, p. 212). Thus they propose a separation of people and object interactions for the first 9 months as it is found in today’s narrative.

However, different from today’s version of the separation, Bates and colleagues made this statement in the context of the Piagetian instrumental action framework which they take as a model for communication: under its terms it is the more general challenge for infants in all instrumental thinking to combine two hitherto separate action schemas. And this problem seems indeed to be the same in principle as well for both proto-imperatives and proto-declaratives to occur. In the context of Bates and colleagues’ specific setup and their research question concerning “communication with people about objects” the two actions to be combined just coincidentally happen to be actions towards objects and people, respectively.

Thus, having framed (intentional) communication in terms of Piagetian (compositional) instrumental action, the occurrence of proto-performatives only after 9 months is exactly in line with what would be predicted by the Piagetian framework, stating that 9 months is the time any two motor schemes start being combined into compositional instrumental actions.
And more than that, taking a closer look at proto-imperatives and proto-declaratives in terms of the Piagetian framework, they actually seem to go well beyond the above described stage IV instrumental acts, combining two familiar motor schemes (where infants still directly engage in familiar motor schemes as they normally would, only chaining them together sequentially in novel ways). Rather, proto-performatives can be considered to constitute a more indirect, truly tool-mediated instrumental act, analogous to Piaget’s example of fetching an object with a stick, characterizing stage V “tertiary circular reactions”, where action targets are not acted on in the familiar way anymore, but in novel and sometimes indirect ways, which, according to Piaget, only become possible through new stage-V-type co-ordinations: employing novel means discovered through systematic experimentation, (compare section 2.2.1: Piaget).

And indeed, it was around the (same) time, 10 months, (when) infants tested positively for pulling a cloth to attain a desired object standing on the cloth (qualifying for Piaget’s stage V), that Bates and colleagues reported observations of proto-imperatives and proto-declaratives, and at 11 months, in a situation similar to the one described above, the infant now “looked the adult intently in the face” before again taking up her efforts towards the object.

Thus the Piagetian framework seems sufficient to account for the development of proto-performatives, of “social tool use” along with “tool-use” in general. And while Bates and colleagues go along with this in their later book, in the earlier paper they seem to go one step further and stress the differences between people and object interaction schemes making the gulf between them larger and harder to be bridged.

The difference between the use of a simple instrument as an intermediate means, and the use of an adult, is that the child himself is not longer the agent. In appealing to the other as agent, the child must pass through an indirect causality, a means that is not under his direct control, and whose results may be delayed and unpredictable. The adult-tool is hence quite different from such malleable instruments as cloth-supports or sticks and handles. (Bates et al., 1975).
This gulf between people and objects will grow still larger later on, either or implicitly assumed from the start as e.g. in Adamson and Bakeman (section 2.3.2), or explicitly with Trevarthen’s notion of two entirely separate “physiological action systems” or “motives” for people and objects (see 2.3.1).

Also in other cases regarding social aspects of interaction, Bates and colleagues feel they have to move beyond Piaget: First, they newly describe proto-declaratives, which they first tried to model after a Piagetian stage framework but eventually found hard to account for exclusively in Piagetian terms. They see the pre-stages of proto-declaratives in showing off, which they interpret as Piagetian circular reactions, as the infant seeks to prolong and re-create the adult’s affective response (e.g. laughter) by repeating the behaviour. When infants begin to hold up (“show”) objects, Bates and colleagues consider this to be the onset of proto-declaratives proper, interpreted as deliberately using novel means (showing the object) to achieve a familiar end (amusing the adult and receiving attention). With time, infants begin to give the object to the adult, and later still they begin to point, i.e. engage in the prototypical proto-declarative. As they cannot account for this development in Piagetian terms, they turn to Werner & Kaplan’s concept of “distancing” where infants beginning from an undifferentiated primordial sharing situation increasingly differentiate between self, other, and object (see Werner & Kaplan 1963). While “distancing” is also seen as a general development, it is here understood quite “literally”, as showing off, showing an object at arms length, etc., to finally pointing at distal objects “carry the inclusion and progressive distancing” of objects “beyond the length of the child’s arm.”
2.2.3 Sugarman-Bell (1978): A succession of overlapping waves moving from simple to complex single person or object acts (and) finally to co-ordinated person and object acts.

2.2.3.1 Background, interest, and theoretical framework

Sugarman-Bell’s pre-doctoral research conducted at the beginning of the 1970s is similar to the research of Bates and colleagues – rooted in the growing awareness of the time that infants express meaning prior to speech and language through non-verbal action. In order to investigate the development of preverbal communication the classic referential language acquisition scenario of “labelling or communicating to people about objects” is again translated into a preverbal, behavioural analogue of infants conveying their mothers their intentions towards objects, which provide visible targets of behaviour to the observer.

However, rather than looking for specific communicative acts as Bates and colleagues do, Sugarman puts the question in slightly more general terms:

*How do infants organize and co-ordinate their actions and vocalisations towards people and towards objects within social contexts and how does this change over time?”* (Sugarman-Bell, 1978, adapted, merging two paragraphs)

She adopts a slightly more general analytic stance, devising quantitative measures for investigating the general “complexity” of the actions and how it changes over time, albeit restricted to small pilot study.

Sugarman takes as an impetus the same situation Bates and colleagues were wondering about: as an object was held on the mother’s lap, younger infant A kept trying to pull the object off the mother’s lap without ever looking at her face, whereas older infant B first touched and made eye contact with the mother and only then grasped for the object. Sugarman isolates two aspects, along which A’s actions (only directly pursuing his/her object intention) and B’s actions (making B’s intentions towards the object clear to the adult and then reaching for the object) can be differentiated:
1) The **complexity** of actions in terms of number of different sub-actions employed, and

2) the **orientations** towards, or “targets” of, the action: object and/or mother.

(This for our purposes could also be called: “dimensionality of action”, see chapter 3)

As she built her study around these two aspects, complexity and orientation, these in turn became the relevant aspects in terms of which the development of pre-verbal communication was accounted for.

Extending Piaget’s and Duncker’s findings about physical object interactions, that simple actions directly affecting an object (secondary circular reactions) precede more complex ones (means-end actions), Sugarman predicts the following sequence for social object interactions: 1) Simple actions within each of the two orientations (towards people and objects, respectively) – which hence remain mutually exclusive. 2) Increasingly complex, differentiated actions within each of the orientations – still kept separate. 3) Action sequences integrating both orientations (Sugarman-Bell, 1978, p. 50).

**2.2.3.2 Empirical study**

In an exploratory study 8 infants were visited at their homes once a month for 5 months (4 from 4/5 – 9/10 m and 4 from 8/9.5 – 12/13 m) following mother and infants through their everyday routines and recording social interactions with a pen and paper coding system for about 1 hour at each visit. Social interactions were defined as: a) 2 people being active and b) at least one participant at least partially orienting toward the other.

Finally, the Uzgiris and Hunt (1966) “Means for Achieving Desired Events” and “The Development of Causality” scales were administered (Sugarman-Bell, 1978).

**2.2.3.3 Analysis**

The mother-infant social interactions were coded in terms of 1) single orientations toward object or person, which were further divided into simple vs. complex single orientations (paralleling Piaget’s distinction
between unitary secondary circular reactions (stage III) and means-end actions (stage IV), and 2) co-ordinated people-object orientations, resulting in the following coding scheme:

Single simple orientation (1, 2):

1) Simple person-oriented acts: one or two types of behaviour directed at person (e.g. adult looks at and vocalizes to child – child smiles at adult).

2) Simple object oriented acts: child manipulates or turns attention to object, person only involved incidentally (e.g. adult holds object over child’s cot – child beats at object without making direct social contact ...)

Complex single orientation (3, 4):

3) Complex person-oriented acts: approaches or responses to a person involving a combination of different behaviours, e.g. child looks, smiles at adult, touches, tugs at adult clothing – adult lifts child, reciprocal imitation

4) Complex object-oriented acts: different discrete behaviours are combined and directed at external object, e.g. adult gives child box and lid – child takes object without acknowledging adult – child puts lid on box.

Co-ordinated person-object orientation (5):

5) the child marks his objective regarding the object, and approaches the person involved socially (e.g. child holds jar toward adult, looks at adult and vocalizes – adult takes jar from child, vocalizes to child; Child vocalizes to adult, reaches toward chair, adult places child in chair.

2.2.3.4 Results

Results showed a 3 step or rather wave like succession of action types, where the emergence of a new type of engagement co-occurred with or immediately followed a sharp rise of the previous one.
1) simple single orientation sequences towards people or objects each present since beginning of the study and peaking between 4.5 and 7 months,

2) complex single orientation sequences towards people or objects emerging between 4 and 7 months and peaking around 10 months, and

3) co-ordinated orientation interaction emerging between 8 and 10 months.

While remaining separate, single person acts and single object acts developed in parallel through the first months with changes occurring in the same order and the same time: simple single person acts peaked at the same time as secondary circular reactions with objects (sensorimotor stage III), later means-end actions on objects (sensorimotor stage IV: combining familiar means), e.g. pushing away a cushion to get to object (but not yet using an object to get to another object) co-occurred with a peak in complex single person acts, combining multiple distinct actions (e.g. tugging at clothing and reaching up towards adult’s face) as well as the emergence of multi-step play routines. Despite the complexity of complex single object acts occurring in social context social bids for objects were not observed until later, when subjects began to subordinate the use of one (physical) object to actually manoeuvring a second (e.g. using a stick to retrieve an object, sensorimotor stage V) suggesting that they first had to learn to – rather than directly acting on the desired object – to set in motion either instruments or agents as a means to indirectly attain the object.

In addition to the increasing hierarchical differentiation of instrumental thinking and acting structures in general Sugarman, similar to Bates and colleagues, notes another development within instrumental actions: infants increasingly differentiated the way they approached objects (through manual manipulation) and the way they approached people (through social cueing) in instrumental acts, with Sugarman reporting transitory in-between states where infants were directly acting on parts of the adult (e.g. hands) to make them work as instruments – a phenomenon already discussed by Piaget and later often restricted to discussions of children with autism and great apes, as an argument for illustrating their supposed lack of social understanding (but see Gomez, 2004). Sugarman notes that: “Subjects’
social-interactive manoeuvres developed not only toward the elaboration of instrumental relationships but also toward the differentiation of another person as an agent in those relationships” (Sugarman-Bell, 1978, p. 61).

Here the object-person dichotomy already built into the framing of research on preverbal communication in analogy to language – now makes visible the infants’ different responses to people and objects. Here we already see the seeds for postulating a priori a gulf between people and objects needing to be bridged as done in later accounts. Though here, in contrast to more recent research, the general differentiation between communication and manipulation seems to be considered a rather late development:

[I]t seems reasonable to speculate that while an instrumental basis evolves for social transactions, the child eventually distinguishes communication from manipulations, certainly by the time he begins to talk (Sugarman-Bell, 1978)

Interestingly, the formulation “the differentiation of another person as an agent” already foreshadows Tomasello’s account (first take, see section 2.4.1), where the development of specifically social person-knowledge: “understanding people as intentional agents”, rather that an increasing general ability to deal with (and enact) increasing complexity, will be considered the crucial factor (to account for) for the ability to share a world in triadic interactions to develop.

Of particular interest for our purpose of understanding the development of triadic interaction are in this study 1) the efforts to systematically capture the complexity of actions, 2) the suggestion of a wavelike succession of increasingly complex action forms, but also 3) which aspects were explicitly excluded as “unscorable” with the coding scheme applied: “Additional unscorable sequences include sequences in which both the person- and object-

4 language and reference framed as: people communicating about objects; preverbal communication framed as: infants nonverbally conveying their object-intentions to adults.
orientations were present, but in which [...] the adult clearly elicited each focus in succession [...] A final unscored group of sequences comprises rituals (for example ‘peekaboo’), i.e. reciprocal exchanges built up over time by adult and child in which there was a fixed set and order of actions. These sequences involved a training element and/or a reliance on the mother’s stereotyped behaviors. For a variety of reasons, it was decided that these sequences ultimately require a different kind of analysis.” (Sugarman-Bell, 1978, p. 56)

This note of exclusion of adult-elicited shared orientations and trained routines highlights challenges and limits of this particular kind of behavioural coding; at the same time we only here are informed of the existence of particular classes of observed interactions which might be of particular interest for the development of attention and action co-ordination but are not described further, inviting us to have a more detailed look at these classes of interactions in our naturalistic longitudinal study.

2.2.4 Cognitive complexity accounts: a comparison

What these three approaches have in common is that they all address and account for the developmental changes over the first year, in particular towards the end of the first year, concerning the phenomena they are interested in (emergence of intelligence and instrumental actions, onset of intentional communication, respectively) in terms of the differentiation and increasing complexity of the infant’s actions and cognitive structures (in particular an increasing means-end differentiation and causal understanding). Additionally Bates as well as Sugarman suggest an increasing differentiation between the way infants used objects as instruments and the way they used people as instrumental agents, already foreshadowing later categorical distinctions between object and people engagement, but let us for now stay with the general point and unpack it into its parts.

2.2.4.1 The infant as the unit of analysis: The focus and unit of analysis of all three approaches is the infant as an active, autonomous individual.

Piaget went so far as to focus and restrict his observations almost exclusively on infant-object interactions, neglecting any social context (only
acknowledging social influence for culture specific “content” such as arbitrary symbols or morals, as well as some more complex logic operations). This overemphasis might be understood in the context of his efforts to carve out the infant as an active individual developing cognitive structures out of him-/herself through autonomous sensorimotor interactions with the (object) world from early on. Piaget thus separated the infant’s intelligence out of the “logos complex” (see section 2.7), that is, freed it from language and hence sociality, seeking to counter approaches rendering the infant’s intelligence completely dependent on sociality (including his own previous overemphasis on socialisation to overcome the child’s assumed egocentricism), or even denigrating it to being passively shaped by “drill” and social reinforcement.

Bates and colleagues, as well as Sugarman, were part of the pioneering language researchers interested in preverbal communication, who separated certain not-exclusively-verbal aspects of language (e.g. “reference”) out of the “logos-complex” pulling them downwards and looking for them at earlier ages within early social interactions. Looking for a framework for early development they drew on Piaget and – seeking to understand aspects of language – extended and applied his framework to the social realm. Thus, while (in contrast to Piaget) explicitly focusing on infants’ social interactions, their level of analysis still remained exclusively the infant, only analysing the infant’s actions, or even infant initiated actions (Sugarman’s study explicitly excluded actions where the mother established the shared focus, as well as familiar “rituals”) to check for specific characteristics (non-verbal performatives, simple/complex single/integrated person or object actions) and from this determine infants’ abilities, and potentially infer capacities and cognitive structures.
2.2.4.2 Actions and cognitive structures: All three approaches share a rather abstract logic- and language-oriented conception of action and thinking, which are partitioned into units defined by their functional roles, in particular in terms of specific effects on objects or people, observable for a researcher, supposedly in the service of instrumental reason.

While pointing out the continuity between sensorimotor action and intelligence and cognition (see below), Piaget – being mainly interested in logic and mathematics – tended to look at early psychological development as leading towards reversible logic operations (state-based, atemporal) as their end point. Due to this orientation and reflecting his observational method, he considered processes of thinking and acting primarily in terms of acts of instrumental intelligence, partitioned into units defined by their functional roles, i.e. their specific observable effects on objects or people performed step by step in the service of instrumental reason. While he seeks to ground intelligence in sensorimotor interactions, even basic sensorimotor schemas\(^5\) are – due to the constraints of his method (and era) – conceived of in quite abstract form, removed from physiological processes of motor planning and control (though less so than later cognitivist sense-think-act approaches) (Piaget, 1962).

Likewise, Bates and Sugarman interested in (the cognitive aspects of) language (i.e. reference) look at early development already with these later emerging complex language structures in mind. Looking for analogies in pre-verbal actions and hence selectively paying attention to the cognitive aspects they are interested in, they find a good match in Piaget’s model of instrumental actions. Again due to this orientation and apparently having

\(^{\text{S}}\)

Schemas of action are defined as coordinated systems of movement and perceptions, which constitute any elementary behaviours capable of being repeated and applied to new situations, e.g., grasping, moving, shaking an object (Piaget, 1962, p. 273).
not yet been able to integrate the then new research on early interactions and proto-conversation, they see infant’s early communication as merely automatically caused reactions to inner states, and conceive of later social interactions such as showing objects exclusively as deliberate intentional, indeed rational instrumental acts – with “getting attention” as a rather abstract and late acquired goal – rather than e.g. as affect imbued exchanges and acts of sharing between interaction partners, co-creating and regulating the interaction and affecting each other.

Conceiving communicative intentions, and intentionality in general, exclusively in terms of instrumental means-end thinking in which “the goal” and the “means” of an action are clearly distinguished as separate discrete units, paves the way towards a rational conception of thinking and acting as a series of discrete steps: 1) forming a goal or intention, 2) ahead-of-time planning, and 3) executing the plan. This idealized conception may to some extent capture (aspects of) conscious rational thinking, but may be misleading as a generalized frame for understanding complex multi-dimensional processes of control and coordination, artificially separating planning and control of action (a tendency we will find exacerbated in later cognitive approaches, see e.g. Tomasello section 2.4).

In addition, their research methodology based on verbal transcription may further reinforce this one-dimensional, serialised conception of (in particular communicative) action and cognition: by recording behaviour by taking notes and transcribing video recordings into linear text as their first step, and exclusively relying on this text-based version for further analysis, they inadvertently collapse a complex multi-modal and multi-strand social interaction into a serial, single channel, and one dimensional sequence of selected discrete steps, reflecting the selective interest and attention (Tomasello’s “intentional perception”) of the researcher. In general, managing, enacting, and coordinating activities of increasing complexity here is reduced to singling out and selecting the correct action unit from an ever larger repertoire and sequentially chaining them together into longer and longer stacks, thus missing some of the ongoing continuous, analogue, aspects (including affect) of the multi-modal multi-stream co-ordination.
2.2.4.3 The focus on (inter-)action and its complexities:

Though rendered common place by countless quotations and textbook descriptions/renditions, Piaget's considerations on sensorimotor processes – held in high regard in the embodied action currents in cognitive science (Bickhard, 1992; Di Paolo, Barandiaran, Beaton, & Buhrmann, 2014) – as well as his biological process model of assimilation and accommodation merit another fresh look as they – rather than later abstract “information-processing models” – offer a rich source for building a framework considering “developing in interaction” as a process, where all poles in an interaction, people and objects, actually affect and change each other on many levels and at many timescales in a process of ongoing transformation. And yet, these processes seem to remain rather abstract in Piaget’s account – giving rise to Vygotsky’s criticism that what Piaget’s infants lack is “reality, and the [...] relationship to that reality. What is missing is the child’s practical activity” (Vygotsky, 1987, p. 87). This peculiar dissonance in Piaget’s thinking might be caused by a clash between conceptualizing the irreversibly transformative historic change processes of evolution on the one side and reversible logic operations performed offline and thus not affecting the "real world".

All three accounts address development over the first year of life, including the challenges and major changes in the development of triadicity, in terms of complexity of action, which later tends to fade into the background in favour of accounts in terms of specific abstract knowledge. While these three accounts are all characterized by a somewhat abstracted and sequential conception, they place a major focus on the challenges of coordination of actions, which gains a new relevance when considering actions as multi-modal and multi-strand processes of joint praxis.
2.3 Accounts II: Social relating and co-ordination of social and object realms (Trevarthen & Hubley, Adamson & Bakeman)

The research programs discussed in this section are examples of research pursued in the generations after Piaget, countering Piaget by emphasizing in particular social relating and interacting. For these approaches, the challenge and achievement of infants’ development over and towards the end of the first year primarily comes to mean coordinating and integrating engagement with the two domains, now considered separate, of people and objects. We discuss 1) Trevarthen and Hubley's account of infants' developing intersubjectivity from “primary” to “secondary intersubjectivity”, and 2) Adamson and Bakeman's account of the gradual emergence of “coordinated joint engagement” between people and objects.

2.3.1 Trevarthen and Hubley: from primary to secondary intersubjectivity
(Trevarthen & Hubley 1978, Hubley and Tevarthern 1979s)

2.3.1.1 Infants’ early intersubjectivity

Trevarthen's early research conducted together with Martin Richards and Barry Brazelton at Jerome Bruner’s Center for Cognitive Studies in Harvard in the late 1960s, where they observed infants from birth to their 6 month of life, convinced him of two things: First: the behavior newborns showed was not reflex like but purposeful, integrating the whole body, including highly co-ordinated eye-head or hand-to-mouth movements, which he also experienced as being “aware” and “conscious.” And secondly they seemed to behave in two distinctively different ways depending on whether they engaged with objects or with people:

“They seek physical objects as sources of perceptual information or interest, and also as potentially graspable, chewable, kickable, step-on-able or otherwise usable. But persons are communicated with.”(Trevarthen, 1979, p. 323)

Trevarthen describes these – coming from a biological systems, physiological, and motor control background – as two different “regulatory states”, “inner contexts of action” and later simply “motives of behavior”. In
"The self born in intersubjectivity" (1993) he writes about the motivation for this concept:

[T]he word schema suggest a static pattern or plan that can keep a record in its shape. It has no efficiency of its own. It brings to mind something seen from outside and described, or “schematized”. We need a term to designate an active process or system more closely identified with the life of the subject. Program is the term used in cognitive science and artificial intelligence, but this is a set of logical or mathematical instructions that determines how information is processed through the system defining its input-output relations. Again we have a word that does not convey the inner generative and developmental aspect of psychological activities. A program, even one called generative, has to be written and then put into effect by some agency that delivers the right input language of energy and information to the programmable system. It is a rational, symbolic insertion. Psychological activities, with other life activities, have their own agency: They are autopoietic, or self-creating, like a growing organism, and they generate both awareness and the action in it. The word motive seems more appropriate to describe psychological functions that develop in the subject’s mind in readiness for perceiving the information needed for acting. The innate releasing mechanisms of von Uexküll and Lorenz […] is a motive, but it is defined as triggered or released by a “sign stimulus.” We need a term that indicates, more directly, the psychological function that explores and orients towards the consummatory response or goal. (Trevarthen, 1993, p. 123)

“Motive” in the sense used here, designates a mental function that is a cause and director of movement and, at the same time, a seeker of information to direct and confirm movement – to make it work for a purpose […]. A motive causes a subject to be curious and exploratory, as well as purposeful and effective, to be prepared to react selectively to the information that will be taken up in perception and to seek immediate influences that are appropriate for direct, ongoing control of acts and their effects. The motive regulates what will be chosen for uptake in perception and for retention in memory. Motives originate in largely inaccessible cerebral activity, but because they generate a wealth of movements for
aiming and focusing perception as well as for acting on the world, they are as real and readily observable as any regulatory principle in behavior. The central energy and self-regulating quality of motives are expressed in emotions. (Trevarthen, 1993, pp. 123–124, emphasis ours)

For Trevarthen, the remarkable thing about infants’ different modes of acting towards people and towards objects is - especially with regard to various theoretical positions within developmental psychology, in particular Piaget’s exclusive focus on object interactions and neglect of the social discussed in the previous section - that there indeed seems to be a distinct social motive active in the infant from the beginning, which becomes manifest in neonatal imitation and proto-conversation (Bateson, 1975, 1979; Brazelton, Berry, Koslowski, & Main, 1974; Bullowa, 1979; Lock, 1978; Snow, 1977; Stern, 1971; Tronick, 1982)

To account for these intricately coordinated behaviours Trevarthen formulates his “Theory of Innate Intersubjectivity”: “a child is born with motives to find and use the motives of other persons in “conversational” negotiation of purposes, emotions, experiences and meaning. [...] the idea of infant intersubjectivity is no less than a theory of how human minds, in human bodies, can recognize one another’s impulses, intuitively, with or without cognitive or symbolic elaborations” (Trevarthen, 1998).

Thereby Trevarthen also clearly positions himself within the field of various theoretical approaches: he sets himself apart from Piaget and cognitivist theories in claiming that there is an active social motive, an understanding of and meaningfully interacting with others from the beginning, not necessarily depending on cognitive or symbolic mediation. He distinguishes himself from cultural psychology approaches by stressing the active role of the infant, and from clinical approaches by insisting that this active role goes beyond responding for being-cared-for or being regulated but rather is a genuine seeking for engagement and sharing experiences with others and engaging in joint regulation and also implying a genuine intersubjective understanding. What exactly does he mean by “intersubjective understanding”?
The term “intersubjectivity” is drawn from Habermas’ theory of communicative action. Trevarthen repeatedly stresses the impact Habermas’ work had – along with Austin’s speech act theory – on the field of early infant communication: how his efforts to put language in a broader pragmatic context of communication and to develop a framework describing more general characteristics and preconditions for mutual understanding were readily taken up for example by Ryan, as they could fruitfully be applied to investigating the development of communication before language (compare last section). In Bullowa’s “Before Speech” Trevarthen explains what work the term "intersubjectivity" does for him:

*It is not a graceful word, but it does specify the linking of subjects who are active in transmitting their understanding to each other. The relating is "interpersonal" but we need to penetrate the psychological process by which conscious intending subjects relate their mental and emotional processes together. I feel that the "intersubjective" emphasizes this.* (Trevarthen, 1979; note Trevarthen’s emphasis on consciousness and awareness)

For Trevarthen following Habermas the concept “intersubjectivity” defines the very requirements people need in order to participate in communication:

*First they must be able to exhibit to others at least the rudiments of individual consciousness and intentionality. This attribute of acting agents I call subjectivity. [...] By subjectivity I mean the ability to show by coordinated acts that purposes are being consciously regulated.* [examples given: focusing attention on things, handling and exploring objects with interest in the consequences, orientating or avoiding while anticipating the course of events and meeting or evading them.]

*In order to communicate, infants must also be able to adapt or fit this subjective control to the subjectivity of others: they must also demonstrate intersubjectivity.* (Trevarthen, 1979 emphasis ours)

How intersubjectivity changes over time was investigated by Hubley in her dissertation research (Hubley, 1983; Hubley & Trevarthen, 1979; Trevarthen & Hubley, 1978)
2.3.2 From primary to secondary intersubjectivity

Penelope Hubley began data collection for what was initially intended to form the basis for two distinct studies on the development of 1) mother infant social interaction and 2) early object reaching and manipulation, however, it soon "became apparent that social and object related activities did not remain separate after 4 to 5 months" (Hubley, 1983, pp. 6–7). Hence the focus of attention moved to look into how this joint activity developed and changed over time.

2.3.2.1 Descriptive Qualitative Longitudinal Case Study: Trevarthen and Hubley (1978)

For their study, Trevarthen and Hubley followed free play interactions of a young girl (Tracey) with her mother across 32 lab visits from 3 weeks to 12 months of age. The pair brought their own toys from home including a wooden pull-doll, a multi-part globe rattle, and a wooden truck on a string with 2 wooden figures. Tracey was placed in a specially designed seat, supporting her trunk in a near vertical position and allowing free limb movement. Tracey’s and her mother’s actions and facial expressions were recorded on 16 mm film projected side-by-side using a mirror.

In the first month Tracey - in her brief periods of alertness - showed tracking and pre-reaching movements towards a ball suspended in front of her face as well as communicative actions (eye-contact, cooing, and body and arm movements) towards her mother, which flourished into proto-conversations over the next two months, though being slightly weary with people, paralleling her mother’s slight insecurity, which also led to a longer break of several weeks in recording. Data collection was resumed with a happier dyad at 4 months, and by 5 months the infant’s interest in - and grasping movements towards - objects had increased, as well as towards her mother’s hands and mouth. By 6 months she manipulated objects under close visual attention with an expression of serious intent with hardly paying attention to her mother. She accepted or pushed away offered objects, in the first case without any motion towards giving them back, though dropping them became a game for which she and her mother shared
amusement. Tracey also expressed joy when successfully setting a jumping-jack doll in motion pulling on its chord, without however directly looking at the mother or her reaction. While during these months interest in direct social face to face communication had decreased, out of the mother’s touching her and animating objects two mediated forms of social play emerged: 1) games of the person from around 4-5 months, where the bodies’ palpable, object-like characteristics as an object to be physically engaged with or a means for displaying exaggerated actions came to the fore, and 2) games with objects from around 5-6 months, where the mother animated objects (analogously to moving body parts in games of the person), and rhythmically moved, loomed, or made them emit noises, following and adapting to Tracey’s interests in the effects of her own and her mother’s actions on the objects.

As with games of the person, Tracey co-operated closely with her mother, but she did so without giving more than an occasional glance to her mother’s eyes and without looking at her mother’s expressions to observe feelings or interests concerning herself. She appeared unable to attend to the other’s purpose directly, or else resistant to it. (Trevarthen & Hubley, 1978)

Around 7-8 months, within games with objects, a reversal occurred: as Tracey shook and banged objects, and her mother replied with “bang-bang-bang” she realized the effect her object actions had on her mother and the effect became the goal of a new game leading to smiling and laughter. However, Tracey still failed to act reciprocally in giving objects back to her mother’s open hand and was not truly co-operative, failing to attend directly to the intentions of her mother.

From around 9-10 months on, Tracey interrupted her actions to make eye-contact and exchange affect and paid attention to demonstrations of object acts, showing willingness to share her experiences. She followed when her mother pointed to an object and at the command “pull it!” made a move to pull the toy trolley towards her. A little later she engaged in giving and taking and around 11-12 months was able to select objects named by her mother in a naming game.
“With babies in the second and third month, most mothers we have filmed played games that involved touching the infant's body, like pat-a-cake with the hands, bouncing the legs, shaking the cheeks, prodding the nose or stomach. Gradually, it would seem the mother herself is accepted as a game object as she mirrors the infant's acts of expression. After this the play incorporates objects that the infant has accepted as foci for interest. We found that by 6 months these games via objects, or with parts of the mother's body treated as objects, became the infants' preferred form of play. Then, at 9 or 10 months, they started the deliberately co-operative form of interest in objects which transforms play into exchange of acts of meaning.” (Trevarthen & Hubley, 1978, pp. 211–212)

2.3.2.2 Quantitative Longitudinal Study: Sharing a Task in Infancy (Hubley & Trevarthen, 1979)

For a follow-up study aiming to investigate the emergence of cooperation, 5 mother-daughter pairs were videotaped at the lab for 6 play sessions when the infants each were 34, 38, 42, 46, 50, 54 weeks (~8-12 months) of age. The mothers were instructed to engage in 4 conditions, of 4 minutes duration each: 1) play without toys, 2) join in play with toys (a similar set of toys as in the pilot) 3) refrain from initiating play, just watch the infant play, but respond with gaze or smile when addressed, and 4) teach the baby a simple manipulation task: putting 3 wooden figures into a toy truck. Analysis was restricted exclusively to this forth condition.

All behaviours coded were considered as “acts of communication”, as they “regulate communication between partners because they are directed in relation to a manifest interest or activity of the partner, or because they respond to acts of communication.” Two types of communicative acts are distinguished: 1) interpersonal acts: acts directed to the other person making no reference to objects, including looks, smiles, laughs towards each other, vocal imitation, movements towards or touching each other; and 2) acts of joint praxis: acts on objects that are oriented to the attention or action of the other person. Thus, interestingly, Hubley's and Trevarthen's “acts of joint praxis” by themselves already include coordination between people and objects, that is, already show some aspects of triadicity.
Hubley andTrevarthen reported a stable number and duration of communicative episodes across infants and ages, with a quite stable composition of such episodes, around 2/3 of which were categorized as “joint praxis only”, 10-20% “interpersonal acts only”, with the combination of “joint praxis and interpersonal acts” in the same episode increasing from 15% in week 34 to around 30% in week 54. Only infants in the older age groups began to (sometimes) imitate demonstrations and comply with instructions, which became more frequent and more elaborate with age.

*Our results show that between thirty-four and fifty-four weeks infant subjects began to engage with their mothers in communication about using objects. […] As the infants grew older they began to integrate interpersonal acts and acts of joint praxis together, they complied with the mothers instructions, and the imitated her acts on objects. […] Mothers and infants together started communicating, using objects in a way that transmitted messages and invoking people as agents to help in a task. (Hubley & Trevarthen, 1979; emphasis ours)*

For our current purposes, the results - even if restricted to condition 4 – are highly suggestive in that infants from the first session (at 8 months) already abundantly engaged in acts of joint praxis, i.e. coordinate between persons and objects. However, the authors unfortunately do not further describe or analyse joint praxis, nor look at its development (before 8 months) in this study. To shed light on the development of joint praxis itself, it would have been interesting to learn about how joint praxis is organized, expanding on their qualitative observations from the case study (see above).

### 2.3.2.3 Accounting for the emergence of secondary intersubjectivity

Apart from the behavioural description of the emergence of secondary intersubjectivity as the integration of joint praxis and interpersonal acts, Trevarthen and Hubley account for these developments in terms of 3 different systems, the seeds of which are already in place from birth and which develop in alternating waves and reach a new level of integration between 9 and 12 months:
These modes are probably three real systems of the brain that achieve functional differentiation by interaction with each other and with the environment. Forms of action and perceptual processing appropriate for 1) knowing and using objects (praxic mode), for 2) communicating with the human world (communicative mode), and for 3) acting in self-directed or thoughtful manner (reflective mode) appear as distinct rudiments in the new-born. [...] Once free interaction between communicative and praxic modes of action is achieved, the infant suddenly shows behavior that is unique to man in its complexity, and full of potential for the development of knowledge, joint enterprise, and language. (Hubley & Trevarthen, 1979; emphasis ours)

While distancing themselves from Piaget, Trevarthen and Hubley indeed consider the infant as more or less the sole source of agency in development, not unlike Piaget; only with infants not developing exclusively in interaction with objects but also through interaction with social partners who adapt to them, thus providing an even better source of interaction. And, crucially, adult social partners engage with infants as companions. With regard to the concrete course of development however, Trevarthen and Hubley do not credit adult social partners much influence beyond what the infant’s development and actions dictate:

Even though the mother’s adaptations were undoubtedly influenced by knowledge and ideas imported from her cultural experience away from Tracey, the strong regulation of the play by Tracey’s acts makes it unlikely that the games were invented by the mother. The games appear to follow development in the object-seeking and person-recognizing functions Tracey has exhibited since her first month, and to depend on the mother being interested and sufficiently aware to adapt to these functions. (Trevarthen & Hubley, 1978)

Changes are initiated, not by the adult rule bearers, but by the infant and child. [...] The vocabulary of language, games, toys and all other cultural artefacts enrich the possibilities of life of an infant because they meet the infant’s habits of intersubjectivity. Changes at certain ages, such as the change at 9 months, caused the rules of interaction to change. People in the
social world react by becoming affectionate, co-operative, interested and talkative, adapting to the forms of social action that seem most natural to an infant at each age. (Trevarthen & Hubley, 1978)

In addition to their general account in terms of 3 developing and increasingly integrating action systems, they repeatedly mention a change in understanding or knowledge as a key factor the qualitative changes in interaction, namely “the infant’s changing understanding of her mother as a person”:

At the beginning of the study [34 weeks] the infants did not understand or have clear expectations of the mothers’ behavior with objects, hence the mother was not able to communicate about objects to them. […] We believe that the meaning of the mother’s activity changed because the babies started somewhere in the 20 weeks over which the study extended to conceive their mothers as praxic agents doing particular things. […] Perceiving someone as an agent implies seeing that person as having plans and goals in action; it implies a degree of subjectivity. Cooperating with someone in joint action on objects involves taking up at least some part of the other person’s plans and goals. In cooperative actions on objects the partners combine communicative intention and praxic agency having effect both on each other and on the physical world. (Hubley & Trevarthen, 1979; emphasis ours)

This focus on understanding or knowledge about someone, her plans and goals, using cognitive terms now, we will later find again in Tomasello’s account as the crucial aspect suddenly leading to the different ways of relating in the course of the “9 months revolution” (see section 2.4.).

2.3.3 Adamson and Bakeman: passive and coordinated joint engagement
(Bakeman & Adamson, 1984)

2.3.3.1 Background, interests, and theoretical framework

In a seminal study and a series of resulting papers appearing throughout the early 1980s Adamson and Bakeman set out to explore “the emergence of the ability to coordinate attention toward a social partner and an object of
mutual interest”, here already dubbed “a developmental milestone” (Bakeman & Adamson, 1984).

In their 1982 reflective “research-in-progress-chapter” they share some of the background: Interested in the development of infant communication they found themselves facing a gap between 2 different research traditions interested in different aspects and age groups: early communication researchers looking at arousal and affect regulation in 0-5 month olds, and language researchers looking for the first cases of reference from around 9 months on. This gap with regard to age as well as concepts made it hard to compare and relate the approaches and hence made an understanding of the process of development of communication difficult (Adamson & Bakeman, 1982).

Adamson and Bakeman became convinced that to understand the development of communication it was necessary I) to look at the same parameters across the whole age span, II) to take into account the interaction of all participants, and that III) sharing and guiding attention (in a situation where people and objects come together, which Adamson also dubs “triadic”) was at the heart of language and communication – as Adamson elaborates in more detail in a later textbook (Adamson, 1995) 6.

6 In her 1995 textbook Adamson reminds the reader that this idea should not be taken for granted, as it e.g. is not found in Piaget’s account, and traces aspects of it back through history:

1) The idea that a sharing situation, where people and objects come together in close vicinity so (attention to) objects can be shared, is crucial for language development, is common to both Werner and Kaplan’s (“primordial sharing situations”), as well as Jakobson’s language model, a similarity Bakeman traces back to their common teacher Karl Bühler and his organon model of language.

2) Turning to the process of sharing and directing attention itself and underscoring its role for language (development), Adamson appeals to Bruner: “The problem of how reference develops “ can be restated as “the
2.3.3.2 Empirical Study

problem of how people manage and direct each other’s attention”, quoting from his book Child’s Talk. This book is dedicated - and hence again links these thoughts - to Bruner’s “teacher and friend” Jacobson [and thus in turn, indirectly, to Bühler].

3) Interestingly, the discussions in her 1995 textbook strongly suggest that, for Adamson, the role of sharing and directing attention for language goes well beyond being an expression of a general prerequisite-capacity of “joint attention” reflecting specific social knowledge necessary for sharing, as it is conceived of by knowledge-based cognitivist approaches. Rather, following Jacobson, she points out the **concrete functional role** of jointly shifting attention for ongoing communication: **each concrete attentional shift** towards a particular aspect within an ongoing communication matters, as it effectively changes and newly **determines** the joint topic, and with it also the **function of communication** – being **emotive** with focus on the speaker; **conative** with focus on the addressee; **referential** with focus on objects; **phatic, metalinguistic, and poetic** with focus on the connection, the linguistic mode of communication, or its concrete form, respectively. She also adapts this model to capture crucial characteristics of communicative development in terms of the shared topic of communication: the **connection** is shared in the first “**two months of shared attentiveness**”, **each other** and the **lines and messages of communication** are shared in in **“interpersonal communication”** (2-6m), **objects** in **“joint object involvement”** (notably here starting already from 6m), and, finally, around 13 months – with increasing ritualization – the **code** comes into shared awareness going hand in hand with the emergence of **symbolic communication**.

Note one **difference** between the model presented here and Bühler’s model: what is not explicitly emphasised here, while it lies at the centre of Bühler’s model appears to be **jointly acting together**.
Thus they devised a study to investigate I) the development of various parameters across an age range from 6-18 months, and used II) a relational coding scheme taking into account both interaction partners and their relation as a basis for III) focussing in particular on the development of attention coordination (and relate it to multiple other parameters, see below).

In their longitudinal study two cohorts of 14 infants each were visited in their homes at 3 months intervals, from 6-15 and from 9-18 months of age, respectively. At every visit infants were videotaped in three different contexts of 10 minutes duration each, in order to investigate the role and contributions of different interaction partners: a) in free play with mothers using toys provided by researchers (toy telephone, picture book, wooden puzzle, nesting cups, doll, rattle, soft toy with wheels), b) with peers using the same toys, and c) placed alone and provided with toys (with their mother in the room).

1) In their first report on this study, they particularly focused on a) the appearance of attention coordination (its first emergence as well as routine use) as well as b) the role of the adult partners (as opposed to peers and solitary engagement) potentially fostering coordination (Bakeman & Adamson, 1984). In subsequent reports they 2) analysed whether and how the mothers’ communicative acts changed over time in accordance with the infant’s development (Adamson & Bakeman, 1985), as well as 3) how the display of affect changed when infants’ “involvement in social interaction shifted from participation in predominantly ‘expressive’ face-to-face exchanges of early infancy to participation in the preverbal ‘referential’ dialogues of late infancy” (Adamson & Bakeman, 1984).

For analysis the infant’s attention was coded using a state-based (relational) coding scheme with 6 distinct mutually exclusive “engagement states” which allows – with respect to report 1) – to directly address and measure attention coordination, and provides – with respect to reports 2) and 3) contextual parameters to relate the measures used in reports 2 and 3 to. The 6 engagement states were: a) unengaged: uninvolved with any specific person, object, or activity b) onlooking:
observing another’s activity, but not taking part in it c) **person engagement** (e.g. face-to-face, person play, cooing at tickling, vocalizing and reaching towards person), d) **object engagement** involved in playing with objects alone, attending just to the toys at hand, e) **passive joint engagement**: infant and caregiver actively involved in the same object, but the infant evidences little awareness of other’s involvement or even presence (Mothers often attempt to induce this state in their babies. They will manipulate objects (e.g., shake rattles, ring toy telephones, etc.) in ways that seem designed to capture their infants’ attention by making the objects “come alive” for them), f) **coordinated joint engagement**: “infant actively involved with and coordinates his/her attention to both another person and an object that person is involved” (Bakeman & Adamson, 1984).

The results Bakeman and Adamson reported with regard to question 1) the development of attention coordination, included: a) **age effects**, as well as b) **partner effects**. A) Age effects: the relative percentage of **person engagement** (dyadic) steadily decreased between 6 and 18 months (roughly from 12% to 5% with mothers, from 12% to 3% with peers), while the relative percentage of **coordinated joint engagement** (triadic) increased with age (with mothers starting from 2% (at 6m), then jumping from 4% at 12 months to 11% at 15 months and 27% at 18 months, with peers it grew from 0.3% to 7%. The duration of coordinated episodes also increased with age, on average from 7 seconds to 34 seconds with mothers and from 2 seconds to 9 seconds with peers, respectively. The percentage of the other engagement states remained stable over time (with onlooking between 8-11%, object engagement between 37-43%, and passive joint engagement between 17-24% with mothers, whereas between 2-4% with peers). b) Partner effects: as shown above, both passive as well as coordinated joint engagement were more likely to occur when infants played with their mothers than with peers.

In addition, a sequence analysis of the respective engagement states showed that **object engagement** was most likely to be followed by, and to follow, both passive and coordinated joint attention. Moreover in the condition with mothers, **passive joint engagement following a period of**
onlooking was also a likely sequence. Person engagement, in contrast, was not a likely precursor of either passive of coordinated joint engagement.

With regard to question 2), how mothers’ communicative acts changed over time, they reported a change from self- to object-marking and within object-marking from literal marking (e.g. animating an object) to conventionalized object marking (words, gestures and ritualized games, i.e. give and take, book sharing), with these ritualized games greatly increasing after 12 months and with over 20% constituting the by far largest share of activities at 15 and 18 months). Note how their results come alive in their vivid verbal description:

The mothers we observed gradually altered their actions in a way that reflected their infants’ emergence as increasingly able users of conventional, object-focused acts. They seemed therefore to be dynamic, developing partners who tailored their actions to suit their infants’ developmental path.

The age effects observed were striking. Although mothers always tended to act and to chat about half the time during a play period, they were not always enacting the same plot. To a 6-month-old partner, they presented an animate, literal world […] and were most likely to highlight themselves and their social link to the infant. If objects were acted upon, it was often to make them “come alive” with sound and movement as if extending their own perceptual qualities outward beyond the boundaries of their social interchange. Rarely were acts that gained meaning only through social convention used, and when they were they were as likely as not to occur in concert with literal actions which, by their actual force, might draw the infant’s attention. Mothers of older infants, in contrast, seemed less like stars in a social relationship and more like narrators of the world surrounding them and their infants. They rarely marked only themselves. Yet their many object-markers were now often socially mediated. Conventional acts, performed without literal support, were both numerous and patterned to enact shared rituals such as “bookreading” and “give-and-take.” Thus while the attentional focus of the mothers’ actions was primarily away from
the immediate social interchange, the way it was marked was fundamentally shared and social. (Adamson & Bakeman, 1985)

A systematic analysis of question 3: infants' frequency of affective displays, found that affective displays generally increased with age, but mean duration of displays decreased (from over 3 to under 2 seconds) and shifted relatively to the vocal modality with age, and was most frequent in person play (with mothers and peers) and in object play, especially with mothers when affect sharing was sustained throughout the activity.

Generally, Bakeman and Adamson conclude that coordinated joint engagement emerges around 9 months, but becomes routinely used only from around 15 months, and – as suggested by the sequence analyses – that infants only begin to coordinate engagement after they begun to focus on objects on their own. Bakeman and Adamson align themselves with Nelson (1979) in that this surprisingly late mastery of attention-coordination might be the reason for language developing only relatively late – well after its prerequisites already seem to be in place, and they affirmatively take up her suggestion that infants indeed seem to need that time to meet the challenge to – in Nelson’s and Adamson & Bakeman’s words – “co-ordinate social and object realms”.

In the spirit of Mead, Vygotsky, and proponents of a social interactionist perspective, such as Bruner, the Newsons, and Trevarthen, who Bakeman and Adamson refer to, they interpret their results as affirming their perspective that “new communicative forms occur first embedded in and supported by social context”. In particular, they see the results as strongly suggestive of mothers playing a crucial role in the development of communication: engagement states that offer some potential for referential communication, namely passive joint engagement and coordinated joint engagement, were observed much more frequently when infants were with mothers than with peers. The biggest differences were found for passive joint attention, where infants spend around 20% of time in when with mothers (none with peers) irrespective of age, that is, it was observed even in the youngest infants of the study. Mothers seemed to deliberately induce
passive joint engagement (accounting for the difference with peers, as mothers “were more willing and able to complement the infant’s attention”) to some extent freeing infants of the need to establish co-ordinated attention to objects and people.

2.3.3.3 Accounting for the development of coordinating attention to people and objects

Seeking to find in the process of observing real life data that coordination ability, which later gives rise to referential communication, and trying to operationalize it for coding, the concept “attending to the same object” apparently turned out not to be a sufficient coding category: neither to appropriately capture this ability nor to differentiate interactions at different ages, as “attending to the same object” was already observed from the beginning of data collection (6 months).

This seems to be the reason that they split up this class of interactions into two categories, singling out joint (= triadic) engagement “proper” – what they now called “co-ordinated joint engagement” defined as “the infant is actively involved with and coordinates his attention to both another person and the object that person is involved with” – separating it from all other cases of “not-yet-proper-joint-engagement” – dubbed “passive joint engagement”, defined as: infant and caregiver being actively involved in the same object, but the infant “evidences little awareness of other’s involvement or even presence”.

What is striking is that, rather than relying on gaze measures to assess “visual joint attention” as has become the default strategy pursued in most later studies, Bakeman and Adamson deliberately chose the broad, inclusive notion of “joint engagement” to capture the ability in question and accordingly constructed a coding scheme based on broadly defined relational states, importantly giving room to include a variety of different forms of joint engagement, including but not restricted to gaze-based coordination. In Adamson and Bakeman (1985) they explicitly reflect on these methodological considerations:
A major advantage of such state-based schemes is that they provide descriptions of attention that are not based solely on patterns of visual gaze. Of course, there is a high degree of correspondence between the direction of gaze and engagement state. For example, by definition, infants do not gaze repeatedly toward a person in either object engagement or passive joint engagement, while visual regard is expected during onlooking, person engagement, and, to a lesser degree, coordinated joint engagement. But gaze alone may not provide a fully adequate index of the intensity of interest—a quality that often differentiates states such as onlooking from others such as person engagement or passive joint engagement. Moreover, by the end of the first year, infants appear to continue to attend to some portion of the surroundings while momentarily looking elsewhere, as when in "coordinated joint engagement" they gaze at an object but seem aware of their partner as well. (Adamson & Bakeman, 1985)

And indeed their definition of coordinated engagement stressing “active involvement” does not only seem to allow, but to actually call for, engagement beyond mere looking, e.g. physically-effective manual engagement. While the inclusive, holistic relational coding categories of “passive and coordinated joint engagement” seem vague, they apparently are strikingly robust to implement, as human observers seem to reliably recognize “joint engagement” when they see it.

While relationally defined, that is, taking into account both partners’ actions, the main drawback of the coding methodology in terms of these categories is that they actually cover up how the participants interact in detail. It would be very interesting to learn more about the actual range and modes of joint engagements as – despite their broad definition – their own example again reverts back to the typical “gaze-checking while pushing a toy truck” scenario. By pooling interactions into engagement states and then applying sequence analysis to extract global probabilities of transitions between states, Adamson and Bakeman are left with only speculating about why these sequences occur, or that the “passive joint engagement state seems to be closely tied to the mother’s actions”. In short, local as well as global information about the meaning of the interaction is lost. Therefore we
here will use qualitative microanalysis to shed some more light on these local and global aspects of meaningful interaction: how exactly interaction is co-ordinated, and what functional role specific actions play in the context of a larger meaningful interaction (see chapters 3 and 4).

2.3.4 Comparing Hubley & Trevarthen with Adamson & Bakeman

In Hubley and Trevarthen’s, as well as in Adamson and Bakeman’s approach, both focusing on “relational” aspects of the interaction, the challenge and achievement of the development of interacting over the first year and particularly over the final months of the first year are considered as the co-ordination of relating to **two different realms: people and objects**. Trevarthen – countering Piaget’s object-centredness – stresses infants’ intersubjectivity from birth – setting it apart from their way of relating to objects (based on his early observations of neonates responding differently to people and suspended objects), and insists that infants employ two separate “motives” or neural “action and regulation systems” for people and for objects (plus actually a third one, directed at the self), which need to be “assimilated” and integrated over time. On the other hand, Adamson and Bakeman, who, interested in the development of communication, found themselves in a gap between early-infancy-focused proto-conversation researchers and late-infancy-focused seekers of language-precursors, frame their research question from the outset as: “coordinating attention between people and objects” as key to understanding communication development. This may be in part due the dichotomous structure inherent in the common conceptualization of their phenomenon of interest: communication with people about objects and its framing by previous research (Bates, Sugarman-Bell etc., Nelson), as in their discussion they affirmatively take up Nelson’s formulation that the main challenge for infants to master, is to coordinate separate “**social and object realms**” (Bakeman & Adamson, 1984). Apart from this framing, Adamson and Bakeman largely remain descriptive, focused on the behavioural level, without speculating about inner psychological or physiological mechanisms.

Both approaches, rooted in observing semi-naturalistic interactions, do not restrict themselves to gaze measures for assessing “joint visual
attention”, but use broad, complex terms to capture the complex ways of relating which develop towards the end of the first year: “secondary intersubjectivity” in its operationalized form encompasses: “joint praxis”, that is, “acts on objects that are oriented to the attention or action of the other person”, in combination with “interpersonal acts” (see section 2.3). Similarly, “coordinated joint engagement” is constituted by “the infant being actively involved with and coordinating his/her attention to both another person and the object that person is involved with” (Bakeman & Adamson, 1984). With regard to their conceptualization of (triadic) relating, both of these notions point towards complex multi-strand, multi-modal actions, and “joint praxis” and “active involvement” evoke actual physically-effective (typically manual) object engagement, thus also fitting definitions of joint action and collaboration which typically require “jointly bringing about a change in the world” (compare Sebanz, Bekkering, & Knoblich, 2006).

Strikingly, both these conceptualizations also imply different forms of jointly engaging with objects before full-fledged secondary intersubjectivity, or coordinated joint engagement, occurs: In Adamson and Bakeman’s approach this is described as “passive joint engagement” – localized on the attention or attentive-communicative strand of joint engagement – effectively scaffolded by the mother following and guiding the infants’ attention. Conversely, in Trevarthen and Hubley’s approach this is described as “joint praxis” – localized on the physically-effective strand of joint engagement. Both categories make up a large part of the engagements of mother-infant pairs: “passive joint engagement” constituted consistently around 20% of the play interactions observed in Bakeman and Adamson (1984) from 6-18 months and “joint praxis” accounted for a large majority of communicative acts (joint praxis, interpersonal acts, joint praxis combined with interpersonal acts) in Hubley & Trevarthen (1979).

While Adamsom and Bakeman speculate on the potential importance of “passive joint attention” and the role of the caregivers establishing it, both categories have unfortunately not been investigated in more detail at this point, although this would be of key interest for understanding the development of triadic interaction. For both of the approaches these early
not-yet-full-fledged forms of joint engagement with objects are part of a larger account framing the development of triadic interactions as a gradual process (rather than a sudden emergence), with Hubley and Trevarthen sketching the development from proto-conversations in primary intersubjectivity, first via “games of the person”, then “games with objects” and “inverted games with objects” to finally infants imitating demonstrations and following instructions in full fledged secondary intersubjectivity. While both accounts propose a gradual development, in accordance with their different focus and perspectives, this development follows to some extent inverse trajectories: In Adamson and Bakeman’s account – focused on the attentive-communicative side and interested in the role of the caregiver – the infant moves from being “passively” attentionally engaged in “passive joint engagement” to actively participating in “co-ordinated joint engagement”, whereas in Trevarthen and Hubley’s account – focused on endogenous infant action and development – the infant moves from only interacting cooperatively as far as the mother adapts to the infant’s interests and purposes, to becoming more receptive and malleable accepting adults in new ways and understanding and taking account their intentions – in particular with regard to objects, thus giving rise to truly co-operative forms of interaction.

What seems particularly relevant about these two seminal approaches for future research is that they – in contrast to later research narrowing in on visual joint attention – develop broad, complex notions of relating, pursue longitudinal investigations of semi-naturalistic interactions and provide qualitative descriptions of gradual development, and in particular hint on different forms of early joint object engagement. All these aspects merit to be taken up again and looked into in more detail using current data collection, visualization, and analysis techniques.
2.4 Accounts III: Cognitive Accounts (Tomasello and colleagues)

Cognitive approaches have in common that they operate from a framework based on (classical) cognitive science and analytic philosophy and account for development by testing an individual’s cognitive capacities considered to be based on (and hence to some extend indicative of) specific forms of knowledge. They conceptualize the challenges of social interactions – and hence social development – in terms of problem-tasks to be solved by individual rational agents using instrumental reason, that is, thinking, planning and deciding in advance, then simply executing the plan, monitoring, and in the end evaluating the result. Analogously, they conceptualize their own approach to the phenomenon under investigation as conceptually analysing it in advance (drawing from literature and using thought experiments and conceptual analysis), devising experiments to test and decide between different theories and hypotheses about capacities, and interpreting and evaluating the results. With regard to the development of triadic interactions, many researchers have made a range of important and quite varied contributions from a cognitive perspective (such as Barresi & Moore, 1996 who are primarily focusing on executive functioning; Rochat, 2009, and many more); for our current purposes we primarily focus here on the approach pursued and popularized by Tomasello and colleagues because of its far-reaching comparative scope, systematic development over an extended period of time and refined through multiple revisions, and high impact with regard to the current interpretation of the narrative.

2.4.1 Background, interests, and theoretical framework

Michael Tomasello and colleagues (in particular at the Max Planck Institute in Leipzig) have been pursuing the (ever) notorious “what is special about the human species” – question (section 2.7.), and in particular – asking from a cognitive science perspective – how have such complex human forms of cognition such as language, mathematics, complex tool use, and social institutions – typically associated with culture – come about?

Expanding the cognitive science perspective with a comparative program, Tomasello and colleagues tackle this question by contrasting typically
developing human children at different ages with autistic children, and with non-human primates (as well as with dogs, with whom we share a cultural history of domestication rather than a close phylogenetic relationship). They effectively seek to construct a theoretical framework, a “matrix”, aiming to find features, which allow to logically and conceptually demarcate those groups from each other, at multiple levels: palpable cultural products, processes of cultural transmission, and – as the preferred level of “explanation” of cognitive science – (species-specific and typical) individual cognition. On the surface level of palpable cultural products, Tomasello and colleagues (1993) foreground human culture being “cumulative” as a main “distinctively human” feature, that is, cultural products are building on each other and getting more complex across time and generations. The authors then link this cumulative character of culture to specific forms of cultural creation and transmission: collaborative invention and high fidelity transmission, which in turn they seek to link to specific forms of “cultural learning” and “social cognition”.

For his theoretical framework Tomasello partly draws on (his interpretation) of Vygotsky’s ideas of “culture”, adopting in particular the idea of two different lines of development: a biological one, shared by humans with their great ape relatives (practical intelligence), and a cultural one, socio-culturally mediated, – for Vygotsky starting with language, in Tomasello’s version starting earlier with specific forms of cultural learning: learning through other people – which then thoroughly re-organizes the cognitive abilities of the biological line. The idea that cultural forms of learning are achieved “through” other people is the second aspect drawn from Vygotsky, from his notion of cultural mediation, and is reminiscent of Vygotsky’s formula: *The path from object to child and from child to object passes through another person* (Vygotsky, 1978a, p. 30).

In contrast, however, to Vygotsky’s focus on cultural mediation and how it transforms individual (biological) cognition, and, in contrast to the goals of cultural psychologists who - also drawing on Vygotsky - seek to develop methods to adequately capture how exactly cultural processes (also going beyond language) unfold, Tomasello selectively adopts some Vygotskian key
concepts (also see section 2.5.1: Vygotsky) and gives them a cognitive twist to assimilate them to the capacity- and individual-focused cognitive Psychology framework (also heavily drawing on Piaget) and asks: what are an individual organism’s cognitive capacities, that are necessary to engage in culturally mediated activities such as cultural learning?

2.4.2 Tomasello and colleagues, first take: cultural learning, joint attention and understanding people as intentional agents (1993, 1995, 1999)

Inspired by Vygotsky, Tomasello and colleagues begin to analyse the general notion of “social learning”, conceptually organizing processes of learning found in various forms across the animal kingdom, seeking to identify and set apart more specific forms of human-typical “cultural learning”: Common to the basic forms of “social learning” (= any form of learning which is influenced by the social context in any way, including local and stimulus enhancement, as well as emulation) is that learning itself, however, can still be considered to be achieved by an organism individually through simply focussing on and interacting with the physical environment alone, without taking another’s perspective or specific action strategies into account. In contrast, in “cultural learning” in the narrow sense, learning is defined as occurring “through the other” (which in this early account is conceptualized very broadly and tentatively as occurring via some form of “intersubjectivity” and/or “perspective taking”). They differentiate three types of cultural learning: 1) imitative, 2) instructive, and 3) collaborative, and go on to propose that these develop in a 3 step sequence. A main aim of their research program is to link each type of cultural learning to specific forms of infants' developing social-cognitive capacities and knowledge, and

\[ \text{7 Tellingly, Tomasello dismisses cultural psychology in passing for exclusively focusing on the role of culture and neglecting the individual – a characterization not agreed with by proponents of cultural psychology who explicitly seek to overcome such dichotomies (Rogoff, Chavajay, & Matusov, 1993).} \]
to contrast this developmental sequence with atypically developing children with autism and great ape relatives.

1) **Imitative learning** (emerging between 9-14 months): in which the learner reproduces a model’s specific actions or “actual behavioural strategies” novel to the learner in an appropriate context. This form is considered to be prominently used in learning object directed actions and conventional linguistic symbols, and is considered to be achieved through some basic form of “perspective taking”.

2) **Instructed learning** (emerging around 4 years and characterized as “Vygotskyan learning”): in which the learner internalizes the instructions of a teacher and later re-enacts them overtly in similar situations to self-regulate their attentional mnemonic or other cognitive functions, in the form of e.g. performance monitoring, meta-cognitive strategies, or self-regulating speech. This form is considered to be achieved by recognizing two perspectives (the learners’ own and their teacher's) and coordinating them.

3) **Collaborative learning** (emerging around 6 years): in which two symmetric peers work together to solve a common problem co-constructing knowledge neither of them had before, and which they then individually internalize. This is considered to be achieved by reflective integration of different perspectives.

Tomasello and colleagues (1993) proposed **differences** between typically developing human children and atypically developing children as well as great apes **already on the first level of (true) imitation learning**, let alone on the more complex levels.

**2.4.2.1 Joint attention and social cognition**

To better understand **imitation learning** Tomasello et al. (1993) draw attention to a number of other interactions appearing around the same time and take their synchronous appearance as an indicator for a newly emerging underlying capacity. The factor all these interactions have in common, and which thus seems to be a visible marker of this underlying capacity, is “joint attention”, sometimes taken to be the underlying capacity itself. This
concept is further developed theoretically in Tomasello (1995) and empirically investigated with Carpenter and colleagues (Carpenter et al., 1998).

In his 1995 “Joint attention as social cognition” paper, Tomasello seeks to establish “joint attention” as a marker for social cognition, thus pointing out the cognitive basis of joint attention he sees neglected in more behaviour oriented discussions of joint attention, primarily gaze following, and also making the case for earlier, more basic forms of social cognition than were proposed by “theory-of-mind”-theorists. To this end he seeks to clarify and define what joint attention is:

“Joint attention is primarily a social, social-cognitive phenomenon: two individuals know that they are attending to something in common” (Tomasello, 1995)

... and also what it is not, and basically claims that everything infants do before 9 months basically is NOT joint attention but merely “simultaneous looking”, including:

1) orienting to the same spatial location, but not to the same object aspect
2) looking to the same object, focusing on the same aspect, but simply watching the other person engaging with an object without engaging themselves (Bakeman & Adamson’s onlooking)
3) mother and infant both being drawn to the same object fortuitously, as when barking leads each of them to look at a dog out separate windows of the house.
4) cued looking, when infants have learned that looking in the direction another individual is looking, often results in interesting sights.
5) even gaze alternation is not necessarily a marker for joint attention, as it might be cued (Bakeman & Adamson’s passive joint attention)
6) looking up spontaneously to check whether to expect reward or punishment, i.e. not monitoring the other’s attentional focus on the object.
This definition of joint attention already implies that this process involves some kind of knowledge. According to Tomasello, the knowledge joint attention depends on is a specific understanding of the other person as an “intentional agent” whereby he understands “intentional” in the sense of Piaget as differentiating means and ends in their own actions (Piaget stage IV-V). This uniquely human knowledge is supposed to emerge from the confluence of two different cognitive aspects: 1) the uniquely human tendency to identify with others as being “like me” (e.g. Meltzoff & Gopnik, 1993), exercised from early on in imitative and – extending Meltzoff – reciprocal interactions, and 2) the understanding of the self as an intentional agent, differentiating means and ends in instrumental actions, which is shared with great apes and emerging in humans around 9 months of age (see also section 2.2.1: Piaget).

2.4.3  Tomasello and colleagues, second take: “shared intentionality” (2005-)

However, over the course of the progressing research program, a considerable number of studies, in particular studies from their own research group, suggested, that apes do understand others “as intentional agents” to a far greater extent than Tomasello and colleagues previously assumed, in particular they understand that other people have goals and perceptions (they e.g. differentiate between humans being unwilling or unable to give them food and respond more patient in the latter case etc.) (see e.g. Tomasello, 2011 for a summary), and, as has recently been demonstrated, even beliefs (Krupenye, Kano, Hirata, Call, & Tomasello, 2016).

In short: they were supposed to at least know the “peripheral aspects” of people’s minds, which can be inferred from observation. Apes seem to indeed understand intentions, which had been regarded as the crucial factor for engaging in triadic or co-operative activities, yet they still do not seem to engage in sharing, triadic, or co-operative activities. How come? Tomasello and colleagues thus had to revise their theory:

Whereas the previous developmental trajectory exclusively focused on and framed infants’ understanding of other people in terms of the
perception and knowledge of an uninvolved bystander, passively observing from a distance, in their revised version (Tomasello et al., 2005) they 1) refined their definition of “intentional action” previously based on Piaget by drawing on analytic philosophers such as Bratman, 2) further differentiated the development of “understanding intentions” into 3 levels (see below), and 3) extended the first perceptual-cognitive and thus rather passive line of development by adding in parallel a second, more action- and participation-oriented one, labelled “shared intentionality”. This latter line refers to different levels of actually engaging in collaborative activities which (following Bratman, 1989, 1992) are defined as: being mutually responsive, having a shared goal the partners are knowingly committed to, and participants coordinate plans (including potential role-reversal) (Tomasello et al., 2005).

This “double-trajectory” suggested for development now reads as follows: During the first 9 months infants are considered to understand people as animate agents (recognizing autonomously-produced action and being able to predict actions in a familiar situation, without knowledge of their internal structure) and are motivated for and hence engage in dyadic engagement, sharing behaviour and emotions with an animate agent, that is, merely “expressing emotions” and performing “behavioural turn taking”. By comparison with the diagram depicting full blown collaborative interaction (with the supposedly involved hierarchical representations superimposed on the interaction partners’ heads, see figure 2.1), Tomasello characterises this earlier phase as: “nothing inside the heads” (Tomasello et al., 2005, p. 681), thus apparently not considering these processes to be “cognition”, to involve any “knowledge” or “representation”, yet.
Figure 2.1: “Each partner’s conception of a collaborative activity in which a shared goal and joint intention (with complementary roles) are formed”. Tomasello and colleagues (2005) illustration of (visibly adult) shared intentionality in terms of coordinating individual knowledge. Note the (rationalist) depiction of the interaction partners as 2 completely symmetric large heads with small hands in the background, with all arrows linking hierarchical representations of increasing complexity and recursivity pointing from the top down. For details see Tomasello et al. (2005).

From 9/10 months, infants are considered to understand people as goal directed agents, that is, persistently pursuing a goal after failed attempts, accidents, and around obstacles, and perceptually monitoring their actions (infants show surprise when an agent takes a detour in the absence of an obstacle (Csibra, 2003), distinguish purposeful action from accident in responding more impatiently to unwilling compared to trying but unable action partners, and complete other’s action after only having seen failed trials (Behne, Carpenter, Call, & Tomasello, 2005)). Correspondingly, they are supposed to engage in triadic interactions with people understood as goal directed agents. Triadic interactions are here characterized as involving child, adult and some outside entity towards which they both direct their actions (and in particular their gaze), sharing goals (which they are jointly committed to) and perception (“joint perception”). This is inferred from infants (but not chimpanzees)
now attempting to reengage a recalcitrant adult in joint activities (Warneken, Chen, & Tomasello, 2006).Projected into the heads of the interaction partners (see figure 2.1) are now shared goals and perceptual monitoring.

**Between 12-15 months** infants are considered to understand people as intentional agents, choosing one plan over another and employing selective attention. Suggested evidence: 1) presented with a model pressing a button in an unusual, awkward way with their head, infants imitated this style: however they did so only if the adult-model’s hands are free (shown placed flat on table), but not when the shivering adult uses their head while hands are occupied holding a blanket (Gergely, Bekkering, & Kiraly, 2002); 2) 12 month old infants, robustly – upon hearing an adult exclaiming an ambiguous “Oh, wow! That’s cool! Can you give it to me?” – handed over that one out of three objects they were familiar with, which the adult did not know yet (Tomasello & Haberl, 2003).

Now infants get involved in collaborative engagement with joint intentions and attention (equated with Adamson and Bakeman’s coordinated joint engagement, see section 2.3.2), they are credited with understanding specific action plans, understanding the situation from a “birds-eye-view”, and are supposedly able to reverse roles.

Thus, interestingly, emerging results shifted the research program from its exclusive focus on understanding others based on visual observation from a distance to taking into account actual engagement in interaction, where the “uniquely human” aspects of cognition seem to reside. However, rather than taking a closer look at what happens in interactions in general, the peculiarities of human engagement in interaction are quickly translated

---

8 However, the Authors’ claim that the study shows infants’ understanding of “rational choice”, has by now been largely deconstructed, as the effects can be accounted for by more basic factors such as postural difficulty (Paulus, Hunnius, Vissers, & Bekkering, 2011) and perceptual saliency (Beisert et al., 2012).
into cognitive psychology’s language of hidden forces characterizing it as requiring: a (specifically human) **motive to share experience**, and a (specifically human) **knowledge of intentions being shared**. Thus the focus of investigation still remains (on) the individual; “social interaction” merely enters as content into his/her individual knowledge and motives.

Placing this new version of the theory into a larger comparative and evolutionary context, Moll and Tomasello (2007) propose what they call the "**Vygotskian intelligence hypothesis**", in contradistinction to the Machiavellian Intelligence hypothesis (Byrne & Whiten, 1989). The latter claims that the evolution of cognition, intelligence, and the brain was – particularly in primates – largely driven by (the requirements of) social competition. In contrast, the Vygotskyan intelligence hypothesis suggests that the evolution of human cognition and intelligence – in contrast to other primates – was primarily driven by (the requirements of) their uniquely collaborative ways of living. Tomasello and Carpenter (2007) go on to propose – again taking their cue from Vygotsky (see section 2.5.1) – how shared intentionality in the course of evolution “transforms” older primate forms of individual cognition into more powerful socio-cultural ones: it is argued to transform 1) gaze following to gain knowledge about objects into initiating truly intersubjective joint attention for the sole purpose of sharing, 2) social manipulation (gesturing to get things, proto-imperatives) to co-operative communication (pointing for sharing and getting attention, i.e. proto-declaratives, and for giving information), 3) group activity where everyone pursues their own goal to collaboration with shared goals and commitments, often apparently performed for its own sake, and finally 4) social learning, where information is gathered individually and unilaterally, to forms of cultural learning involving imitation, and in particular instructed learning, which later on is going to be performed collaboratively and often involves the establishment and enforcement of social norms.

In these and later publications Tomasello and colleagues lay out various evolutionary scenarios, how human forms of collaboration might have come about through several steps (Tomasello & Carpenter, 2007; Tennie, Call, & Tomasello, 2009; Tomasello, 2011):
1) A change in temperament: becoming more tolerant (in particular in food sharing situations) and more social-comfort-seeking. This might have been brought about by self-domestication, where aggressive individuals are selected against in small social groups or by the requirements of cooperative breeding.

2) The now more tolerant and pro-social individuals would do more things together (potentially fostered by situations like group hunts where collaboration in combination with food-tolerance creates a win-win situation for everyone) thus further developing the skills of shared intentionality and collaboration.

3) In a third step group level processes may have developed: such as establishing and enforcing group norms and constituting social institutions, processes which from then on were driven further by group competition.

In their later developmental research Tomasello and colleagues primarily focus on aspects of social cognition, which are related to these group processes and appear well after the first year, such as e.g. enforcing norms (Rakoczy, Warneken, & Tomasello, 2008).

2.4.4 Analysing and comparing the approach of Tomasello and colleagues to other approaches

Putting forth a wide-ranging framework spanning typical and atypical human development as well as great apes, tentatively accounting for evolution of culture in terms of (cultural) learning and cognition, Tomasello and colleagues provided 1) a palpable structure for multiple disciplines to discuss a range of phenomena connected to cultural learning, and 2) a grid of concrete hypotheses – reminiscent of Wittgenstein’s chessboard-image from the Tractatus (Wittgenstein, 1921) where each field represents a proposed “fact” to be tested to be the case or not – laying out the course of a systematic experimental research program. The drawback of putting forth such a speculative, largely conceptually guided – and hence logically intriguingly consistent – framework, however, is that it captures attention, binds the rational mind in its apparent completeness, potentially letting the world appear more alike to a logical grid than it might actually be, and
thereby aspects missing in the framework (e.g. interaction flow, affect, etc.) might easily be overlooked, making it difficult to see and adopt alternative perspectives and frames: e.g. conceiving of co-actors not as rational, equal agents co-ordinating their stepwise actions by rational thought in advance, but as potentially diverse partners continuously co-ordinating and negotiating their interaction throughout various levels on the flow as they go.

Tomasello and colleagues’ approach was widely welcomed for drawing attention to “sociality” within cognitive psychology, thus extending it at a point, where its exclusive focus on individual cognition would leave cognitive psychology at an impasse in how to account for higher cognitive functions. However, they introduced “the social” into cognitive psychology in terms compatible with its conceptual and methodological framework, simply adding “social cognition” as a new topic, a specific individual cognitive capacity to be investigated, with the level and unit of analysis still firmly on the individual. And social interaction and collaboration are conceived of as rational stepwise instrumental actions (also compare Piaget and Bates, section 2.2) understood by observation from a distance and planned and negotiated in advance by rational agents as equals.

A research program which frames social interactions in such rationalistic terms as driven by conceptual knowledge alone, runs the risk to be incomplete and even potentially misleading when addressing how rich embodied interactions between adults are coordinated and unfold moment-by-moment. In combination with a conception of development based on abruptly emerging new capacities and forms of knowledge (giving rise to a “revolution”), this approach renders it futile to even look into earlier interactions at all, thus making it difficult if not impossible to address and understand the developmental trajectories of social and cultural engagement.

A related problematic aspect of such a rationalist and concept-centred approach is that it tends to blind out anything which apparently does not fit into the concept-grid. A particularly striking example is arousal and affect:
E.g. when trying to conceptually delineate joint (visual) attention and applying this concept to real-life situations, Tomasello admits that it is difficult to determine from gaze alternations whether joint attention really is achieved, but notes that cases become much more convincing when occurring spontaneously and integrated with ongoing social interaction in appropriate ways. His examples of such “convincing” instances of joint attention almost exclusively happen to be situations where affect plays a crucial role, such as in social referencing, when children check for the adult’s emotional response to an object and adapt their own relation to it. Tomasello acknowledges that in general “the child does not look to the adult or attempt to secure their attention at random moments, but rather looks to the adult or points to an object when surprised by something, pleased with something, afraid of something, or in other socially meaningful situations” (Tomasello, 1995). The meaningfulness in those situations is based on sharing affect.

While the theory stays rather silent about the role of (shared) affect, many of Tomasello and colleagues’ experimental paradigms actually heavily depend on it: For example, in the Tomasello and Haberl (2003) study, the experimenter engages the infant exclaiming: “Oh, wow! That’s so cool! Can you give it to me?” while gesturing ambiguously to elicit the infant to hand over that of three objects equally familiar to the child, which is new to the adult. Tomasello and Haberl’s analysis and interpretation focuses exclusively on the kind of knowledge and understanding that can be inferred from the infants’ robust responses, without considering the potential role of affective communication even for understanding what is relevant in the situation (similar to Piaget). Generally, earlier social interactions before 9 months are considered to be dyadic and in Tomasello’s terms merely an “expression of emotion and behavioural turn-taking” having nothing to do yet with cognition or knowledge, which is also illustrated by Tomasello’s specification for a schematic depiction of co-actors at this phase as: “nothing [that is no representations] inside the heads” (Tomasello et al., 2005, p. 681).

When shared affect finally does enter Tomasello’s account it does so in the form of a uniquely human “motivation to share experience”, that is, as an
external force, an ad hoc deus ex machina, invoked to maintain the conceptual grid after testing of great apes and autistic children fails to establish clear-cut deficits of cognitive capacities of intention understanding. This motivation remains difficult to address within the limits of the approach, as long as affect continues to be seen as either as random content, not affecting or informative for cognitive structures or an inscrutable force pushing the cognitive system “from the outside” towards sociality; rather than as an integral part of the system, an organizing and connecting force, a container or scaffold for developing cognitive structures. We suggest that jointly participating in early affect-imbued social interactions, rather than being merely an expression of affect and automatic behavioural regulation, may help to learn about action structures and provide an important carrier structure for cognitive development (see section 6.7).

What role do social interactions and their cultural embeddedness play in Tomasello and colleagues’ approach, who investigate culture, albeit with a focus on individual social cognition and capacities? Their treatment of social interactions takes different (and not necessarily coherent) forms in different places: Generally, they repeatedly stress the crucial role of the cultural niche with its rich social interactions as the second, indispensable pillar, in addition to the capacities brought along by the individual, for culture to thrive (Tomasello et al., 1993; Tomasello, 1999; Tomasello et al., 2005). “Let us be very clear on this point. Participation in these interactions is critical. A child raised on a desert island would have all of the biological preparation in interactions involving shared intentionality, but because she did not actually participate in such interactions she would have nothing to internalize into perspectival cognitive representations. Ontogeny in this case is critical” (Moll & Tomasello, 2007). Looked at more closely, they seem to consider social interactions primarily as 1) a source of culturally specific empirical “content” (similar to Piaget), and 2) secondarily as forms of supporting activities for children.

1) Social interactions provide content a) from early on through a specific “habitus” (now drawing on Bourdieu) understood as a specific way of life
acquired by the inhabitants of a specific cultural niche, directing, and constraining content, but not considered to play a distinctive role for cultural learning in the narrow sense and hence disregarded⁹. And b) later on, cultural content conveyed in social interactions includes language or (conventional) object actions. 2) Social interactions particularly provide support for infants e.g. in the form of a) scaffolding, which, however, is considered to promote learning only on the individual rather than the cultural level, and b) in the form of active instruction; as this already operates on the cultural level it still does not necessarily help to bring about cultural learning, but may be utilized by children once they have achieved cultural capacities.

Generally, when Tomasello and colleagues stress the importance of social interactions, and how their richness can contribute to e.g. language learning, this cultural resource can only be “accessed” once the infant has already acquired the knowledge of others as intentional agents, and of shared intentionality, implying an already present capacity of joint perception and later joint attention.

*A child in Quine's (1960) famous "Gavagai" situation has no way of figuring out for itself the referent of a novel linguistic item. But in the real world young children learn new pieces of language almost always in highly contextualized, often routinized, mutually understood (i.e., intersubjective) nonlinguistic formats such as the feeding situation, diaper changing, book reading, taking a walk, or playing a game of peek-a-boo [...] These contexts are so replete with information about adult intentions - from the child's past experience in similar situations, as well as from the adult's current* 

---

⁹ The particular habitus into which a child is born determines the kinds of social interactions she will have, the kinds of physical objects she will have available, the kinds of learning experiences and opportunities she will encounter, and the kinds of inferences she will draw about the way of life of those around her. The habitus thus has direct effect on cognitive development in terms of the *raw material* with which the child has to work [...] (Tomasello, 1999, p.79)
direction of gaze, tone of voice, and specific behaviors toward objects- that they even support the acquisition of words for referents that are not perceptually present. (Tomasello et al., 1993)

Spelling out in more detail how this might work, Tomasello points out that for establishing reference and sharing information (whether via pointing or words), and in particular for learning (e.g. language or perspective taking etc.), it is crucial to participate in what he calls “joint attentional scenes”\(^{10}\) in which a socially shared reality formed of a shared subset of each partner’s perceptual world, provides the broader context for, and hence grounds, language and acts of reference.

In later texts “joint attentional frames” are primarily invoked as a crucial factor distinguishing apes from humans: their lack is used to account for apes’ apparent inability to utilize pointing, and, more generally, they serve as a demarcation line between instrumental communication and communication for the sharing of information considered unique to humans.

In order to successfully participate in creating a “joint attentional scene”, children need, according to Tomasello, to take an “outside” perspective on the situation and consider on one and the same conceptual plane: that which is shared, the adult, the child him/herself, as well as the (interchangeable) roles inherent in the situation. So, again this already requires a lot of complex “cognition” to be in place in order to benefit from social interactions.

But how do children get there in the first place? And what about early socio-cultural interactions? Do they play any role?

As already mentioned above, Tomasello and colleagues acknowledge “habitus”, but quickly dismiss it as not interesting for cultural learning in the narrow sense. In addition, we indirectly learn about the importance of socially guiding attention at an early age, or in Tomasello’s words:

\[ \text{__________________________} \]

\(^{10}\) In other places they are also referred to as joint attentional “frames”/“formats” or “common ground”, taking up earlier labels by other researchers such as Goffman, Bruner and Clark.
“socialisation of attention”, as it is proposed as the potentially distinctive factor setting apart enculturated from wild apes, enabling enculturated apes to e.g. perform some basic forms of imitation, nota bene not by helping them develop social cognition, but by extending their individual capacities (Tomasello et al., 2005).

While later social interactions are considered crucial for cultural transmission, the question, whether early social interactions play a role in human infants’ development of social motivation or knowledge (understanding others as intentional agents or shared intentionality) seems to remain inconclusive: confirmed in some places (Tomasello, 1995), left open or denied in others (Tomasello, 1999): “We do not know exactly how much of an understanding of intentional action is necessary for children to participate in collaborative activities. And conversely, we do not know whether the kinds of collaborative activities that exist in cultures before children are born are a necessary or only a facilitative component in the ontogenetic process – or whether they play no effective role at all at the outset (though clearly they play a crucial role later)

Taking a closer look reveals that Tomasello acknowledges that some amount of social interaction is necessary – (not unlike Chomsky’s view that a self-maturing language acquisition device (LAD) still does need input/exercise to mature):

“It is very unlikely, in our view, that a human or ape kept in social isolation for the first year of life would suddenly understand others as goal-directed or intentional agents on its initial encounter with them; presumably, the developmental pathway for understanding intentional action depends on species-typical social interactions early in ontogeny. This does not necessarily mean, however, any specific experiences.” (Tomasello et al., 2005, my emphasis)

But he further claims – referring to the lack of research demonstrating cultural variation – that particular ways of social interaction do not make a relevant difference and hence do not hold any explanatory value:
“As noted above in this section, there has been almost no research – not even training studies or correlative studies – that establishes a solid relationship between any kind of particular social experience infants might have and individual differences in the unfolding of this developmental pathway. In the absence of such studies, we might tentatively conclude that this is a very robust, heavily canalized ontogenetic pathway in humans that emerges in all “normal” human environments.” (Tomasello et al., 2005)

Thus, while the general notion of “social interaction” does play an active – and for the argument effective – role within the theoretical part of Tomasello’s approach, supposedly constituting the second pillar – in addition to the capacities brought along by the individual – for culture to work, in his concrete statements he seems to be downplaying its role in early development – reflecting the tension inherent in looking at culture from a cognitive, capacity-based perspective. The notion of “social interaction” does not play any effective role in the empirical part of the approach, but remains basically unaddressable, and its relevance untestable in the empirical program focussing exclusively on individual capacity, thus actually leaving essential parts of the theory empirically ungrounded and suspended in mid-air.

The approach of Tomasello and colleagues shares with other cognitive approaches, primarily focusing on the individual and his/her capacities, (what sometimes seems to amount to) an outright disdain for social learning by training, as they construe a supposed dichotomy between mindless passive associative learning by drill and active learning by insight and understanding (compare Piaget, Bates, section 2.2.1-2 as well as section 6.7). Thus Tomasello dismisses learning by what he calls “ontogenetic ritualization”¹¹, a form of learning ascribed to chimpanzees and very young

¹¹ In earlier texts Tomasello refers to Smith (1977) with respect to this notion. Discussing the process of “formalization” i.e. the “specialization of behavior to become informative” in the context of displays in animal communication, Smith (1977) articulates the need to complement the notion “ritualization”, which in biology had come to mean genetically
children, which is socially trained, develops within particular social relations, thus showing individual variations, thus is not conventional, and hence – in contrast to insightful rational and faithful imitation – is not considered relevant for cultural learning and transmission by Tomasello and colleagues:

In ontogenetic ritualization a communicatory signal is created by two organisms shaping each other’s behavior in repeated instances of a social interaction. For example, an infant may initiate nursing by going directly to the mother's nipple, perhaps grabbing and moving her arm in the process. In some future encounter the mother might anticipate the infant's desire at the first touch of her arm and so become receptive at that point – leading the infant to abbreviate its behavior to a touch on the arm with response waiting (cf. Tinbergen 1951, on 'intention movements'). Note that there is no hint here that one individual is seeking to reproduce the behavior of another; there is only social interaction that results eventually in a communicative signal (Tomasello, 1996). “This is presumably the way that most human infants learn the “arms-over-head” gesture to request that adults pick them up, that is, first as a direct attempt to crawl up the adult’s body, and then, as the adult anticipates their desire and picks them up, as an abbreviated, controlled phylogenetic change of behavior, with a notion capturing learned change and for this suggests to again take up the notion of “conventionalization”. In earlier texts Tomasello uses the notion “individual conventionalization” (e.g. Tomasello, 1994) where the qualifier “individual” indicates the default meaning of conventional to be non-individual, a little later he uses both “ritualization (or conventionalization)” (Tomasello 1996a), and after that mostly “ontogenetic ritualization” (e.g. Tomasello 1996b, 1999) walking back Smith’s move to use “conventionalization” but instead using ritualization with the qualifier “ontogenetic” to distance it from the default phylogenetic interpretation of ritualization in biology, thus keeping the term “convention” completely out of this, presumably so as to reserve it for aspects operating beyond particular inter-individual relationships at an impersonal level of social groups.
Ritualized version of this crawling activity performed for communicative purposes only (Lock, 1978).” (Tomasello, 1999)

However, human infants and caregivers frequently enact jointly performed routines in which learning by training may occur (Negayama et al., 2015; Rączaszek-Leonardi, Nomikou, & Rohlfing, 2013; Reddy, Markova, & Wallot, 2013); rather than dismissing those interactions outright, we should look at them in more detail and investigate the potential role they may play in the development of social understanding and joint acting. Indeed, learning by training and learning by understanding do not seem to be mutually exclusive, on the contrary, jointly performed routines with their frequently repeated, affect-imbued, and highly structured action sequences may provide a rich source and container for understanding and mastery of actions and people to develop in the first place (see section 6.7).
2.5 Accounts IV: Cultural Accounts (Vygotsky, Rodríguez, Zukow-Goldring)

Cultural approaches to development are convinced that processes of learning, understanding, mastery and participation – rather than originating in the individual and to be accounted for exclusively with the individual as the sole unit of analysis, let alone in terms of individual capacity – are a product of socio-cultural history. Therefore these approaches “zoom out” with regard to their unit of analysis and take into account various aspects of the larger culturally shaped and culture-manifesting context, emphasising and actively investigating the role or caregivers, siblings, artefacts, and cultural routines.

This chapter's analysis will focus on: 1) Lev Vygotsky, as his work not only serves as the major reference point for socio-cultural approaches to trace themselves back to, but has also had tremendous influence on numerous approaches, many of the ones discussed here, cultural (see below) and beyond (e.g. Bruner tradition, and see also Tomasello). 2) Christiane Moro’s and Cintia Rodríguez’ explorations into triadic interactions to find out how infants learn the conventional meaning of objects, and 3) Patricia Zukow-Goldring’s Socio-cultural Realism, spelling out in detail how infants learn to participate in close dialogue with more proficient members of a culture through “educating attention” and “assisted imitation”.

The last two approaches were selected as they are particularly close to the topic of triadic interaction as well as may be particularly suited as boundary objects for interdisciplinary dialogue: their units of analysis and conceptual frameworks seem relatively compatible with more individually oriented approaches of cognitive psychology as they both - drawing on additional theoretical sources outside the cultural-psychology tradition - in a way seek to bridge the gap between psychology and the socio-cultural traditions (see below).

However, there are numerous other cultural approaches beyond the scope of the present chapter which should not be overlooked when seeking
to investigate and better understand infant-caregiver-object interactions: in particular Barbara Rogoff’s approach investigating “guided participation”, working out multiple (non-exclusive, but interrelated) levels or better perspectives of analysis for thoroughly understanding participation in cultural activities (Rogoff, 1990; Rogoff, Mistry, Radziszewska, & Germond, 1992; Rogoff, 1995; Rogoff et al., 2007), or Valsiner’s cultural historical study (drawing on activity theory and Lewin’s field theory) on the development of the activity of feeding/eating (Valsiner, 1987), etc.

2.5.1 Lev Semyonovich Vygotsky

2.5.1.1 Background, interests, and theoretical framework

Vygotsky – who had studied law, history and literature/semiotics – was interested in the specifically human ways of thinking and relating to the world (e.g. through labour), which he called “higher mental/cognitive?” functions and which he considered as formed (and hence potentially formable) in a dialectic socio-historically process. It was in the context of the Russian revolution that these questions gained a particular urgency as well as practical relevance, as Vygotsky and his colleagues were actively involved in building a new psychology and education aimed at contributing to building a new, socialist society.

The young Vygotsky grappled with and sought to free himself and Psychology from the (established) Psychological schools in Russia (Introspectionism, predominant Reflexology, and combinations thereof) and abroad (American behaviourism not unlike Reflexology), and, in particular, to bridge the gulf which he traced back to their one-sided and incompatible perspectives: between (passive) stimulus-response reactions (central concept of predominant Reflexology) and human volitional, conscious action – or, more generally, between “behaviour” and “mind”: “This is the other half of the same dualism. Previously we had mind without behaviour. Now we have behaviour without mind. In both cases, we have “mind” and “behaviour” understood as two distinct and separate phenomena” (Minick, 1987, p. 19)
2.5.1.2 Bridging the gap between passive stimulus-response behaviour and human volitional and conscious higher mental functions by postulating the use of signs as “psychological tools” or “artificial stimuli” to mediate and actively control behaviour.

To bridge the gap, Vygotsky theorized that humans, rather than being forced to only respond automatically to conditioned stimuli as is characteristic of non-human animals, are also able to actively and volitionally regulate their own behaviour by instrumentally using “signs” (in particular in the form of speech), presenting them as artificial stimuli to themselves, which then mediate their behaviour in analogy to physical tools as “psychological tools”, a new form of “instrumental act” which characterizes higher cognitive functions. "For higher functions, the central feature is self-generated stimulation, that is, the creation and use of artificial stimuli which become the immediate causes of behaviour.” (Vygotsky, 1978a, p. 39)

Vygotsky characterizes the difference between physical tools (analysed by Engels) and “signs” as “psychological tools” (which Vygotsky now develops in analogy to Engels) as a difference in direction from “other-/object-directed” to “self-directed”:

The tool's function is to serve as the conductor of human influence on the object of activity; it is externally oriented; it must lead to changes in objects. It is a means by which human external activity is aimed at mastering, and triumphing over, nature. The sign, on the other hand, changes nothing in the object of a psychological operation. It is a means of internal activity aimed at mastering oneself; the sign is internally oriented. These activities are so different from each other that the nature of the means they use cannot be the same in both cases. (Vygotsky, 1978a, p. 55)

Because this auxiliary stimulus possesses the specific function of reverse action, it transfers the psychological operation to higher and qualitatively new forms and permits humans, by the aid of extrinsic stimuli, to control their behavior from the outside. The use of signs leads humans to a specific structure of behavior that breaks away from biological development and
creates new forms of a culturally-based psychological process. (Vygotsky, 1978a, p. 40)

The use of artificial means, the transition to mediated activity, fundamentally changes all psychological operations just as the use of tools limitlessly broadens the range of activities within which the new psychological functions may operate. In this context, we can use the term higher psychological function, or higher behavior as referring to the combination of tool and sign in psychological activity. (Vygotsky, 1978a, p. 55)

2.5.1.3 Vygotsky’s mediation triangle and the infant-caregiver-object triad

![Diagram of Vygotsky’s mediation triangle]

Figure 2.2: Vygotsky's depiction of mediated action, details see text, redrawn from Vygotsky (1978)

In Vygotsky's graphical sketch (figure 2.2) of the sign-mediated instrumental act we recognize another three-pole structure, akin to the depictions of the infant-caregiver-object triad, both of which are considered to mark the transition into culture. It shows a “stimulus” (equivalent to either the object or the caregiver in the infant-caregiver-object triad), a “response” (the equivalent of the child and his/her actions), only here we do not have a human partner but a product of culture, a sign or artificial stimulus X, mediating the interaction. The depiction illustrates processes, which find their analogy in the triad as well: it shows how by mediation of and engaging via a third pole, distance is created, breaking up the tight connection between stimulus and response, making them discernable and, by inhibiting the immediate response, creating more action flexibility (also reminiscent of Piaget’s instrumental act).
2.5.1.4 Socio-historical origins of sign mediation and the process of internalization

Vygotsky regards speech as the primary mediating sign/psychological tool, and investigated in detail its role in children’s thinking. Noting that “Signs and words serve children first and foremost as a means of social contact with other people” (Vygotsky, 1978a, p. 28) he proposed that “subsequently, the individual "begins to apply to himself the same forms of behavior that were initially applied to him by others" (quoted in Minick, 1987, p. 21), and, conversely, “the child begins to master his surroundings with the help of speech” – foreshadowing Bates notion of “social tool use” – “prior to mastering his own behavior” (Vygotsky, 1978a, p. 25). Thus he came to regard the sign mediated action characteristic of higher mental functions as originating out of social interaction, from where it later gets “internalized” to intrapersonal thinking: thus speech is first used in conversation, then as self-directed vocalization for problem solving (countering Piaget’s interpretation of ego-centric speech), and finally fully internalized as silent inner speech in thinking.

An interpersonal process is transformed into an intrapersonal one. Every function in the child's cultural development appears twice: first, on the social level, and later, on the individual level; first, between people (interpsychological), and then inside the child (intrapsychological). This applies equally to voluntary attention, to logical memory, and to the formation of concepts. All the higher functions originate as actual relations between human individuals (Vygotsky, 1978a, p. 57).

The internalization of socially rooted and historically developed activities is the distinguishing feature of human psychology, the basis of the qualitative leap from animal to human psychology. As yet, the barest outline of this process is known (Vygotsky, 1978a, p. 57).

Thus a first integration of the stimulus-response behaviour with conscious-volitional thought into one framework is achieved via appeal to the origins of thinking in social interactions and hence a product of socio-cultural history.
2.5.1.5 Vygotsky’s puzzling lack of empirical investigations/studies of social interaction

However, this social origin remains a theoretical claim, supported with a few anecdotal or common-sense examples, since, as Minick notes: “Noticeably absent in this work was any attempt to carry out empirical research on the development of mental processes in social interaction” (Minick, 1987, p. 22).

Rather than looking at social interaction per se, Vygotsky’s own systematic experimental studies remain confined to how products of culture such as signs (in form of speech or pictures etc.) are used by children of different ages and degrees of experience in individual problem solving activities, and how they impact on the children’s thinking.

According to Minick, this may have been due to 1) their primary concern being to show in principle the difference in quality and origin between natural and cultural thinking rather than their earliest development and 2) that for the semiotically well versed Vygotsky the conceptual framework used back then, based on the concept of conditioned reflexes (albeit extended by artificially mediated ones) ultimately did not seem to be sufficient to address a complex phenomenon like meaningful language.

2.5.1.6 Claiming two separate lines of development with different roots in ontogenesis to motivate the demand for a new method – recognizing a familiar narrative

Vygotsky discussed this methodological question in detail in his 1930 Problems of Method and explicitly puts himself in line with Wundt who considered experimental research methods only applicable to elementary function while for higher mental functions turned to developing an alternative approach (his Völkerpsychologie). In this paper, the close connections between theory, methodology, and politics also become palpable in Vygotsky’s reference to Engels:

The keystone of our method, […] follows directly from the contrast Engels drew between naturalistic and dialectical approaches to the understanding of human history. Naturalism in historical analysis, according to Engels,
manifests itself in the assumption that only nature affects human beings and only natural conditions determine historical development. The **dialectical approach**, while admitting the influence of nature on man, asserts that man, in turn, affects nature and creates through his changes in nature new natural conditions for his existence.' This position is the keystone of our approach to the study and interpretation of man's higher psychological functions and serves as the basis for the new methods of experimentation and analysis that we advocate. All stimulus-response methods share the inadequacy that Engels ascribes to naturalistic approaches to history. Both see the relation between human behavior and nature as unidirectionally reactive. My collaborators and I, however, believe that human behavior comes to have that "transforming reaction on nature" which Engels attributed to tools. We must, then, seek methods adequate to our conception. In conjunction with new methods, we also need a new analytic framework. (Vygotsky, 1978a, pp. 60–61; emphasis ours)

It is against this background of a naturalistic cause-effect framework, experienced as all-encompassing and rendering humans passive victims of inevitable historic processes, that Vygotsky's next move has to be seen, effectively separating and hence freeing humans from nature, narrowly determined by mechanistic laws:

Responding to Köhler's and Bühler's empirical findings that some kind of practical intelligence can already be found in language-less apes and in a similar form in pre-verbal infants and their resulting claim that hence practical intelligence/"Werkzeugdenken" and language seem to constitute different, relatively independent systems, Vygotsky counters that the tool-use of apes and the tool use of children already capable of language [let alone the adult shaping his world through labour] are not expressions of the same kind of intelligence only differing in degree, but rather represent two different lines of development originating from separate roots in ontogenesis: a natural line (subject to stimulus-response laws and hence addressable by a naturalistic framework) and a second, cultural line, which **cannot** be reduced to stimulus response laws, but – being part of socio-cultural history – necessitates a different framework and methodology.
Vygotsky, somewhat confusingly, conceived of the cultural line as itself emerging when two hitherto separate lines of development: “natural” pre-verbal thinking (associated with some basic form of intelligence) and “natural” pre-intellectual speech (associated with emotion) converge in children around 2 years of age (while remaining separate in apes and continuing only on the natural line of development), transforming the child’s natural thinking and speech into their intertwined cultural forms: giving rise to meaningful speech and verbal thought and new forms of learning.

“1. As we found in our analysis of the phylogenetic development of thinking and speech, we find that these two processes have different roots in ontogenesis.

2. Just as we can identify a "pre-speech" stage in the development of the child's thinking, we can identify a "pre-intellectual stage" in the development of his speech.

3. Up to a certain point, speech and thinking develop along different lines and independently of one another.

4. At a certain point, the two lines cross: thinking becomes verbal and speech intellectual.”

“conclusion: the most significant moment in the course of intellectual development, which gives birth to the purely human forms of practical and abstract intelligence, occurs when speech and practical activity, two previously completely independent lines of development, converge.” (Vygotsky, 1978a, p. 24)

Here we clearly recognize a variant of today's familiar developmental narrative, only in this older version one of the lines is labelled with a slightly narrower “speech”, and the convergence and resulting transformation is supposed to take place later, only around 2 years of age resulting in “meaningful language and verbal thinking”, whereas in the modern version “natural speech” has been replaced by “(dyadic) social engagement” (and social knowledge) and the point of convergence and transformation has been pulled downward(s) to 9 months, giving rise to “cultural learning” “through the other person” (in Tomasello's 1995 words). Interestingly, we
now see that Vygotsky's original separation of lines arguably can be seen at least in part as serving the function of justifying the rejection of a limiting stimulus-response framework and demanding a new method. In today's narrative, the twofold separation (nature vs. culture, speech vs. practical intelligence) remains firmly in place (dyadic vs. triadic, social vs. object engagement), though the practical, methodo-ideological functions it had arguably served in Vygotsky's version, and indeed the awareness thereof – have mostly been lost.

Separating out a specific area of research, declaring its independence, and seeking to develop an additional specifically tailored research methodology (as practiced here by Vygotsky or previously Wundt) is one specific strategy to deal with a framework experienced as inadequate. The advantage of this strategy is that it momentarily allows to productively continue and expand work into new territories while ideally minimizing the friction with the opposed established framework. The drawback, however, is that in the long run it leaves Psychology and its related disciplines with a seemingly unbridgeable conceptual and methodological fault line running right through the middle of it, separating nature from culture, behaviour and affect from (rational) higher cognition, learning by training from learning based on knowledge and understanding, etc., also playing out in an illustrative way in investigating the development of triadic interaction, which we will explore in more detail in the discussion in section 2.8.

2.5.1.7 Vygotsky’s take on children’s way of relating to objects

Indeed, children's way of relating toward objects (as towards their environment in general) is changed by the proposed cultural transformation, which – according to Vygotsky – occurs around 2 years of age once practical thinking and speech have converged: children now conceive of objects as part of a socially-constituted networks of meanings [which Vygotsky considers primarily linked to language]:

A special feature of human perception – which arises at a very young age – is the perception of real objects. This is something for which there is no analogy in animal perception. By this term I mean that I do not see the
world simply in color and shape but also as a world with sense and meaning. I do not merely see something round and black with two hands; I see a clock and I can distinguish one hand from the other. Some brain-injured patients say, when they see a clock, that they are seeing something round and white with two thin steel strips, but they do not know it is a clock; such people have lost their real relationship with objects. These observations suggest that all human perception consists of categorized rather than isolated perceptions. (Vygotsky, 1978a, p. 33)

Also for actively engaging with objects (whether concrete or abstract objects), and in particular when encountering problems with the object of engagement, children who enact cultural ways of engaging have a whole range of flexible interaction opportunities at their disposal, enhancing their chances to solve the problem: “In summary, children confronted with a problem that is slightly too complicated for them exhibit a complex variety of responses including direct attempts at attaining the goal, the use of tools, speech directed toward the person conducting the experiment or speech that simply accompanies the action, and direct, verbal appeals to the object of attention itself” (Vygotsky, 1978a, p. 30). Here we encounter in one paragraph all the various ways of engaging with objects (now at the disposal of a slightly older child), which were emphasized in one or the other approach discussed: Piaget’s instrumental act, Bates’ et al.’s social tool use, Vygotsky’s self-directed speech (taken up again by Rodriguez as well as by Tomasello), and, interestingly, one additional way of engagement: “direct verbal appeal to the object”, typically used for relating to the social world: “Our research has shown that very small children solve problems using unique mixtures of processes. In contrast with adults, who react differently to objects and to people, young children are likely to fuse action and speech when responding to both objects and social beings” (Vygotsky, 1978a, p. 29). This complicates current notions of distinct people vs. object engagement.

However, how children come to understand objects conventionally and engage with objects in the complex and flexible ways described remains unclear. In the often-quoted final paragraph of “Tool and symbol in child development”, Vygotsky writes:
From the very first days of the child’s development his activities acquire a meaning of their own in a system of social behavior and, being directed towards a definite purpose, are refracted through the prism of the child’s environment. **The path from object to child and from child to object passes through another person.** This complex human structure is the product of a developmental process deeply rooted in the links between individual and social history. (Vygotsky, 1978a, p. 30, emphasis mine).

A formulation which – with intriguingly capturing imagery – suggests the important role of adults and social interaction in developing object understanding and use, but seems to stand alone at the end of Vygotsky’s chapter as a rather isolated theoretical thought, and does not further specify how this is concretely achieved. This is also one of the passages where certain inconsistencies and tensions in Vygotsky’s writings become apparent. On the one hand, this passage is located in a context referring to slightly older children and Vygotsky states that the natural lines converge into culture only around 2 years of age. Whereas here he notes that the infant is from the beginning embedded in a system of social behaviour.

**2.5.1.8 Specific forms of cultural learning – the zone of proximal development and imitation**

In another manuscript, he proposed forms of learning typical for human children, which may address our question about how conventional object understanding and mastery develop. In particular, he directs our attention to the role of what he calls the **zone of proximal development** in children’s learning, where they can learn what is just outside their current abilities by utilizing the help of an adult guide:

[The zone of proximal development] is the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers (Vygotsky, 1978a, p. 86)

He further points out that to understand what is going on in this learning process might require a reconsideration of “imitation”, which at the time
was considered as rather mindless parroting. This conception is also illustrated in Köhler's experiments, referred to by Vygotsky, in which Köhler tried to find out, whether the apes were learning autonomously or “merely copying” what someone else was doing, and which ended up indicating that the apes were only imitating what they were already able to do. Vygotsky now comments:

*Köhler failed to take account of an important fact, namely, that primates cannot be taught (in the human sense of the word) through imitation, nor can their intellect be developed, because they have no zone of proximal development. A primate can learn a great deal through training by using its mechanical and mental skills, but it cannot be made more intelligent, that is, it cannot be taught to solve a variety of more advanced problems independently. For this reason animals are incapable of learning in the human sense of the term; human learning presupposes a specific social nature and a process by which children grow into the intellectual life of those around them.*

Children can imitate a variety of actions that go well beyond the limits of their own capabilities. Using imitation, children are capable of doing much more in collective activity or under the guidance of adults. This fact, which seems to be of little significance in itself, is of fundamental importance in that it demands a radical alteration of the entire doctrine concerning the relation between learning and development in children” (Vygotsky, 1978a, p. 88).

In these lines we readily recognize where Tomasello – after already taking up the ideas of two separate lines of development, their convergence resulting in specifically human forms of cultural learning, including imitation and internalization – takes part of the mission of his research programme from Vygotsky, and where – in line with his cognitivist psychology perspective – it is heading: as Tomasello’s programme, seeks to investigate this “special social nature” presupposed in imitation learning, and to find the “capacities” considered to be developed out of the “natural or biological” lines which make this “special social nature” and cultural imitation learning possible.
However, another quotation makes clear that Vygotsky certainly was not prepared to cede all developmental agency to individual capacities but considered learning and social interactions to be intertwined in both directions:

*We propose that an essential feature of learning is that it creates the zone of proximal development; that is, learning awakens a variety of internal developmental processes that are able to operate only when the child is interacting with people in his environment and in cooperation with his peers. Once these processes are internalized, they become part of the child's independent developmental achievement (Vygotsky, 1978a, p. 90)*

However it still remains unclear, how infants come to be able to imitate and learn, and what this “specific social nature” necessary for imitation learning exactly entails. Also, the age of the children Vygotsky is talking about again remains open, but most certainly he is speaking about children who have already undergone the cultural transformation and are already capable of language. This only moves the challenge forward in time: how, why, or through which processes does the convergence of speech and practical intelligence – supposedly leading to this transformation – come about, also remain unclear.

In “Mastery of Memory and Thinking” Vygotsky makes some tentative, suggestive, but also, at least in the translation, slightly cryptic remarks:

*This means that sign-using activity in children is neither simply invented nor passed down by adults; rather it arises from something that is originally not a sign operation and becomes one only after a series of qualitative transformations. (Vygotsky, 1978a, p. 46)*

So he rejects the notion that signs/language are a taken-for-granted part of human inherent “logos/reason”, which simply need to be discovered as assumed by some earlier accounts of language acquisition (e.g. W. Stern), as well as the idea that signs are simply and directly “transferred” probably by instruction or simple copying and points out instead that we have to take into account a more complex process of “qualitative transformations” going on. Again leaving in the dark how these transformations go about.
If we include this history of higher psychological functions as a factor in psychological development, we must arrive at a new concept of development itself. Within a general process of development, two qualitatively different lines of development, differing in origin, can be distinguished: the elementary processes, which are of biological origin, on the one hand, and the higher psychological functions, of sociocultural origin, on the other. The history of child behavior is born from the interweaving of these two lines. The history of the development of the higher psychological functions is impossible without a study of their prehistory, their biological roots, and their organic disposition. The developmental roots of two fundamental, cultural forms of behavior arise during infancy: the use of tools and human speech. This alone places infancy at the center of the prehistory of cultural development (Vygotsky, 1978a, p. 46).

These passages spark some suggestive ideas, but again leave many things open, thus opening up spaces and potential paths in multiple different directions.

Tomasello, interpreting Vygotsky from a cognitivist perspective, goes for the “biological roots” and “organic disposition”, and investigates the early supposedly exclusively “natural” lines of development looking for individual capacities (to emerge) which could account for the “convergence of the lines” and the resulting cultural transformation, or in Tomasello’s words “revolution”. That is, he investigates the in his opinion overlooked aspect, what capacities individuals bring to cultural forms of interaction and learning. In contrast, cultural approaches move in a different direction, seeking to spell out other parts Vygotsky himself was only hinting at: e.g. how socio-cultural interactions exactly interact with development, or in more general, how cultural processes unfold in a culturally structured environment through ongoing socio-cultural interactions. Some of them (see e.g. Rodriguez and Zukow below) also seek to extend these investigations to early ages, as proponents of these approaches typically relativize the formulation from Vygotsky’s middle period, that early development is strictly “natural”: 
After all, even in children at the very earliest ages mental processes are being formed under the influence of verbal social interaction with the adults who surround them. Consequently these mental processes are not ‘natural’” (Leont'ev and Luria quoted in Wertsch, 1988, p. 44)

And indeed in the latest phase of his work, Vygotsky – as laid out by Minick (1987) – apparently sought to understand how changes in activity practices (and in particular social practices and ways of relating), go hand in hand with changes in thinking, how – from the beginning – they affect each other, or – seeking to overcome the “personality”-“environment” distinction altogether – how they are actually two (only conceptually separated) sides of the same ongoing process of experience in transformation:

“Fundamental to all Vygotsky's work during this period was the notion that psychological processes develop in connection with transformations in behavior, whether these transformations are the consequence of biological maturation [...], the internal development of the child's own activity [...], or the inclusion of the child in new forms of social interaction and social practice [...]” (Minick, 1987, p. 26)

The social situation of development, which is specific to each age, strictly defines the child's entire mode of life, his social existence... Having clarified the social situation of development that forms at the beginning of a given age and defines the relationships between the child and the environment, we must then clarify how the new formations characteristic of this age necessarily arise and develop from the child's life in this social situation” (quoted in Minick, 1987, p. 30).

We have inadequately studied the internal relationship of the child to the people around him... We have recognized in words that we need to study the child's personality and environment as a unity. It is incorrect, however, to represent this problem in such a way that on one side we have the influence of personality while on the other we have the influence of the environment. Though the problem is frequently represented in precisely this way, it is incorrect to represent the two as external forces acting on one another. In the attempt to study the unity, the two are initially torn apart. The attempt is then made to unite them” (quoted in Minick, 1987, p. 32).
“The child's experience is the kind of simple unit of which it is impossible to say that it is the influence of the environment on the child or a characteristic of the child himself. Experience is a unit of personality and environment as they exist in development.... Experience must be understood as the internal relationship of the child as an individual to a given aspect of reality” (quoted in Minick, 1987, p. 32).

Thus Vygotsky has come full circle in his vision to come to an integrative understanding of artificially torn apart “behavior” and “mind”. However in psychology researchers have had and will have a long way to go and a busy time to put such a vision into practice to more thoroughly understand how we jointly create, sustain and transform the world and ourselves in continuous inter-/transaction. The approaches discussed below can be seen as efforts towards such a vision.

2.5.2 Moro and Rodríguez, and Mureno-Núñez: How do infants acquire the social/conventional meanings and uses of objects?

2.5.2.1 Background, interest and theoretical framework

Standing in the tradition of Piaget's Geneva School and interested in materiality, semiotics, as well as education, Moro & Rodríguez seek to explore how infants come to master the conventional meanings and uses of objects. They criticize developmental psychology for largely taking objects for granted, that is, for considering object meaning as self-evident (“natural”): Infants are considered to interact with objects from early on merely as part of their physical world, and supposedly learn about objects mostly spontaneously through interaction with them, where the object discloses its meaning by affording one or another action (Gibson).

Instead Moro and Rodriguez insist that objects are by no means “obvious”, but are on the contrary “opaque”, that their meaning and function are not directly visible - “objects do not say what they are, or what they are for” (now countering Gibson) - that “besides having physical properties objects also have [opaque] functional properties of use” which are subject to social convention, that is, they are part of culture. Consequently, infants do not spontaneously start to appropriate the canonical uses of objects, rather
these canonical uses, as Moro and Rodríguez put it: “need to be mediated by additional ‘signs’” and infants need adults as guides to “provide them with semiotic mediators” (Rodríguez, 2007).

Whereas, in the case of language, it was established in the course of the pragmatic turn that understanding the meaning of a word is rooted in understanding its use and function in interaction (Wittgenstein, Austin, Bruner) – initiating a whole field of research to figure out how this is achieved by young infants (see e.g. Bates, Sugarman, Adamson section 2.2.2-3, and 2.3.2) – Rodríguez (2007) wonders why this has not been the case with the “meaning” of objects with regard to their conventional uses in interaction.

They point out that even Vygotsky, who famously stressed the importance of cultural mediation, came to neglect the role of conventional object meaning before language. According to his framework, infants, when interacting with objects, first exercise their “natural” practical intelligence (shared with our ape relatives), which only after the onset of language is finally reorganized and transformed into its uniquely human cultural forms (again refining what was at the time counted into the human logos-culture-complex). Thus – according to Vygotsky – all object interactions based on practical intelligence prior to their reorganisation through language are explicitly not part of culture (Moro & Rodríguez, 2004; Rodríguez, 2007; Moro, 2011). However, despite this distinction between early physical tool and later cultural sign, Vygotsky’s analysis of the sign and how it mediates cognition in principle opened up a pathway to perform a semiotic analysis for earlier, preverbal processes as well (Moro, 2011).

Now Moro & Rodríguez moved one further step and responded to and refined Vygotsky’s framework by in effect separating “conventional object meaning and use” from the logos-language-culture complex, suggesting that language already rests on a network of these conventionally and hence culturally established object meanings and practices broader than language and occurring prior (and hence transforming thinking already prior) to language.
Assuming that objects are indeed opaque and given that the child “towards the end of the first year [and prior to language] [...] begins to use objects according to their social and conventional everyday-life functions.” (261), the question arises how infants acquire these meanings prior to language or in Rodríguez’ terms: “At what point and through which semiotic processes do objects become signs of their conventional use?” (Rodríguez, 2007).

They criticise existing research on triadic interactions (Trevarthen & Hubley, Bates, Tomasello) for only telling part of the story when focussing on triadic interactions from 9 months when infants are already communicating “intentionally”, missing out on previous triadic interactions (from which the later ones, so Rodriguez, “necessarily emerge”). Whereas Tomasello et al. consider such earlier interactions as “dyadic”, the caregivers actually establish and sustain triadicity in didactic interactions, as well as demonstrate object use from early on; this key role of the adult as guide is absent from standard accounts of early development.

We seem to forget that, from the moment children are born, adults use objects in their daily chores, in a space of public and everyday uses children continuously attend to and participate in these “performances” well before they themselves are able to perform the first conventional uses of objects. (Rodríguez, 2007)

Moro’s and Rodríguez’ conception of triadicity thus differs considerably from the more cognitive approaches: The latter focus on the infant’s individual capacities and accordingly define triadicity in terms of the infant’s actions, requiring an active “co-ordination between people and objects” on the part of the infant for an interaction to be counted as triadic. It has to be pointed out, however, that the exact nature of this co-ordination despite extensive treatises ultimately remains unclear and hence the criteria for attributing active coordination to infants remain theoretically underspecified – while, at the same time, for practical purposes infant triadic engagement is typically operationalized in terms of infant-initiated, (uncued) gaze alternation, hence being methodologically over-specified, focusing on the singular aspect of gaze alternation which is neither
necessary nor sufficient for defining joint attention or co-ordinated engagement; as gaze alternations by themselves do not necessarily indicate joint attention/co-ordinated engagement and there may be episodes of co-ordinated engagement/joint attention utilizing other modalities apart from gaze (Carpenter et al., 1998; see section 2.8: conclusions and outlook).

In contrast, Moro and Rodríguez use “triadic” in a much broader sense: 1) they explicitly include interactions in which it is exclusively the caregiver who actively establishes and sustains triadicity, and indeed criticise cognitive approaches for “missing” these kind of triadic interactions, which they claim to be particularly crucial as a necessary foundation for later (infant-initiated) triadic interactions. (It has to be noted, though, that the criticized cognitive approaches are by no means unaware of the existence of these caregiver-led interactions, rather they go to great lengths to operationally distinguish and explicitly exclude all interactions which don’t allow to deduce the infant’s capacities clearly12).

In keeping with the cultural tradition taking into account and starting their investigation from the material aspects and situatedness of the interaction, Moro and Rodríguez 2) use “triadic” to characterize interactions whenever infant-caregiver and object are present, and they 3) – drawing from Vygotsky and semiotics – even speak of “a triadic perspective” that researchers may adopt: “From a triadic perspective, the role of the adult as

12 e.g. Tomasello (Tomasello et al., 1993) painstakingly carving out by thought experiment “true” joint attention, ruling out a range of situations where a shared visual focus is present, but which do not allow to clearly infer the infant’s awareness of this shared focus; Adamson and Bakeman (1984) differentiating “passive joint engagement” from “co-ordinated joint engagement” thus explicitly introducing a “not-yet-proper-joint-engagement” category. These examples do not illustrate an oversight on their part but their different focus and evaluation of the situation.
guide in providing the child with the semiotic mediators, which allow him/her to incorporate increasingly complex levels of meaning."

With this move they interestingly transfer the concept “triadic” from the level of phenomena to the level of observation and linking it to semiotics: conceiving of [any?] interaction as involving three poles and taking into account all three poles in analysis: the participants, in this case infant and caregiver and a third pole – the latter may include any “meaning” or anything which serves as a “semiotic mediator”, be it a gesture or an object.

From this perspective – taking up Vygotsky – essentially all “construction of knowledge” (echoes of Piaget), which is what they seek to understand, is necessarily triangular (extending Piaget with Vygotsky and Peirce):

*The construction of knowledge is triangular. This means the child holds an indirect relation with his environment, one that is, instead, mediated through the semiotic systems adults produce. The unit of analysis for the construction of human psychological processes has to be triadic. In other words, without communication and education, cognitive development simply does not – could not – take place."* (Rodríguez, 2007)

With this agenda, Moro & Rodríguez seek to extend Vygotsky and Bruner by looking at earlier object interactions, rather than later ones based on language, and to extend existing research on triads again by starting at earlier ages and emphasising in particular the adult’s role as a guide. To extend the theoretical framework, they draw on Peirce's semiotics, as it 1) provides a theory of meaning broader than language and is hence suited for looking at pre-verbal infants, 2) is not restricted to “intentional communication” as meaning is defined by its effect and interpretation (the interpretant) rather than by the speaker’s intent, and 3) is dynamic, that is, meaning changes in an ongoing process of semiosis.

### 2.5.2.2 Empirical study

For the empirical part they conducted a qualitative longitudinal study, visiting 6 infants (3 male, 3 female) at home at 7, 10 and 13 months of age (starting when infants already handle objects and ending just before speech), and each time videotaped infant-caregiver-object play with 2 toys
(5 min each) provided by the researchers: a toy phone (which suggests acting in an symbolic space) and a truck with blocks to insert (which suggests manipulating blocks in a canonical way), instructing the caregivers to “play with your child as you usually do”.

In order to gain insight into “processes of construction rather than results”, they performed a microgenetic analysis aiming to focus on “all parts of the triad”, asking how the use of objects (whether non-canonical or canonical use) and the communication about objects change, and in particular, how signs (ostension (showing), indexical pointing (distal, immediate, multiple), demonstrations (distal or immediate), verbalisations) are used in the process (Moro & Rodríguez, 2004).

2.5.2.3 Results

Moro and Rodríguez report a gradual socially guided development for infant’s appropriation of social (canonical and symbolic) uses of objects: before 9/10 months infants spontaneously performed only non-canonical uses such as sucking or banging. Adults produced a lot signs, including ostensive gestures (showing and giving, e.g. shaking a noise-producing hoop), demonstrations of canonical uses, immediate ones (putting the receiver to the infant’s ear and saying “hello”), distant ones (putting a hoop around a stack, inserting a block into the truck, saying “Hello” on the phone) and conventional indexical pointing as well as verbalisations. These demonstrations often prompted infants to engage with objects, albeit more complex signs like complete demonstrations or conventional pointing were in the beginning only understood ostensively, resulting mostly in infants focusing on and engaging with the objects, but in their own mostly non-canonical way. However, it was following demonstrations that infants first performed (to some extent) canonical object uses whose meaning they claim was thus agreed between infant and adult. Those first canonical uses were single holistic acts (e.g., receiving a noise-producing hoop and shaking it, or when the receiver is held to the infant’s ear by the adult, adopting an upright “phoning” posture), slightly more complex ones at 10 months (putting a receiver to one’s ear), and it was only later (13 m) that sequences of object-object relative movements (e.g. inserting blocks into truck or dialling and
phonning) were performed (Moro & Rodríguez, 2004; Rodríguez, 2007; Moro, 2011).

Looking at infants’ actions in terms of interpretation and production of signs, the phone example illustrates how they move from iconic to indexical: the action of “adopting an upright position when holding a receiver to one’s ear” moves from functioning as an icon as the mother repeatedly demonstrated it (referring to itself by virtue of similarity) to functioning as an index as the infant accepts the receiver held to his ear and adopts an upright “phonning” position, thus having established a connection between the action and the receiver: if receiver: put it to the ear and adopt a phoning position. Similarly a little later in the infant’s own actions: as the receiver accidentally ended up next to the infant’s ear during play, the infant’s vocalisation “euu” suggested an iconic recognition of his earlier similar experience and a little later, as the boy completes the action by adopting an upright “phonning” position, the just before iconically recognised action of holding the receiver to ones ear can now be seen as functioning/being interpreted as an index pointing to the [rest] of the “phonning” action. Already realising aspects of the canonical use without, however, being aware of it (Moro, 2011).

This example also demonstrates how actions first encountered in situations led by the caregiver later on show up in infants’ spontaneous behaviour. Taking up this Vygotskyan theme, Rodríguez also reports how gestural signs first used by adults to guide infants were later on adopted by the infants as “private gestures” regulating their own behaviour in a longitudinal study focussing on a girl with Down Syndrome: the girl, who at 12 months only was able to put hoops around a pivot after adult’s exaggerated ostensions, repeated pointing, and demonstrations, by 18 months used such ostensive and pointing gestures as private gestures to herself to guide herself towards the conventional use of the object (Rodríguez, 2007; Rodríguez & Palacios, 2007)

Extending their observations to earlier ages, Ana Mureno-Núñez in her doctoral research (2014) performed further longitudinal studies visiting typically developing infants and their caregivers at their homes,
video of object play sessions, again instructing the parents to “play with your child as you normally would”: 1) 3 infants at 2, 4, 6 months of age were recorded engaging in 5-minute play sessions with a stacking cone with sonorous and non-sonorous rings. 2) 6 infants at 2, 3, and 4 months of age with the same procedure but with a different object: a rattle (maraca). 3) Infants at 9, 11, and 13 months for 10-minute play sessions each with 8 different (sets of) toys (Moreno-Núñez, 2014; Moreno-Núñez, Rodríguez, & Del Olmo, 2015).

A main focus of the analysis was again on the adults’ actions: object uses and sign acts, although the infants’ actions were coded as well. She reports that caregivers interacting with these young children rather than performing complex demonstrations of canonical uses (stacking hoops on a stacking pole), or engaging in conventional pointing actions, preferred ostensive uses (showing, giving), in particular making use of the sonorous rings, rhythmically shaking them to capture the infants’ attention. The older infants (study 3) engaged in a host of communicative gestures, but mainly with objects themselves – (self-)ostensive as well as giving and taking – and, to the extent that they engaged in pointing, they mainly used proximal gestures (“touch-pointing”) rather than the distal type of pointing emphasized in the cognitive literature (e.g. by Tomasello). Taken together, her research underlines the importance of ostensive actions in particular and communicative object actions in general, pointing to the role of objects as means rather than merely as targets of communication (Moreno-Núñez, 2014; Rodríguez, Moreno-Núñez, Basilio, & Sosa, 2015).

2.5.3 Patricia Zukow-Goldring: Educating attention and assisted imitation.

2.5.3.1 Background, interest, theoretical framework

Patricia Zukow-Goldring has been interested in, and conducted research on, how infants become adept members of a culture, “how they learn what everyone else already knows” (Zukow-Goldring, 2012), how they learn to participate in cultural activities, and, in particular, how they master language (how they understand the word-world fit), where the former is also seen as the foundation for the latter, as also illustrated by her SEED
(Situated, culturally Embodied, Emergent and Distributed) model of language. Her studies investigate and document cross-culturally “how more skilled members of a culture, older siblings and adults, assist infant-novices to notice, participate in, and finally communicate about ongoing events” (Zukow-Goldring, 2012) through the processes of “educating attention” and “assisted imitation” (Zukow-Goldring, 1997, 2006, 2012).

Looking for a way to address her questions, the existing approaches did not quite seem to be sufficient:

*Cognitive theories concentrate on autonomous mental mechanisms untouched by direct contact with ongoing events to explain the growing abilities of children (they examine what children put in their head), whereas ecological [cultural] approaches emphasise that the social world guides development (elaborate the active where). Neither tells us the particulars of how children make sense, and thus learn, from the details of actual events.*


Therefore Zukow-Goldring combines aspects from multiple different approaches, each contributing one crucial piece to a possible working methodology to tackle this question: 1) Linguistic anthropology and ethnomethodology take a detailed look at how (mostly language competent adult) members of a culture negotiate meaning in the unfolding interactions of daily life, “continuously embodying their verbal messages by making them visible to one another in a seamless ensemble of spectacle and talk” in order to reach a working “consensus for all practical purposes” (Zukow-Goldring, 1997). These approaches, however, do not specify how exactly perceiving and acting create knowledge and thus inform action to negotiate meaning and act together, in particular across a discrepancy between the knowledge and communicative competences of novice-infant and proficient adult.

2) Vygotsky and sociocultural approaches in general emphasise the role of proficient members of a culture in guiding the infant-novice to achieve cultural competence. Vygotsky even constructed a rich theoretical framework, stating that new competences and knowledge first “arise for the child within social interaction” – along a constantly moving zone of proximal
development, just beyond what the infant could hitherto do on his own – and “only later appear as an individual achievement” (Zukow-Goldring, 1997). Still, they again are lacking in details how exactly this is achieved by caregivers and (in particular pre-verbal) infants (compare section 2.5.1).

3) In contrast, Ecological Psychology, (J.J. & E. Gibson) finally details a) how organisms through interaction with the (physical) environment detect the perceptual structure that specifies what the world affords for action (depending on their respective phylogenetic and ontogenetic history), guiding them what to do, and b) how experience, as we continuously move through perception-action loops “educates attention”, increasingly refining affordances and effectivities. The effect of the social world on these processes, however, is “conspicuously absent” (Zukow-Goldring, 1997).

Integrating these approaches into what she calls “social ecological realism” – Zukow seeks to investigate 1) how caregivers and infants in everyday interactions negotiate meaning and “practical consensus” by educating each other’s attention to perceive a common ground for action in everyday activities across the knowledge and capacity gap between them, and, in particular, 2) how this gap is bridged: “how more skilled members of a culture, older siblings and adults, assist infant–novices to notice, participate in, and finally communicate about ongoing events” (Zukow-Goldring, 2012) through the processes of “educating attention” and “assisted imitation”. In turn, 3) how this is achieved through caregivers “making prominent culturally relevant possibilities of perceiving and acting”, making visible previously unnoticed affordances (possibilities for action) in the environment and their dynamically-coupled relation to the infants’ own bodily abilities (effectivities) (Zukow-Goldring, 1997)

2.5.3.2 Empirical study

As an empirical basis for pursuing these questions, Zukow-Goldring conducted a cross-cultural naturalistic longitudinal study visiting 12 infants - from 6 European American middle-class families and 6 Latino low-income families - from 6 to 30 months of age, using video recordings, field notes and diary entries for data collection.
Engaging in qualitative and quantitative analysis, she investigates the role of educating attention and creating invariances across modalities in the acquisition of the lexicon, thus moving beyond (standard) labelling-by-association accounts of word learning and their inherent ambiguity problems, how to know which aspect of a situation a word stands for (often, but as Zukow shows, inappropriately, associated with Quine’s indeterminacy of reference: “gavagai”) (Zukow-Goldring, 1997).

Beyond her focus on lexicon acquisition, Zukow-Goldring also investigates more generally how children come to participate in and gradually master specific cultural activities being guided step by step by parents and siblings - more skilled members of the culture – through “assisted imitation” and “cultural embodiment”, continuously refining affordances and effectivities. (Zukow-Goldring, 2006, 2012)

2.5.3.3 Results 1: Educating Attention in lexicon acquisition

Qualitatively analysing example situations, she specifies how exactly caregivers gather and then “direct attention to one specific element, relation or event, over the multitude of other possibilities present in a specific place and time”, using multiple modalities and creating invariances across them so the infant can “detect invariance within a unitary whole” and “equivalence across dissimilars”. For example, a caregiver shakes an open bottle of sweet smelling vitamins:

“When she says and does “shakey = sha::key = SHA:::KEY!, she coordinates in action, vision, odor, and sound the target of attention, cultivating perceptual pickup of amodal invariant relations through tempo, rhythmicity, synchronous initiation-termination, accelerating intensity, and so on.” (Zukow-Goldring, 1997)

She develops an “ethnography of attention directing”, differentiating different ways and targets of directing attention, pointing out affordances and effectivities to the infants:

In act-ons caregivers put infants through the motions of some activity (caregivers pull infants up as they say up); In shows caregivers control the infant’s line of sight with a translational movement in which they loom-
magnify an object toward the infants (saying ribbit while looming a frog) in demonstrations infants who are monitoring their caregivers must detect or pick up in the perceptual flow the action to be repeated or completed (saying hi when catching gaze, smiling, and greeting) of infants may be invited to see-feel-say the texture of bristles in a broom while caregivers say stick; finally in points infants must detect the intersection as a gesture’s trajectory through space to the place where it intersects with some target of attention (caregivers pointing to and saying over there) (Zukow-Goldring, 1997 emphasis mine)

Quantitative measure of occurrences showed the way of guiding attention and the target attention is guided to (and its complexity) changed over time: the semantic complexity of the target increased in both US and Latino families (in the beginning, simple non-dynamic objects, initially dominant, decrease in comparison to actions, complex relations, and events. Specific ways of guiding attention correlated with specific targets: shows were mainly used for non-dynamic objects and animate beings (supposedly helping the infant to link words/sounds to structural invariances), while for actions, act-ons and demonstrations were used preferentially (displaying transformational invariances potentially linking words to action affordances). In general, “what caregivers communicate foreshadows what infants express several months later”. (Zukow-Goldring, 1997)

2.5.3.4 Results 2: Learning to participate in cultural activities through assisted imitation and cultural embodiment

Beyond her focus on lexicon acquisition, Zukow-Goldring also investigates more generally how children come to participate in and master specific cultural activities. She reports from cross-cultural analysis that children’s learning and mastery of cultural activities – be it pulling the string of a vibrating music box, peeling an orange, or kicking a football – do not usually occur by children spontaneously copying complex actions they have observed from a distance, as e.g. suggested by approaches like Tomasello’s. Indeed, such “spontaneous imitations” occurred in less than 10% of cases, with the remaining 90+% of imitations being examples of “assisted imitation”. Infants were guided through a longer process of interaction and
dialogue with (more) proficient members of the culture, as parents again set up scaffolded situations, demonstrated and invited infants into actions, and “culturally embodied” the activity by putting infants through the motions including acting on the infant’s body, continuously refining affordances and effectivities. (Zukow-Goldring, 2006, 2012).

A particularly striking example illustrates how the cultural activity of “playing soccer” was “culturally embodied”, assisting the child in an ongoing process of guiding attention by his mother and family members continuously refining affordances and effectivities, to develop from soccer novice to expert in the short time span from 6 months to 3 years:

1) At six months, a soccer ball was repeatedly shown and loomed towards the boy, as well as thrown to family members, accompanied by “going up!”, carving out the ball as a relatively stable noticeable pattern. 2) At 8 months, the infant was gently swung like a pendulum – stiffening his body – by the mother shouting “Andale!” to hit the ball the brother rolled towards him, making experienceable his effect on the ball (effectivity). The mother also supported him in a sitting position and was acting on his leg, moving it and thus providing him with the first culturally embodied practice of kicking the ball – which was held in front of him by his brother – accompanied by “kick!” and “goal!” which came to be among his first words. 3) At 10 month, before he could walk autonomously, he was held on both hands and (while attentively tracking but failing to kick a moving ball) he successfully adapted his posture to an emerging situation, taking a preparatory step towards the ball at rest and kicking it. 4) At 13 months, he was held on one hand only and placed the ball himself and even used the ball instrumentally to remove obstacles getting in the way so he could kick it, now flexibly adapting to dynamic circumstances. 5) At 22 months, while tossing the ball to and fro with his grandmother, he accidentally drop-kicked the ball and his following joyous jump could be seen as indicating a positive self-evaluation of his success 6) by 3 years he also shows his built-up “expertise”, criticising his older brother's technique and showing him the correct way to kick the ball with the inner foot (Zukow-Goldring, 2012).
2.6 Accounts V: Current and interactive accounts

This section focuses on studies with recent and ongoing data collection, whose reports are easily accessible, and are the subject of current discussions within the field. Therefore, unlike the previous sections, this section will give a much briefer overview of and localize selected current research programmes investigating infant-caregiver-object interactions in semi-naturalistic settings. As before, we will provide a summary of where the researchers are coming from, what phenomena of interest they focused on, and how they exerted control to make the semi-naturalistic settings more amenable for observation and analysis.

While there have been sections within developmental psychology which have continued conducting (qualitative) home visits and developing theoretical frameworks to address and capture the richness of multi-modal multi-strand interactions (see previous section), most of the field has focused on data from experimental investigations spelling out ever finer distinctions of scripted, normative interaction patterns such as those involved in “joint attention”. However, (semi-)naturalistic investigations are currently enjoying a (small) renaissance, with multiple research groups extending and actively developing new methodologies to address infants everyday life social and object interactions. In the meanwhile, developmental science – as well culture more generally – have been rethinking the pace of cognitive development and are much more open about attributing (precursors of) quite sophisticated cognitive capacities to even very young infants. Therefore, the studies discussed in this section tend to begin data collection at earlier ages, looking towards describing and relating the experiences and abilities of infants much younger than 6 months – which was the starting age of e.g. the seminal studies of Adamson and Bakeman who categorized interaction patterns of infants at this age mostly in rather non-descript terms as “passive joint engagement” (see section 2.3.2).

The studies discussed here, while heterogeneous, all have in common that they reflect one or more of recent trends in cognitive science (see Reichelt & Rossmanith, 2011), here in the context of developmental
psychology. The embodied action trend, here especially impacting on infant research in form of collaborations with (developmental) robotics (see e.g. Vollmer et al., 2014) has put a firm emphasis on the role of sensorimotor control of an agent’s own actions, including eye movements supportive of action control, in particular as a basis for learning. At the same time, machine learning approaches have led to a renewed interest in the rich perceptual world of infants (an emphasis they share with Gibsonian ecological psychology), guided by a general appreciation of the extensive learning opportunities inherent in large sets of time-series data, both for unstructured learning mechanisms (model-free) as well as for learners driven by prior expectations of particular patterns, markers, and relationships within the dataset (model-based; see e.g. Smith, Suanda, & Yu, 2014). Such trends have helped re-ignite interest in what was previously mostly relegated to “socio-cultural context”, taking seriously the ideas the cognition develops situated in a (social and cultural) environment and needs to be understood in these terms. Finally, the many ways persons may (directly) engage with each other – which necessarily also includes the relationship the researchers themselves adopt towards their “subjects” of study – have been emphasized by “second person” approaches (see e.g De Jaegher, Di Paolo, & Gallagher, 2010; Hobson, 2005; Reddy, 2008).

2.6.1.1 Mother-infant sensorimotor coordination at multiple timescales: de Barbaro, Johnson, & Deák (2013)

Based in San Diego, Kaya de Barbaro and colleagues bring the “west-coast” flavour of situated (distributed) cognitive science into developmental psychology, most clearly that of Edwin Hutchins who arguably has been the most influential proponent of situated cognition (Hutchins, 1995, 2008). This development and shift in research direction is most visible in Gedeon Deák’s research who was previously involved in a large-scale project of modelling a well-defined computational capacity, gaze following, as it develops over the first year of life as reflected in experimental tests of this ability (Triesch, Teuscher, Deák, & Carlson, 2006). However, neither these experiments, nor the computational model that accounts for them, actually address the relative importance of gaze following for engaging in
observational learning and social interactions. Semi-naturalistic observations (Deák, Krasno, Jasso, & Triesch, 2017), and controlled laboratory studies (Yu & Smith, 2013) led Deák to question whether this capacity is of any, let alone central, importance as implicit in standard accounts of joint attention. Instead, infants simply monitor the activities themselves as they unfold, shifting gaze towards hands and objects rather than focusing on gaze direction to anticipate upcoming actions. This is an example of the general attitude of situated cognition that cognitive processes – rather than being a hidden realm that observers need to painstakingly infer from indirect cues – are more often than not publicly available as they expand into the world through engagement with activities (see also Leudar & Costall, 2009; Reddy & Morris, 2004).

From the series of studies de Barbaro and colleagues have conducted, we have selected the one investigating triadic behaviour most closely. Here they report from home visits of 5 mother-infant pairs at the ages of 4, 6, 9, and 12 months. At each session, a free play activity was staged lasting several minutes, with identical sets of objects brought in and infants’ posture stabilized with infant walkers (for the 4-9 month olds) and captured with a set of 3 synchronized cameras directed at infant, caregiver, and overall scene, for later frame-by-frame microanalysis of sensorimotor coordination in the mother-infant-toy triad. For their analysis, de Barbaro and colleagues stressed the importance of taking a comprehensive look at multiple modalities involved, multiple parties, and multiple time scales (see also de Barbaro, Johnson, Forster, & Deák, 2013):

At the microsecond scale, we observe shifts of gaze, facial expression (e.g., gleeful smiles), and hand movement. At the macroscale, a particular look or grasp is positioned within an ongoing routine where, for example, it may repeat (as in peek-a-boo), or change, or organize with other events. At the historic-developmental timescale, the dyad’s long-term experience with such routines (e.g., a playful father’s tendency to initiate exciting games) comes into play (de Barbaro, Johnson, & Dédik, 2013)

Aside from the measures that introduced a level of control over the free play situation and spatial configuration itself mentioned above, de Barbaro
and colleagues necessarily had to be selective in coding and analysing behaviours between these multiple time-lines. They focused on the role of different modalities in object play, sharing of arousal, and in particular on “maternal bids for the infant’s attention to objects and the infants' responses to them”.

Among their many results stand out in general the gradual character of the development of triadic attention, and in particular, mapping trajectories within these overall gradual changes such as a decoupling of eyes and hands from around 6 months which before had showed a “convergent” mode of attention coordination, with all modalities linked and focused on the actions at hand.

2.6.1.2 Educating attention in early nappy-change activities: Nomikou, Rohlfing, & Szufnarowska (2013)

Nomikou and Rohlfing both contribute their background in linguistics, translation studies – as well in cognitive science (Rohlfing) to qualitative and quantitative longitudinal studies of semi-naturalistic interactions.

In the dataset of the selected study – partially overlapping with (Nomikou & Rohlfing 2011) – the authors conducted home visits to 17 mother-infant pairs first at 3 month and then again at 6 month to look at nappy change interactions. To standardize the setting, they brought in a foldable changing table and 2 cameras to capture both the mother's and infant’s (face to face) communication and actions. In their earlier study relying more on quantitative analyses Nomikou and Rohlfing illustrated the rich, multi-modal character of the infant’s perceptual streams in connection with language (which are emphasized by machine learning approaches to e.g. language acquisition), noting in particular the co-occurrence of language and action, describing how caregivers place maintain tight temporal relationships between their manual actions and vocal actions thus “making the vocal signal both perceivable and tangible” (“acoustic packaging”) in creating multi-modal invariances for learners.

In the present study, Nomikou and colleagues go one step further, now relying more on qualitative analyses, to show the contingent, interactive
character of the mother’s actions in the interplay with the infant. They singled out episodes of mutual eye contact, and in particular how mothers follow their infants’ lead, entrain their actions and use ostensive gestures and vocalizations welcoming them to achieve eye contact and mark particular points of the activity, especially its beginning. The approach taken here is thus in many ways complementary of that of de Barbaro and colleagues discussed above, as they focused on nappy change rather than free-play and looked at eye contact rather than directing attention to objects. Both approaches exert comparability of instances by staging cultural activities and furnishing the required objects (but see Rączaszek-Leonardi, Nomikou, & Rohlfing, 2013 for qualitative analyses of unscripted triadic interactions).

2.6.1.3 Second person approach to triadic interaction: how infants feel minds (Reddy, 2008)

Similar to both the previous accounts, Reddy argues for the gradual emergence of triadicity in attentional engagements. Also in line with both these accounts, Reddy adopts a strongly adualist stance towards the notion of attention; it is posited to be something that is perceptually available within engagement. Somewhat different from these accounts, however, Reddy emphasises affect as key to the understanding of attention. Strongly influenced by Martin Buber’s notion of I-Thou relating as a mode of knowing, Reddy’s account of the emergence of triadic attentional engagements begins with the infant’s affective responses to attention in dyadic engagements. Being addressed as a You by another’s gaze, she argues, is a powerful arouser of emotional response. And this response, experienced within engagement and inevitably intertwined with the perception of the other’s gaze (in the case of visual attention) becomes the key to the emerging meaning of attention for the infant. Merleau-Ponty focused on the perceivability of engagement as the key challenge to Descartes’ emphasis on thought: “I discover vision, not as a “thinking about seeing,” to use Descartes’ expression, but as a gaze at grips with a visible world, and that is why for me there can be another’s gaze” (Merleau-Ponty, 1962, p. 410). Reddy takes the point further, arguing that attention is not
only perceivable, it is, in the typical case, experience-able: “I discover vision [...] as a gaze at grips with me, and that is why for me there can be another’s gaze” (Reddy, 2011). And it is precisely this experiencing of attention, she argues, which draws dyadic attentional engagements into expanding and developing into and beyond the triadic.

The dyadic experience of attention is evidenced in the first few months of life with roughly similar responses – positive, negative, ambivalent and indifferent – to those that can be found later in the first year. What changes is the complexity and scope of the objects of attention that the infant can grasp. Reddy argues that it is not a late appearing ‘discovery’ of attention (the position adopted by Bates and colleagues, for instance) that characterizes changes during the first year but an expanding understanding of the objects of attention. What is needed here to understand developmental shifts is to re-think what can constitute an object of attention, or what the ‘third element’ can be, which enables triadicity. She argues that the triad is present much earlier than the 9 or 10 months assumed typically, and that the ‘object’ can be not only distant things, but also aspects of the infant’s body and actions. These objects of the other’s attention (as grasped by the infant) expand in an outward direction going from the self, to its body parts and then to actions by the self, before involving distal objects. In the second year the expansion continues further, moving to objects which are not currently present or perceivable (see also Tomasello & Haberl, 2003 and compare section 2.2.3: Sugarman). Evidence for these expanding attentional engagements comes from naturalistic observations: of coy responses to mutual gaze between 2 and 4 months (Reddy, 2000, 2003) and of clowning and showing-off beginning around 7 months (Reddy, 1991, 2001).

2.6.1.4 Sampling of infants’ visual world and mapping gaze coordination in caregiver-infant-object interactions: (Fausey, Jayaraman, & Smith, 2016; and Yu & Smith, 2013)

Having co-written the book on dynamical systems theory in (developmental) cognitive science (Thelen & Smith, 1994), Linda Smith in
collaboration with the computer scientists Chen Yu and colleagues in their labs have pioneered an approach that uses infant head cameras to create corpora of the (visual aspects of the) world of everyday life of infants (Smith, Yu, Yoshida, & Fausey, 2015), allowing to explore and map infants’ experience statistically rather than relying on researchers selections of example episodes, thus sketching how the actual perceptual data streams that infant learners rely on look like with direct relevance for statistical or machine learning approaches (Smith et al., 2014). In addition, they also have pioneered a laboratory based dual-eye tracking setup based on the positive science trackers and infant harness that allows describing gaze coordination in infant-caregiver-object interactions. While this system samples eye movements only 30 times per second identical to frame-by-frame analyses, automated gaze estimations allow describing both the micro-structure and overall statistics of gaze coordination patterns (Yu & Smith, 2013).

While the video-libraries compiled by means of the head cameras reflect a less constrained aspect of the infants’ visual world compared to those created specifically for e.g. object play and nappy change contexts (see above), meaningful coding and analyses of such large data sets clearly still need to be theoretically guided. An initial question posed by Smith and colleagues concerned the relative frequency of faces and hands that infants get to see over the course of the first 2 years of life. Based on their quantitative sampling, they established that young infants of 2-4 month of age have a bias for faces which gradually gets reduced and inverted into a bias for hands for infants 9 month and older (Fausey et al., 2016). Additional controlled studies indicate that this bias reflects active interest of infants at least in part rather than only the presence of faces rather than hands close by (Jayaraman, Fausey, & Smith, 2017). This result goes hand in hand with their main finding, mentioned above, that episodes of shared attention of infants (around 1 year of age) and caregivers on objects are rarely if ever preceded looking at faces (to extract cues to engage in gaze following as assumed in the normative model for joint attention), but rather that infants (and caregivers) directly look at hands and the objects acted on (see also
2.6.1.5 **Comparison**

Juxtaposing these quite different research programs showcases the many aspects of interaction that can be made visible and addressed when investigating interactions systematically, looking at the roles of each person involved, as well as crucially their interplay (see also De Jaegher & Di Paolo, 2007), and how objects and the (cultural) configuration impact on and shape this interplay. In particular, the multi-modal character of interactions – vision, vocalization, and (felt) action – is shown to be fundamental to aspects of development from labelling (“acoustic packaging”) to social and object play.

Multi-modality may be important for all these research programs since they are investigating situations where the infant has much more to work with than in simplified artificial laboratory experiments, which often focus on vision in isolation: (semi-)naturalistic settings are much richer, and in particular involve direct bodily contact and contact with objects (nappies and toys), so gaze coordination occurs in parallel and in relation to ongoing activities.

Perhaps because these accounts cast a much wider net – including the cultural environment, and the infants’, not the experimenter's, time-scales of experience – while at the same time looking into the micro-level of affective exchange and gaze coordination, they each in their own way emphasize the gradual character of development in general and triadic interactions in particular by pointing out earlier engagements that prefigure the infants’ later challenges and abilities.

By showcasing the intricacies of particular activities (nappy-change, object play) or classes of activities (sharing of attention, even the totality of the visual life world over a particular temporal window), starting from one or more central relationships of interest (eye contact, maternal bids, involved engagement, statistical shifts from faces to hands) and expanding outward, these approaches also show their limitations, as results from
particular activities such as nappy change and object play may not generalize and may miss other important interaction dynamics and experience (compare chapter 3: book sharing). Even the focus of interest within a particular activity may lead to a somewhat different account (compare: Nomikou & Rohlfing, 2011; Nomikou et al., 2013; Rączaszek-Leonardi et al., 2013).

While de Barbaro and colleagues (2013) provide an explicit statement of the benefits and methodological challenges inherent in multi-scale analyses, their coding methodology focusing on singular events (maternal bids for the infant’s attention on objects) foregrounds “islands” of events and their immediate context in the configuration to abstract, pool, and compare. While shedding light on the event of interest and enabling statistical analyses, such a perspective may miss activity structures linking the local to the global time scales, such as action arcs (see section 6.7).
2.7 Context box: On the separation of (manual) object - and (language-mediated) people interactions; parallels and interconnections between infant research, human- and world-conceptions, and socio-political developments

2.7.1 Empirical and conceptual roots

Where does the notion that object engagement and (in the long run language-and-reason related) social engagement constitute separate realms for infants through large parts of the first year, with its far-reaching consequences, come from? It echoes through many of the seminal papers on infant-caregiver-object interaction from the 1970s (see below), apparently either considered self-evident, as an unquestioned underlying assumption made from the outset – thus often not even considered to require a reference (e.g. Bates et al., 1979; Bakeman & Adamson, 1984) – or as a conclusion that is quickly drawn or jumped to.

One relatively recent empirical root is constituted by the seminal studies performed by Brazelton, Richards, and Trevarthen in the 1970s (in Jerome Bruner’s Center for Cognitive Studies in Harvard): they have documented differences between young infants’ responses towards their caregiver (communicative attempts and a smooth flow of engagement/disengagement) and towards a suspended object (jerky attention bouts and pre-grasping attempts). Suggesting two separate interaction classes, action-systems, or motives, allowed them to point out and highlight the astonishing social skills displayed by even very young infants, which – since not captured by e.g. the Piagetian model of development – had been neglected at the time. However, differential responding to distinct setups does not mean that objects and people constitute two separate worlds for young infants (Brazelton et al., 1974; Trevarthen, 1977).

Analogously, in approaches such as those of Bates and colleagues, or of Sugarman-Bell (see section 2.2.3) – starting from language and linguistic reference typically associated with the classic labelling situation
“communicating with people about objects” and then seeking to find its first non-verbal forms – the dichotomy was already pre-shaped to some extent in the underlying theoretical assumptions as well.

However, as Trevarthen’s PhD student Penelope Hubley pointed out in her dissertation that while first aiming to conduct 2 separate studies on object and social interaction, respectively, it became soon clear that [at least?] “from 5 months on these lines of development were not separate” (Hubley, 1983). This marked the beginning of Trevarthen and Hubley's seminal studies documenting the gradual development of infant-caregiver-object interaction from dyadic primary to triadic secondary subjectivity (see section 2.3.1). This gradual character, however, was once again lost in later accounts, re-entrenching the image of separate lines of development enduring until the end of the first year.

2.7.2 Division of labour within psychology

This distinction is not a theoretical detail, nor simply an empirical question to be addressed by observation and experiment, but runs far deeper, and the way questions are posed and results interpreted are in part already shaped by our expectations. The distinction runs through the practical organization of introductory textbooks where object and social engagement routinely form distinct chapters, neatly partitioning the infant’s world and mind - as well as the reader’s thinking while reading along - into separate realms (see e.g. Rochat, 2001 The Infant’s World: Chapter 3, The Object World in Infancy; Chapter 4, The Infant and Others). Fundamentally, the division is built in right into Psychology’s own division of labour: object engagement, physical knowledge, and tool use on the one hand, and infants’ social interaction, social cognition and knowledge on the other, have been largely investigated by different fields of research, pursued by different research communities, using different methods, frameworks, etc., and not necessarily communicating with each other. At times, the distinction is even reflected in completely distinct phases in a given researcher’s career. Take as an example again the work of Rochat encompassing his first seminal contributions to object engagement based on a Gibsonian ecological
cognition approach (Rochat, 1987, 1989), and later ones to infant social cognition based on a cognitivist framework (Rochat, 2009).

Why does it matter? Social object interactions involving young infants are in danger of falling through the cracks of this divide and understanding the development of culture – which characteristically involves jointly creating and interacting with and via artifacts – is made difficult if not impossible.

2.7.3 Philosophical background: differentiating “logos” and the ever shifting “man vs. animal” divide

This distinction may reflect even larger (under-)currents in the history of Western philosophy and psychology, in the latter playing themselves out within comparative as well as developmental psychology. It becomes visible in (the turns of) an unfolding argument concerning the nature and thus the status of (white) man among his fellow creatures, which turns to and fro between finding similarities, then new differences, differentiating the concepts of mind and mental capacities slightly further with each turn, maintaining man’s (supposedly) special position among fellow beings by pushing the demarcation lines for “intelligent/sensible being” along the newly emerging distinctions ever further.

For a long time language and reason combined in “logos” have been considered the defining attribute of humanity, only allowing fellow animals the ability to act “as if” from reason. However, researchers at the beginning of the 20th century began to increasingly break up that “bastion” of humanity. Researchers like Otto Köhler and Karl Bühler argued that language-lacking apes and infants nevertheless are capable of some kind of reason, separate from and prior to language: “practical” or “technical” reason/intelligence as exercised in instrumental object actions. Köhler’s successful non-verbal testing methods (Köhler, 1917) were fruitfully adapted by the Bühlers for use with pre-verbal infants (Bühler, 1930, 1936), thus placing apes and human infants within one research framework and thus extracting “practical intelligence” from the unity of “logos”/”logos-conglomerate” and pulling it downward and outward, claiming it for young infants and close relatives of our species.
In the next turn of the argument and research process, Vygotsky at the beginning of the 1930s criticized Bühler’s move to equate apes and young children as simplistic and countered by pointing out that the coming together of practical intelligence and language re-organizes children’s object interactions and tool use, thus further differentiating “practical intelligence” into a “natural/biological stream” outside of and prior to language and a “cultural stream” transformed in human infants from the 2nd year on by language based socio-cultural interactions. Whereas Vygotsky criticized the then standard practice to study tool use and symbol use isolated from each other (with the latter regarded as a spontaneously discovered product of pure reason rather than of developmental history), since he regarded the “dialectical unity of these systems in the human adult the very essence of complex human behavior”, nevertheless his conceptualization of early development claims the existence of two separate lines of development foreshadowing today’s developmental narrative: “The most significant moment in the course of intellectual development, which gives birth to the purely human forms of practical and abstract intelligence, occurs when speech and practical activity, two previously completely independent lines of development, converge.” (Vygotsky, 1978b p. 24).

Similarly, Piaget’s infamous “neglect of the social” can be seen against the background (and result) of an already established separation between the workings of language-bound “reason” and all other motions of life. Countering his own previous “overemphasis” on the social, which he had considered crucial to overcome a child’s egocentrism when doing interview studies with slightly older children, he sought to demonstrate in his work on early sensorimotor development that all cognitive foundations for later cultural activities develop autonomously from the infant’s sensorimotor interactions with the object-world independent of social influence. The latter was only supposed to play a big role later on as arbitrary socially established word labels and morals are learnt, and as logic is developed further through argumentation (Piaget, 1962; Lourenço & Machado, 1996).

Approaching current times the movement continues to result in more complex pictures: Starting at the beginning of the 20th century (e.g.
Wittgenstein, Bühler etc.) and taking off in the second half (Austin, Searle, Ryan) language is further differentiated: apart from its “cognitive” functions related to logic and semantic reference, it’s use and pragmatic aspects and its functional role in interaction come into focus and are further investigated (e.g. conversation analysis) – moving at least parts of language closer back again to “practical intelligence”. Yet again various aspects previously regarded as part of the language-logos complex, such as communication as social regulation (Bateson, 1975, 1979; Brazelton et al., 1974; Trevarthen, 1977; Tronick, 1982), as well as the “transmission” of meaning in various forms, were isolated, extracted and pulled downward and claimed for human infants as well as non-human animals, stimulating further investigation in newly emerging research fields (for infants see Halliday, 1975; J. Bruner & Watson, 1983; Bates, Sugarman: see section 2.2.2-3).

The situation we face today within developmental psychology has become even more complex, as multiple lines of development (or at least complexes of behaviours investigated by specific research communities) can be discerned – depending on the respective research perspective and approach – the interrelations between which remain to be clarified:

These multiple lines of development include:

- early solitary object interactions, including (instrumental) tool use and related object knowledge (e.g. object permanence) associated with practical/technical intelligence (as singled out by Köhler, Bühler) and autonomous sensorimotor object interaction experience (Piaget)

- further extracted from this object-complex and countering Piaget: (“innate”?) knowledge about the physical domain from early on and apparently independent from object manipulation experience (see e.g. Baillargeon, 1987; Spelke, 1985).

- early communication and social mutual regulation (including arousal, affect, and turn-taking) from birth and flourishing from 2 months in proto-conversation (Bateson, 1975, 1979; Tiffany Field & Fogel, 1982; Jaffe, Stern, & Peery, 1973; Trevarthen, 1977; Tronick, 1982).

- (as well as attachment for longer term relating: Spitz, Bowlby,
- early forms of reference in form of proto-imperatives and proto-declaratives as described by Bates et al. from 9/10 months of age (see section 2.2.2)
- (in part) further extracted from language and communication: “knowledge” about the “social domain”, including animate vs. inanimate distinction (biological motion perception, bias for faces, etc.) from birth (Legerstee, 1992), goal directedness of (hand) actions (Woodward, 1998).

This still leaves in the logos-language domain (to emerge at a later point) capacities which can considered to belong to what Vygotsky claimed to be socially mediated and constituting the cultural line of development, such as: learning by imitation, following instructions, conventional object use, and symbol use.

However some of those have also been singled out and claimed to occur prior to language onset:

Moro and Rodríguez e.g. explicitly seek to extend and refine Vygotsky by pointing out that not only language but also conventional object meaning already needs social mediation, and – occurring prior to language – indicates earlier cultural impact and transformation, which can actually be considered to form part of the basis for language (see section 2.5.2).

Tomasello – also draws the start of cultural mediation to an earlier point – but in contrast to Moro and Rodríguez not as occurring through or initiated by mediation but as arising endogenously, with the two lines of development of social communication (from birth) and (instrumental) object knowledge (Piagetian intelligence emerging at 9 m) coming together. They are not mediated later by language but from around 9 months by new forms of social knowledge and motivation (“shared intentionality”) allowing infants to benefit from cultural mediation (see section 2.4).

2.7.4 Socio-political entanglements

The separation between object engagement and language-mediated social engagement has not been relegated to theoretical discussions within
science and philosophy but has pervaded society, culture and politics impacting on everyday life - this may help explain why it is so persistently comes up essentially as a default position, and is difficult to address within the confines of a single discipline. The separation lives e.g. in the (also class-related) distinction between “blue collar” and “white collar” work with its disdain for working with the hands (on material objects), which has a long history, and, e.g. in the early modern age – as argued by Edgar Zilsel in his “Social Roots of Modern Science” – first had to be (at least partly) overcome to make possible experimentation and engineering work thus enabling the birth of science (Zilsel, 2000). The gulf was arguably again further deepened during the 19th and 20th century in the course of industrialization shifting from the skillful expertise of a master craftsman based on extensive training to the monotony of simplistic repetitive conveyer-belt jobs (which also led to Marx’s notion of ‘alienation of labour’, Marx, 1844).

The distinction keeps riding high throughout the 20th century, the age of “formalization” and “operationalization”, where larger and larger parts of thinking and action get formalized and operationalized, thus emphasizing their abstract aspects and moving them further apart from effective interactions with the material world (Graeber, 2015; Mirowski, 2004).

This is reflected in the highly abstract logic-based conceptions of thinking common in cognitive science and psychology such as “information-processing” and “rule based symbol manipulation”. These notions not only point to a highly abstract conception of thinking, but also to how even aspects of thinking became increasingly mechanized and implemented in computers. This 1) leaves apparently less and less room for a “uniquely human” “logos”, condensing the puzzles of thinking – sustained or created by the current conception – into a couple of opaque labels such as “making sense”, “intentionality”, “free will”, “creativity”; and 2) is poised to do the same to white collar jobs what industrialization did to blue collar jobs, as is e.g. experienced in everyday office work with process managing software re-organizing, standardizing and determining human work flows (O’Neil, 2016); it is also reflected in education curricula, which already traditionally valued languages and maths above shop class and solving practical
problems, and which are now centred around acquiring knowledge and skills organized in distinct definable and interchangeable pieces to attain transferable credits.

However in parallel with pushing further the distinction between effective interaction with the material world and language-and-reason(logos)-mediated interactions with people there always have been tendencies in the opposite direction as well, seeking to reclaim and reunite these driven-apart aspects in theory and practice: including e.g. John Dewey’s pragmatism with its emphasis on practical “occupation”, the movements of embodiment and situatedness in cognitive science (Varela, Thompson, & Rosch, 1991; Hutchins, 2008; Suchman, 1987; Núñez & Freeman, 2000; Chrisley & Ziemke, 2006), efforts to bring the material world to the attention of psychology (Costall & Dreier, 2006; Moro & Rodríguez, 2004) appeals for working with one’s hands (e.g. Crawford, 2011), and cultural trends such as maker’s spaces, or a predicted maker’s revolution (Anderson, 2012; Gershenfeld, 2008).

In the meanwhile new aspects of the dichotomy are found and the demarcation lines between humans and the rest of life are continuously redefined. In many approaches like the one, of Tomasello and colleagues, already mentioned above, object engagement – in continuation of Bühler and Vygotsky – is associated with instrumental reason (though not its full flourished cultural forms), whereas social interactions are associated with empathy and co-operation, which are now, rather than language regarded as the uniquely human characteristics ultimately enabling instrumental reason and language to flower into their cultural forms. While these conceptualizations still attempt to squeeze non-rational phenomena into the traditional logos-framework, these phenomena do not quite seem to fit and already start to fracture traditional rationalist frameworks, asking for a larger reconceptualization and integration.
2.8 Conclusions and Outlook

Our analysis has shown that depending on the different historically grown agendas, theoretical assumptions and frameworks, methods, and ways of explanation the narrative has taken various different shapes.

2.8.1 Criteria for defining what counts as triadic interactions

While the analysed approaches all agree on the importance of triadic interactions, taking a closer look, there is little consensus concerning the exact criteria of what actually counts as a triadic interaction in the first place. (In some cases there may even exist differences between the theoretical definition, practical examples given, and how they are operationalized for coding and or experimentation within a single research programme.)

Some cultural approaches, taking into account and starting their investigation from the material aspects and situatedness of an interaction, use “triadic” to characterize interactions whenever infant, caregiver, and object are present, or even speak of “triadic” as a perspective adopted by a researcher, when analysing (any) interaction in terms of 3 poles: the participants and the “object”, the meaning or topic they negotiate (Rodriguez). Perhaps the most central characteristic of triadic interactions which all or most approaches could agree on is two people – in our case infant and caregiver – sharing a common focus of engagement. For most approaches, however, such a shared focus is necessary but still not sufficient, as the criterion of “shared focus” would, after all, already be fulfilled by a mother presenting an object to her infant and watching him play with it, without the infant showing any signs of engagement with the mother and hence not necessarily being aware of the mother at all (Adamson’s and Bakeman’s “passive joint engagement”). This definition might be sufficient for some accounts particularly emphasizing the caregiver (e.g. Rodriguez), however, most relational as well as cognitive approaches require the infant to show, in Adamson and Bakeman’s terms, some active co-ordination of engagement between people and objects, as well. While Adamson and Bakeman’s terminology and theoretical definition is kept
deliberately broad, and remains slightly vague, but leaving room for various different forms of engagement, the practical examples they are giving are restricted to a much narrower understanding of attention co-ordination: gaze alternation while following the mother's instructions to put wooden figures into a wooden truck.

Indeed for most **cognitivist or cognitive psychology approaches focussing on and testing the infant's individual actions**, and (supposedly underlying) **capacities and knowledge**, the criteria for triadic interactions have been narrowed down and condensed to the shorthand formula of “[visual!] joint attention” (or “joint perception”, Tomasello), used sometimes for a specific behaviour, sometimes for a capacity. Joint attention is theoretically defined as “two people looking at something and knowing that they are looking at it together”, and typically **operationalized for coding and experimentation** as “gaze alternation” or “gaze checking”. It is during this process of operationalization that problems with definitions become apparent: as it became clear that “sharing the focus of attention” might not be enough, Bakeman and Adamson introduced the category “passive joint engagement”, for cases where the mother establishes the shared focus, to distinguish it from proper, active “co-operative joint engagement”. However, not even active gaze alternation was sufficient to conclude that an infant was knowingly involved in true joint attention, as it could be the result of external cues or a previously learned routine, leading researchers to further restrict analysis exclusively to non-cued, infant-initiated actions and explicitly excluding caregiver cued gaze and familiar routines (e.g. Tomasello & Carpenter, 2007; Sugarman-Bell, 1978), thus however excluding large parts of typical infant-caregiver-object interaction potentially playing a crucial role in learning and the development of triadic interaction.

Interestingly the early seminal approaches of Hubley and Trevarthen and Bakeman and Adamson, rooted in observing semi-naturalistic interactions, do not restrict themselves to gaze measures for assessing “joint visual attention”, but ended up using broad, complex terms to capture the complex ways of relating developing towards the end of the first year: “secondary
intersubjectivity” which is conceived of as: joint praxis, that is “acts on objects that are oriented to the attention or action of the other person”, in combination with interpersonal acts (Hubley & Trevarthen, 1979), as well as “coordinated joint engagement” which is considered to be constituted by “the infant being actively involved with and coordinating his/her attention to both another person and the object that person is involved with” (Bakeman & Adamson, 1984), both point to triadic interactions being complex multi-strand, multi-modal actions, themselves already based on simpler forms of joint relating to objects (joint praxis and passive joint engagement), which for better understanding should be looked into in more detail.

2.8.2 Underlying conceptions of thinking, acting, and co-ordination: One-dimensional vs. multi-modal, multi-strand

As already briefly mentioned above, more general notions of “co-ordination of engagement” from early seminal studies taking a closer look at (semi-)naturalistic interactions, have mostly been narrowed down to “gaze”, probably in the course of operationalizing action-and-attention-coordination to get observable markers for conducting controlled laboratory experiments. Thus gaze often becomes 1) the only modality looked at, coded, and analysed in observations, and 2) the main and often only modality presented to children in experimental setups testing behaviours such as gaze following, considered as markers for children’s social abilities. Even when gaze cues are not presented in isolation, but are accompanied by an exclamation of surprise “.h!”, or a particularly salient facial expression, it is mostly for the practical purposes of making the experiment work better, and analysis still typically remains restricted to gaze.

While it is interesting in itself that children are able to coordinate their attention and action appropriately and effectively even under/in such impoverished circumstances, and though gaze measures may work as a marker e.g. for predicting later diagnoses of autism (Mundy & Newell,
2007), we still, in order to understand the development of these coordination abilities, need to look at and analyse the much richer multimodal-multistrand action organisation of infant-caregiver-object interactions that infants are participating in from early on in everyday life.

In general, it is difficult to interpret these isolated gaze measures since we as of yet know very little about the role gaze plays in concert and dynamic interplay with the other modalities and strands of action in jointly co-ordinated object practices. The pioneering research of Michael Land and colleagues using light-weight, head-worn eye-trackers has described in detail how gaze is used to guide everyday activities such as tea making, sandwich making, driving... but apart from a few studies on sports has essentially been restricted to solitary activities (reviewed in Land, 2006). The role of gaze in interactions has mostly been looked at in terms of gaze cues for engaging and disengaging from dyadic conversations (Rossano, 2012), with very few if any studies looking at how gaze guides social interactions involving objects, at least outside of highly artificial laboratory contexts (but see e.g. Ho, Foulsham, & Kingstone, 2015; Wu, Bischof, & Kingstone, 2013).

Interestingly recent data from the first few studies looking at gaze interactions in infant-caregiver-object interactions staged in the laboratory (Yu & Smith, 2013; see section 2.2.6.4), indicated that infants in free play activities rarely engaged in the normative steps of “visual joint attention”, not engaging in gaze following but rather monitoring moving hands and objects directly, or – when freely roaming around on the floor – looking at objects and at the legs of talking adults who towered over them (Franchak, Kretch, Soska, Babcock, & Adolph, 2010).

Narrowing down and confining the question how joint acting (and thinking) is co-ordinated to the modality of gaze (in particular specific series of gaze events such as “joint attention” and “gaze checking”), seems to be part of a more general tendency found in (particularly classical) cognitive science and cognitive psychology and reflected in some of the approaches analysed here (see in particular section 4.2.4): It is the tendency to conceive of the processes of thinking and acting as a one-dimensional sequence of
discrete steps – a conception likely stemming from using language and rational instrumental problem solving – both involving (the production of) a highly focused, readily graspable sequence of discrete events - as a model for thinking and acting in general. Consequently the challenge of co-ordination mostly boils down to singling out and selecting the correct action unit from of an ever larger repertoire – and in doing so integrating multiple targets – and sequentially chaining them together into longer and longer stacks. While this conception has led to important insights and not least the take off of computers, to better understand the development of triadic interaction, what its challenges are and what it achieves in the context of the activities associated with it, such as language/symbol use, we need to extend this view: we need to reconsider the challenges and investigate the accomplishments of co-ordination as a continuous dynamic orchestrating of multiple modalities and strands of action, distributing and integrating them flexibly, intra- as well as interpersonally (see chapters 6-7).

2.8.3 Empirical evidence base and ecological validity

The empirical evidence most studies draw upon to account for the development of triadic interactions are actually based on a very small number of specific interaction contexts: Dyadic proto-conversation for early interactions, and (staged) object play for later ones (toys often provided by the researchers: Bakeman & Adamson, 1984; de Barbaro, Johnson, & Deák, 2013, Moro & Rodriguez 2004, Hubley 1983), with the latter also often part of explicitly pedagogical situations or at least guided by normative goals, i.e. word learning or imitation with objects (Sugarman, Zukow-Goldring, see sections 2.2.3 and 2.5.3).

While these activities are lend themselves for investigation as they can be easily re-created with minimal context in a lab and turned into a standardized procedure, and some of them might be particularly relevant for practical questions of education, they most probably constitute, however, only a specific small fraction of and hence may not be representative for the wide range of activities infants and their families
engage in together. These activities might have very different settings and activity structures, enabling and constraining infant-caregiver-object interactions in different ways, giving rise to different forms of co-ordination and hence need to be looked into in more detail.

2.8.3.1 Outlook: how the theoretical frameworks can guide empirical studies

Thus, in order to better understand the early development of joint object practice (how it is established, sustained, and organised), and how triadicity develops, that is how infants participating in these interactions increasingly come to share and collaboratively co-create a world, we suggest to extend the existing research by:

1. Rather than focussing exclusively on (lab suited) staged object play or explicitly pedagogical situations, compare different ecological activities of everyday life and (analyse) how these different ecological activity contexts enable and constrain specific actions.

2. Rather than analysing infants’ actions with regard to a presupposed normative goal, differentiate and characterize different goal characteristics of activities (e.g. goal-oriented task or open exploratory social play), and generally look at how the interaction unfolds, how it is established, and sustained and ended.

3. Rather than focussing on either infant or caregiver, look at the functional roles and contributions of all 3 poles of the interaction, their interrelatedness and how they change over time.

4. Rather than focussing exclusively on (visual) joint attention and conceiving of interactions as a one dimensional serial sequence, investigate in detail how interaction partners jointly co-ordinate complex multi-modal multi-strand interactions as they unfold, and how they jointly structure shared spaces of meaning and action, within which they orient each other towards common points of reference.
Part II An Empirical investigation into infant-caregiver-object interactions over the first year of life: a naturalistic, longitudinal study

Based on the results and implications of the conceptual analysis summarized at the end of the previous chapter, we designed and conducted a longitudinal study accompanying infants and caregivers in their everyday life between 3 and 12 months. Chapters 3 and 4 represent examples of different forms of analysis: first looking at co-ordination in one model-activity “book sharing” in detail, and then extending the analysis to and comparing co-ordination across different activity contexts.

3 Jointly structuring triadic spaces of meaning and action:
book sharing from 3 months on

Nicole Rossmanith, Alan Costall, Andreas F. Reichelt, Beatriz López and Vasudevi Reddy

Abstract

This study explores the emergence of triadic interactions through the example of book sharing. As part of a naturalistic study, 10 infants were visited in their homes from 3–12 months. We report that (1) book sharing as a form of infant-caregiver-object interaction occurred from as early as 3 months. Using qualitative video analysis at a micro-level adapting methodologies from conversation and interaction analysis, we demonstrate that caregivers and infants practiced book sharing in a highly co-ordinated way, with caregivers carving out interaction units and shaping actions into action arcs and infants actively participating and co-ordinating their

attention between mother and object from the beginning. We also (2) sketch a developmental trajectory of book sharing over the first year and show that the quality and dynamics of book sharing interactions underwent considerable change as the ecological situation was transformed in parallel with the infants’ development of attention and motor skills. Social book sharing interactions reached an early peak at 6 months with the infants becoming more active in the coordination of attention between caregiver and book. From 7 to 9 months, the infants shifted their interest largely to solitary object exploration, in parallel with newly emerging postural and object manipulation skills, disrupting the social coordination and the cultural frame of book sharing. In the period from 9 to 12 months, social book interactions resurfaced, as infants began to effectively integrate manual object actions within the socially shared activity. In conclusion, to fully understand the development and qualities of triadic cultural activities such as book sharing, we need to look especially at the hitherto overlooked early period from 4 to 6 months, and investigate how shared spaces of meaning and action are structured together in and through interaction, creating the substrate for continuing cooperation and cultural learning.

3.1 INTRODUCTION

How do we arrive at a shared world? We jointly act in, communicate about, transform and co-create our world. In the process, we smoothly navigate and build complex networks of meaning-making involving persons, objects, and symbols. How do children grow in and into culture? How do they become competent participants in cultural practices, in networks of meaning-making including people and artifacts?

Researchers interested in cultural and social learning mostly start looking from the end of the first year, a period often characterized as a major shift, even revolution (“secondary intersubjectivity” Trevarthen & Hubley, 1978; “9 month revolution” Tomasello, 1999) in development, when infants engage in a number of qualitatively new ways of interacting such as jointly labeling things, following instructions, imitating acts on objects, or frequent gaze checking with their parents. At this point infants are credited
with engaging in true triadic interactions, and are considered capable of coordinating for the first time their engagements with objects and their engagement with people. The transition is often seen as the convergence of two lines of development considered to be separate before this point: dyadic infant-caregiver communication and infant-object interaction. This convergence is supposedly mediated by a newly emerging capacity for visual joint attention only then giving rise to conventional labeling and language use, conventional object use and symbolic activities in general, often associated with cultural learning. Interestingly, the seminal studies which constitute much of the empirical basis of this developmental narrative (Trevarthen & Hubley, 1978; Hubley & Trevarthen, 1979; Bakeman & Adamson, 1984), document early modes of combined social and object engagement termed joint praxis and passive joint engagement, respectively. Looking at the data reported, the studies actually show a gradual rather than revolutionary shift toward active triadic engagement on the part of the infant. Hubley and Trevarthen describe how caregivers first introduce their own body (games of the person) and later objects (marking and animating them) as a third pole into their social engagement with their infants. Adamson and Bakeman (1984) document how caregivers change their marking of objects over the course of the first year toward more conventional forms. These data have begun to be picked up on only very recently (De Barbaro et al., 2013; Nomikou, Rohlfing, & Szufnarowska, 2013; see also Moro & Rodríguez, 2004; Zukow-Goldring, 2012). The standard narrative has also recently been challenged by experimental studies documenting aspects of labeling, and joint attention in infants already at 6 months (Bergelson & Swingley, 2012; Striano & Reid, 2009).

Here we take book sharing as a model activity to explore the development of triadic infant-caregiver-object interactions. In a longitudinal study looking at infants’ everyday life activities from 3 to 12 months, this activity turned out to be one of the earliest social interactions involving a complex object, occurring from as early as 3 months.

This early occurrence raises the question: how can infants who are preverbal, do not yet understand the referential character of pictures, and—
supposedly—do not have command of joint attention, meaningfully participate in a book sharing activity? As one of the earliest jointly practiced cultural object routines, book sharing provides an excellent model for exploring (1) how a joint object activity is practiced and sustained between asymmetric interaction partners; (2) as an inherently semiotic activity\(^\text{14}\), involving the guiding and mutually orienting of attention, and shared meaning, it allows us to explore how triadic interactions involving mutual coordination and orientation toward common points of reference develop over the first year of life. While there is an extensive literature on picture book sharing, most studies start looking toward the end of the first year (Ninio, 1980; Fletcher & Reese, 2005; but see van Kleeck, Alexander, Vigil, & Templeton, 1996), and primarily focus on educational achievements associated with the cultural technology of book reading such as labeling and word learning, picture understanding, and literacy skill.

Here we focus on how the activity of book sharing unfolds, how caregiver, infant, and book respectively guide, sustain, and constrain the unfolding interaction. Taking the interaction as our level of analysis, we draw—in addition to approaches from developmental psychology—on concepts from embodied, situated, dynamical and enactive cognitive science (Fogel, 1993; Thelen & Smith, 1994; De Jaegher & Di Paolo, 2008), adapt methods from ethnography, conversation and interaction analysis (e.g. Goodwin, 2000; Alač, 2005; Streeck, Goodwin, & LeBaron, 2011; Deppermann, 2013), and use qualitative micro-analysis to explore how, from the interplay of multiple

\(^{14}\)The ability to engage with books is typically considered a complex cultural achievement based on the mastery of complex semiotic, that is, sign processes: pictures referring to “things” in the world, words referring to pictures and “things”, or “processes” in the world etc. Young infants’ meaningful participation in book sharing, long before they have mastered these complex semiotic skills, invites us to view them in a different way and investigate how they might be understood as gradually developing processes of increasingly complex mutual co-ordination.
modalities, shared spaces of meaning and action are created around objects and change over time.

3.2 MATERIALS AND METHODS

The book sharing activities documented in this paper have been collected as part of a naturalistic longitudinal study investigating the development of triadic infant-caregiver-object interactions over the first year of life especially focusing on conventional practices and encounters with everyday objects. Ten infants were visited in their homes once a month from 3 to 9 months of age and 7 of them up to the age of 12 months. A smaller pilot study with 6 infants at 3, 4, 5 as well as 9 months of age (3 located in Vienna, 3 in the UK, 4 girls, 4 first ones, 2 of them girls) was conducted in advance of the main study.

3.2.1 PARTICIPANTS

Of the 10 families participating in the study, 7 were from the UK and 3 from Austria. They were recruited from a wider circle of friends and family acquaintances, from mother and infant groups, as well as through word of mouth and flyers. All infants were living in middle class households with two caregivers and were raised in a monolingual (English or German) environment except one boy raised bilingually in German and Russian. The primary caregivers (mothers in all cases) all had tertiary education and took an active interest in supporting the infant’s education. Six of them (all in the UK) returned to either part time or full time work during the course of the study. Of the 10 infants 5 were female and 3 (2 boys and 1 girl) were first born. None of them had medical or cognitive problems.

3.2.2 HOME VISIT OBSERVATION PROCEDURE AND DATA COLLECTION

A typical home visit lasted 3–4 h, spanning 1–2 sleep-wake cycles of the infants. One to two observers accompanied infants and caregivers with a video camera (Panasonic HC-V500 in iframe format: 960 × 540 pixels resolution, 25 frames per second) documenting their everyday activities as they unfolded. For static situations a tripod camera mount was used, though for a large number of cases we switched to a handheld camera approach to
capture dynamic scenes especially after infants became mobile. Also, field notes were taken detailing the behavior of the infants, caregivers and siblings, including object and socially directed behavior, layout of the environment, and availability of objects such as toys and tools. In addition, reports from parents were collected giving additional background information on object use. The study was approved by the Psychology Research Ethics Committee of the University of Portsmouth, and was conducted in accordance with the 1964 Declaration of Helsinki and the Code of Human Research Ethics of the BPS. Parents provided written informed consent for the study.

3.2.3 DATA MANAGEMENT AND ANALYSIS

From these raw data, 300+ hours of video recordings, a video library was constructed in Final Cut Pro X (Apple Corporation). Episodes were tagged with keywords organizing activities into basic ecological activity categories, including (breast) feeding, diaper change, “witnessing,” soothing, social and/or object play, book sharing, sibling interaction, watching TV. In addition, infant-caregiver-object interactions as well as mutual coordination and orientation episodes were marked. For the purposes of this paper, “book sharing” was selected as a model activity for investigating the development of participation in joint cultural activities and coordination of triadic engagements.

In total 124 book interaction episodes (excluding 15 infant-researcher interactions) were identified and described. For an episode to be counted as a book interaction infants needed to be engaged with a book for at least 30 s. If after a period of disengagement—seen here as an integral part of (especially joint) activities (Stern, 1971; Brazelton et al., 1974; Tronick, 1989; De Jaegher & Di Paolo, 2007)—re-engagement did not occur within 30 s, the book interaction was considered to have ended at the point of disengagement. For all episodes, the actors (infant, mother, father, sibling, …), actions and objects used (types of books), as well as spatial configuration were cataloged. We distinguished between 2 different types of book interactions: (1) social book sharing (72 episodes), and (2) solitary book exploration (52 episodes). For a book interaction to count as social book
sharing the participants each had to be engaged with the book (via gaze or other book oriented actions, e.g., grasping, pointing to, or verbally referencing a page) and to coordinate their engagement, that is, to adjust their behavior in response to and in anticipation of each other’s—book or partner directed—actions (Bühler, 1927/2000; Fogel, 1993; De Jaegher & Di Paolo, 2007). For each type of book interaction, the number of occurrences and duration of the episodes was determined across ages and families, and basic analysis and visualization was performed using Python (numpy, scipy, and matplotlib packages, free software).

3.2.4 QUALITATIVE MICRO-ANALYSIS OF SELECTED EPISODES

Of the 72 social book sharing episodes, 20 episodes were selected for further qualitative analysis using the following criteria: (a) only caregiver-infant interactions without siblings to reduce complexity, (b) sampling of interactions from every age group, and (c) richness of interactions including attention and action coordination and communication. These selected episodes were transcribed and analyzed drawing on methods from conversation analysis and interaction analysis, adapted to the study of preverbal infants, with a special focus on embodiment and multimodality (Goodwin, 2000; Alač, 2005; Demuth, 2012; Deppermann, 2013). The analysis was performed in ELAN (free software, The Language Archive, Max-Planck-Institute for Psycholinguistics, Nijmegen, Brugman, Russel, & Nijmegen, 2004) with audio pitch and intensity extraction performed in Praat (free software, by Paul Boersma and David Weenink, University of Amsterdam).

The videos were repeatedly viewed and described in an iterative process looping back and forth between video and transcript (using ELAN), including gross description, and particular tiers for vocalization, audio pitch and intensity, action and gaze of caregiver and infant. Thus a multi-tiered, parallel record of the episode was constructed and visualized similar to a music score sheet, mapping a range of descriptors to the video stream and relating them to each other in time. Using these visualizations, we analyzed the sequential organization of the actions and how the various strands of an action, spanning multiple modalities, relate to each other and play together
in the coordination of action. Transcripts were compared across infants and ages. Some transcription and video stills from ELAN are also used for purposes of illustration.

3.3 RESULTS AND DISCUSSION

3.3.1 GENERAL RESULTS: POPULATION LEVEL RESULTS, THE “UMWELT” OF THE INFANTS AND THREE BOOK SHARING EXAMPLES

3.3.1.1 Population level results

Book sharing was practiced in all 10 participating families (ranging from 2 to 20 episodes per infant). We documented the activity from as early as 3 months (4 families) right from the beginning of the observation period, and no later than by 6 months for all families. To our knowledge, this is the first time book sharing interactions at this early age have been described in the literature. Social book sharing provided the context for infants’ first encounters with books. Later, in the second half of the first year, they also began to approach and interact with books on their own in solitary book exploration. Figure 3.1 (top) shows the number of occurrences of book interaction episodes for all infants observed in the longitudinal study, by age group and type (social or solitary). Note that we include these data to give an overview of the distribution of episodes forming the basis for the qualitative study. Also note the overall small sample size and that key variables such as the frequency of book sharing offers, and presence and comparability of books in the environment were not controlled for in the naturalistic study as would have been the case in an experimental study. Throughout we focus on two relatively robust measures to complement insights about the changing nature of book interactions gained from qualitative analysis: (1) the relative prevalence of social vs. solitary book interactions, and (2) the changes in mean episode duration over the course of the first year. While social book sharing interaction occurred from as early as 3 months, solitary book exploration episodes started to occur at around 6 months, displacing social book sharing as the dominant type of interaction at 8–9 months. From around 10 months on, social book sharing interactions became dominant again until a balance was reached at 12
months. **Figure 3.1** (middle) shows the mean durations (in seconds) of book sharing episodes for all infants, by age group and type. Starting from durations of around 2 and a half minutes at 3 months, mean durations increased considerably from 4 months reaching a peak of over 6 min at 6 months. From 7 months on, mean durations showed a sharp decrease, as book sharing interactions dropped by more than half to around 3 min duration and then stayed relatively constant. Social and solitary book interactions accounted for from around 1% (at 3 months) to around 5% (at 6 months) of the total recorded time that infants were awake on average at each month as shown in **Figure 3.1** (bottom), with their distributions largely reflecting the overall trend from social to solitary to balanced book interaction and the reduction in mean episode duration after 6 months.
Figure 3.1: Social book sharing interactions (red) and solitary book exploration episodes (green) for all infants from 3–12 months. **Top:** Number of episodes for each month. In addition, the absolute number ("n =") of individual infants represented in the sample (out of the total of 10) is also given below the bar graph, first in total and then in brackets for social and solitary interactions, respectively, as one infant may engage in multiple episodes in the course of a home visit. **Middle:** Mean episode durations for each month with standard errors of the mean (SEM). **Bottom:** Total book interaction time expressed as percentages of the total recorded t episodes in the course of a home visit. **Middle:** Mean episode durations for each month with standard errors of the mean (SEM). **Bottom:** Total book interaction time expressed as percentages of the total recorded time infants were awake at each month averaged over all infants. The figure gives an overview of the distributions of the documented book sharing episodes which form the basis of the qualitative study. Note the small sample size.
Before turning to the book sharing interactions in detail, we provide a sketch of the larger context of everyday life with a 3–4 month old infant as it presented itself in the study and is described in the literature. How do infants engage with their world at 3–4 months and what does their world look like at this age? At 3 months of age, infants are getting more and more interested in their surroundings. They have good control over their gaze (with a well developed oculomotor system) and increasingly look at and track objects in their environment (von Hofsten & Rosander, 1997).

At 3–4 months infants are, however, already fluent conversation partners: by then, they have already actively participated in dyadic proto-conversations with their caregivers for several weeks, fully utilizing and practicing all their capacities including gaze and facial expressions, vocalizations, and rhythmic coordinated whole body movements (Trevarthen, 1974; Bateson, 1975, 1979; Snow, 1977; Bullowa, 1979; Masataka, 2003). Not only are they aware of the dialogical, mutual give-and-take character of the interaction—getting upset when the mother’s face became unresponsive (Tronick, Als, Adamson, Wise, & Brazelton, 1978) or when confronted with a friendly but non-contingent (playback) response (Murray & Trevarthen, 1985)—but they are able to regulate their own state of arousal as well as the course of the interaction by turning their gaze and head toward or away from the caregiver (Stern, 1971) and even seem to be able to place their own vocalization exactly at the right time.

An interpretation in contemporary terms: Umwelt refers to those aspects of the environment an organism can interact with—i.e., effectively perceive, distinguish and act on (= the sum of prospective functional action-perception cycles)—and which hence constitute the organism’s meaningful world. This world is subjective, different organisms/subjects who have different histories and possibilities of interaction live in/enact different worlds.

---

15 Notion by Jakob von Uexküll (Uexküll, 1921; Uexküll, Kriszat, & Portmann, 1956).
and place at the right pitch in jointly created vocal phrases (S. N. Malloch, 2000; S. Malloch & Trevarthen, 2009).

As infants now take a wider interest in their surroundings (Trevarthen & Hubley, 1978)—in tandem with their increased waking and attentional periods—, while still lacking the means to pursue their active interests, to explore or manipulate the world on their own—they pose a new set of challenges and opportunities to caregivers. Therefore, at this stage a large part of caregiving activities observed in the longitudinal study—apart from feeding, diaper change and putting them to bed—was to keep infants content and “entertained”: the caregivers in the study responded to this challenge both by taking the infant to the world and by bringing the world to the infant. They did the former by taking the infants along with them, when doing their daily chores, e.g., placing them in a baby rocker, so they had a good view of the activities, regularly addressing them and bringing household objects or food items to their attention (e.g., rhythmically moving and labeling them) and occasionally also within their reach. They did the latter through presenting, looming and animating everyday life objects as well as specifically designed toys. Caregivers also placed them in specifically designed environments such as activity mats and baby-gyms where they were able to interact with objects dangling from toy bars. In contrast to their previous exposure to only a small range of objects, a whole range of new and manipulable objects now enter the infant’s world.

Thus infants were introduced to objects very early at 3–4 months in the context of social interactions. This was also the context in which infants first encountered picture books and book sharing, which took 2 different forms: (1) Their caregivers directly engaged them with books, often specifically designed for young infants. (2) They took part in the picture book reading activities of older siblings and caregivers.

### 3.3.1.3 Three examples of early book sharing interactions

**Figure 3.2** shows three instances of very early book sharing with 3-month-olds. Example A shows a 3-month-old boy vocalizing toward a black and white high contrast face pattern in a book specifically designed to
engage very young infants, even newborns, to meet their particular skills, needs, and interests. In the second example, B, a mother is rustling the crinkly pages of a brightly colored book to soothe her crying 3-month-old daughter. As the infant abruptly stops crying, she begins to engage her daughter in more conventional book sharing, drawing attention to pictures, turning pages, and inviting participation. The infant now and again grasps, holds onto, and crumples the soft pages producing more crinkling noise. In example C, after demonstrating page turning as an action of suspense and release—when a new page is revealed—the book is presented and held in place within the reach of the infant. The book with its rigid pages, solidly bound together at one end, provides a stable structure to interact with that is still highly flexible with easily movable parts along a single degree of freedom. This allows the infant not yet able to properly grasp an object to nevertheless effectively turn pages, thus exerting control over his sensory stimulation.
Figure 3.2: Three examples of book sharing with books specifically designed for young infants. (A) Visually engaging a 3-month-old with high contrast patterns. (B) Soothing a 3-month-old with crinkly pages. (C) Scaffolding a 3.5-month-old’s motor skills with rigid pages.
These three book sharing episodes are examples of early infant-caregiver-object interactions in everyday life, where the object – the book – plays a central role in the interaction. These books have been specifically designed to meet the infants’ needs: their physical properties are adapted to the infants’ perceptual capacities (high contrast patterns, crinkly pages), and serve as a scaffold for their rudimentary motor skills (rigid pages). In contrast to conventional books, this design emphasizes the effective interaction with the medium, the physical properties of the book and pragmatic actions performed on them. The specifically designed books serve as a bridge between the capacities and needs of infant and caregiver, as well as between caregiving and the cultural practice of reading. Indeed, in all three examples specific material aspects present in the book also capture and afford some of the general, mainly pragmatic aspects of conventional book reading; the format of the book itself is present, as is the format of the activity that has a definite beginning and end corresponding to working through a book from cover to cover, as well as the activity of page turning. Even more, already at 3–4 months, infants regularly experienced episodes involving the full range of book sharing typical for older children including more conventional, complex, and semiotic aspects such as pointing, content labeling, as well as reading and narration (Fletcher & Reese, 2005) as will be discussed in more detail in the next section.

3.3.2 EARLY OCCURRENCE OF SMOOTHLY COORDINATED BOOK SHARING INTERACTIONS AT 3–4 MONTHS OF AGE

Given young infants’ inability to interact with objects on their own yet – in contrast to their active role in proto-conversations – and the widely held theoretical view that they are not yet able to co-ordinate their engagement between people and objects (Hubley & Trevarthen, 1979; Bakeman & Adamson, 1984; Carpenter et al., 1998; Tomasello et al., 2005) the question now arises: How do book sharing interactions work at a micro-level, how do they unfold over time? How are they initiated and sustained, and what are the respective roles of the participants?
Figure 3.3: Mother multimodally presenting a book, holding it within reach of her infant: Introducing the book to the infant (A), marking the animals on the title page by dynamical pointing and vocal labeling (B-D), opening the book with the infant attending (E,F), more dynamical pointing drawing the infant’s attention (G,H), who subsequently acts on the book (I,J). Below the camera stills, an ELAN analysis detail documents, from top to bottom: audio traces (pitch in red and intensity in green), and annotation tiers. Tier label abbreviations used (from top to bottom): mothervoc: mother vocalizations, motherract: mother (manual) actions, babyact: infant actions, babygaze: infant gaze, and babyvoc: infant vocalizations. r: right, l: left.
3.3.2.1 The contribution of the caregivers: establishing contact, carving out interaction building blocks, patterning and shaping actions

Establishing contact. As shown above, caregivers were instrumental in introducing objects to very young infants who thus far are unable to approach or handle them on their own. Often caregivers took their cue from the infants’ behavior: either following up on infants’ gaze or action impulses, or, conversely, in trying to divert them out of their current state (e.g. pain) caregivers moved to establish contact between the infant and an object to engage with and build up a shared activity around it.

In the example shown in Figure 3.3 the mother visually presents a book to her 4-month-old son, who is sitting between her legs leaning against her, and puts it in his reach. She starts with a sharp intake of breath indicating surprise (”.h”) (Zukow, 1982), then, pointing dynamically by moving her left index finger up and down over the pictures of the book cover, follows this up with “Look at the cats,” while the infant is looking at the book continuously. (For transcription conventions see glossary).

As shown in this example, establishing contact between infant and object often involved visual presentation, ranging from static “offering,” placing an object into the infant’s view and reach, to more dynamic actions including “animating” the object, such as moving it to and fro, looming, or acting on the object. In the case of books, which were seldom animated by mothers, this prominently included performing dynamical pointing gestures, as in the example above. In addition, caregivers produced a number of different vocalizations ranging from general and unspecific exclamations of surprise (”.h”), via imperatives (“Look!”), questions (“What’s that?”) to specific labels for objects or object parts (“a book!”), and content such as pictures (“an elephant!”). Among these, the most frequently used in the dataset was a sharp intake of breath indicating surprise (”.h”) combined with raised eye brows, wide eyes and open mouth.

Functionally speaking, caregivers are doing two things at once. First, they are capturing and directing the infant’s attention, often utilizing the auditory domain to highlight and mark the visual presentation of an object. Second,
they are making an object available to the infants to interact with “as a unit”—in this case the book itself or one of its parts. Such actions actively foreground—or even create—the object for the infant to interact with “as a unit” by “carving it out” against the background and various other ways to parse a scene, compare Zukow-Goldring’s notion of “educating attention” (Zukow-Goldring, 1997, 2006, 2012).

Thus guiding the infant’s attention and foregrounding or “carving out” “building blocks” to interact with, are two partly overlapping processes. They often involve performing a variety of activities composed of various strands of actions, which appeal to one or another of the infant’s modalities and which can either be used (a) in close succession or (b) simultaneously, adding one on top of each other combining them into a complex multimodal action. It is especially this multimodal structure of the activity, in particular invariant relations across modalities, which provides infants with opportunities to extract coherent perception and action units (Zukow-Goldring, 1997; Bahrick & Lickliter, 2012).

**Carving out interaction building blocks and embodying meaning.**

Book sharing, with its wide range of semiotically rich materials, physical spine-and-page-structure, pictures, spoken words, printed text, rhymes, narratives and referential acts is mostly about learning about, sharing, and negotiating “units” or “building blocks” to interact with, which form the public cultural interaction space. That is, these book related actions are very similar to “guiding attention and making objects available for interaction” described above; only many of the “units” forming the cultural interaction space are more abstract and are not directly graspable. Children become familiar with those “units,” how they relate to each other (pictures to pictures, words to words, pictures to words), and how all of these potentially map onto actions and relations in the world outside, and above all how to jointly manipulate and act upon them.

So how is book sharing practiced with an infant, who is preverbal, does not yet understand the referential character of pictures (DeLoache, Pierroutsakos, & Uttal, 2003) and – supposedly – does not have command of joint attention either? While, as described above, the books designed for
infants highlight particular physical properties adapted to their sensorimotor needs and interests, book sharing even at an early age is not at all restricted to interacting with an “interesting stimulus” or “object for manipulation.” Instead, young infants already encounter the whole range of book sharing actions.

In Figure 3.3 the mother is sitting on the floor supporting her 4-month-old infant boy between her outstretched legs. Throughout, she is closely following the prototypical book sharing protocol: reading out rhymed text, accompanied by additional pointing and labeling, as well as making comments relating the story to the infant’s life. On his part, the infant is intently looking at the pictures, his gaze drawn through dynamical pointing, and from time to time acts on the book, either by banging or grasping the pages, which gets transformed into page turning with the support of his mother.

Neither is the infant in this interaction merely exposed to an arbitrary set of interesting stimuli and action affordances, nor does the mother blindly follow the cultural conventions. Rather, at key points in the activity, the mother is making selected parts and aspects of content and the overarching narrative accessible to the infant, making them meaningful to him through embodying and enacting them and giving them patterns of affective salience and arousal.

Figure 3.4 shows the mother making characteristic animal actions “come alive” and accessible to her 4 month old son through enacting the essence of “leaping” and “jumping” – a rising motion – through a rising intonation contour “This is the speedy kangaroo, she jumps and she LEAPS,” “here’s a smooth gray dolphin jumping in the Air.”

Whereas in the above example the enactment takes place solely within the action medium of speech—typically utilized in picture book sharing—there are also much more extensive and thorough forms of enactment and embodiment.

In Figure 3.5 the mother tells her by now 5-month-old son about baby Humphrey having “a BI::g YA:::wn and a STREtch, going ’UAAAHH.’” First,
she utilizes prosody again, drawing out the words “Bl::g YA:::wn,” thus temporally expressing the extension of “bigness” and at the same time already enacting the yawn. But then, as the text itself goes on to onomatopoetically illustrate the yawn “going UAAAHHH” she adds another layer: turning to the infant, grasping first one hand and then the other and gently pulling them into a stretch while performing the yawn, she is embodying and enacting the meaning directly with the baby’s body.
Figure 3.4: ELAN analysis detail showing pitch (red) and intensity (green) curves. The mother is reading a picture book about animal actions to her 4-month-old son enacting the essence of “leaping” and “jumping” (a rising motion) through a rising intonation contour (highlighted).

Figure 3.5: ELAN analysis detail showing pitch (red) and intensity (green) curves. Mother enacting and embodying a “BI::g YA:::wn and a STREtch” vocally and through acting on the 5-month-old infant’s body (highlighted). Upper case letters (A–D) map upper row stills to ELAN time line.
In this case expressing “meaning” is no longer simply “talking about” something or “depicting” something but rather encompasses fully realizing the action itself. Only that in this special case the action of yawning and stretching, referenced in the book, is now happening in a different context than it usually would, i.e., when the infant is tired or being put to bed. Rather, this context is created and defined by the book. And as the mother is gently acting on her infant’s body, taking him through the motions of stretching and at the same time performing the yawn, mother and infant closely share the meaning and the action in the sense of taking part in and realizing it together (Alač, 2005; Zukow-Goldring, 2006, 2012; Zukow-Goldring & Arbib, 2007).

**Patterning actions and shaping actions into action arcs.** Describing how objects or rather “units for interaction” are carved out to form the building blocks of a shared meaning and action space covers only one aspect of how such a space is created. This section will explore how the actions the partners perform are themselves structured in the course of interaction, highlighting the dynamic form of the jointly structured interaction space.

Two aspects of “structuring of actions” can be distinguished: The first is the temporal patterning, punctuation, and “chunking” of actions, also leading to the creation of “events” in the flow of action (Nomikou & Rohlfing, 2011). Examples include: the rhythmic multimodal performance of a monkey noise (“OohOoh-Ooh-Ooh-Ooh”), the marking and highlighting of action parts by exclamations (“.h!,” “Look!”), the labeling of action parts (“now we TURN the page”), and direct invitations (“Can you turn the page?”). Second, beyond patterning and chunking, caregivers structure actions by continually shaping parts of activities into bigger or smaller dynamic “action arcs” with a beginning, build up, climax, and resolution (compare Brazelton et al., 1974; and notions of ‘vitality contour’ Stern, 2010; ‘narrative’ or ‘shared project’ Delafield-Butt & Gangopadhyay, 2013; Trevarthen & Delafield-Butt, 2013).

To illustrate this we will look at the example of page turning (**Figure 3.6**). The mother sets the stage by drawing attention through the surprise exclamation “.h!” and announcing the action of page turning with the
question: “What’s on the next page?” Then she starts developing the action arc: leaning forward, repeating the question followed by two more “.h!” surprise exclamations of increasing intensity and pitch, she builds up tension which is mirrored in the growing arousal of the infant, indicated by her increasing movement, body tension, and facial expression, culminating in her mouth dropping open and a sharp intake of breath just before the climax. After a short hesitation – drawing forth the tension still further – a sudden quick page turn releases the tension and the arc levels off and comes to a close in a soft, whispered “There we go,” coinciding with the infant relaxing and closing her mouth again.
Figure 3.6: ELAN analysis detail showing pitch (red) and intensity (green) curves. The mother is building up an action arc through surprise exclamations of increasing intensity and pitch before releasing the tension through a quick page turn. Her 3.5-month-old infant is responding with increased movement, body tension, mouth dropping open and sharp intake of breath before relaxing again. Upper case letters (A–E) map upper row stills to ELAN time line.
This shaping of action arcs is found across all kinds of actions and at different levels and multiple timescales within an activity, nested into one another. At a high level, the activity of book sharing as a whole can be considered as an “overarching” action arc structure defined by the physical arrangements of the pages to be turned from cover to cover as well as the organization of the narrative. A smaller scale action arc is defined by each double page, the unit visible at a given time, and often structured by a (rhyming) pair of lines, the first ending in a slight rise continued in one breath (enjambement) to the second one, and coming to a close in a fall in pitch and intensity. At the basic level, action arcs re-occur with any interaction unit, be it the turning of the page itself a literal rise and fall, labeling of a picture, posing of a question, etc. Relevant words were typically placed at the peak of an action arc, and infants often looked at the caregiver’s face at the peak of an action arc, as well as in a pause after an action arc’s closure.

3.3.2.2 What about the role of the infant?

To what extent do infants actively participate in early book sharing interactions?

As briefly discussed above, it was often the infants’ behavior which was prompting the caregiver to introduce an object into the interaction, which – in case the infant let him or herself be engaged – then led to a shared object activity. Such “active interest,” that is, staying content and maintaining attention on the activity might already be considered as a form of “active participation.” Though at this age attention could easily be drawn especially by moving stimuli and also easily wandered away from time to time, infants were already able to some extent to actively control their gaze and hence their engagements. That the shared activity indeed requires an active contribution on the part of the infant became evident from cases when they withhold participation – which did not only happen when they got fussy, but also when they lost interest and kept looking away – and then there simply would not be any shared activity.
When successfully engaged, infants typically were alert and showed “serious intent” with knit brows and widely opened eyes, the type of engagement Piaget (1962) described for the adaptive mode of being absorbed in – and letting oneself be “informed” – in object exploration. Thus – at least for the youngest infants in the study – this shared activity looked somewhat different from other social interactions (e.g., social games) of the same infants at the same age, where more explicit expressions of joy such as laughter were observed.

However, even though not a single case of laughter in relation to a book was observed before 6 months, there was some affective communication going on in book sharing at this age: besides serious intent, a neutral expression, and occasional cases of overall fussiness, there were several instances of infants and caregivers engaging in a mutually attuned build-up of arousal in which infants showed great excitement through their bodily movements (e.g., the example of page turning discussed above, see Figure 3.6). Later, from around 6 months, laughter and a whole range of facial expressions were observed in an intricate emotional interplay going on between book or story, mother and infant (see Section “Ecologies in transformation”).

While caregivers significantly shape book sharing activities with 3–4 month old infants by guiding attention, inviting and scaffolding actions, infants actively participate by showing “active interest” and being responsive, amenable to their caregivers lead, letting their attention and actions be guided, and readily accepting the caregivers’ invitations to engage with objects offered (compare De Barbaro et al., 2013).

Young infants also showed active participation in a more conventional sense in their active movements, especially manual object manipulation as far as it lay within their range of action. Whenever possible, such actions—e.g., getting hold of the edge of a page—were interpreted by the caregiver in terms of the culturally established book sharing framework (“Do you want to hold the book?,” “Can you turn the page?”), and shaped it into the frame of the book sharing activity as far as possible. These actions, however, also sometimes got in the way of the activity, especially when they could not be
made to fit the book sharing frame, as when infants would not let go of a page and their own actions became their primary focus of attention (see Section “Ecologies in transformation”).

3.3.2.3 The interaction unfolding in the interplay between infant, caregiver, and object

After discussing the roles of mother and infant separately let us now look at one example in more detail in order to see how infant, caregiver, and artifact come together and how – out of this interplay – an interaction arises.

In this 13 s sequence (see Figure 3.7) the mother is sitting on the couch with her 4-month-old boy sitting on her knee, facing away from her. Both are looking at an open picture book featuring brightly colored cat pictures and “touchy-feely” textures, which the mother is holding in front of the infant. The sequence begins with the mother rhythmically reading out a line in verse: “I love THIS friendly kitten with the VE::Lvety so::ft NO::::::::se.” thus turning it into a two arc structure: the first arc is dominated by the deictic “THIS” which—with a sudden increase in intensity and a slight ascend in pitch—stands out as a single accentuated peak (accompanied by a slight movement of the left thumb). Thereupon the infant focuses more closely on the left page of the book. The second arc is a more pronounced, with a gradual rise in pitch peaking in “VEL-vety” followed by a slow fall in pitch and a gradual decrease in the intensity of the mother’s vocalizing, during which she turns her head toward the infant. After his mother’s turn toward his face, just as she arrives at the end of an elongated, soft “NO::::::::se” forming the coda of the action arc, the infant turns his head and elevates his gaze toward his mother’s face. As his gaze arrives at her face with a slight delay, her gaze has already moved on to the next page, where her right index finger is now performing a dynamic pointing gesture moving up and down on the velvety textured nose, and the infant’s eyes follow there soon after.
Figure 3.7: ELAN analysis detail showing pitch (red) and intensity (green) curves. This book sharing interaction at 4 months unfolds as smooth interplay between the actions of caregiver and infant: the infant’s attention is drawn by pitch (“THIS,” arrow on the left), and after moving through an action arc looking up at mother’s face, the infant’s gaze is drawn back to book through dynamical pointing (arrow on the right).
There is a sustained social interaction going on revolving around an object. Both mother and infant – acting as autonomous agents – co-regulate each other and the activity – at the same time also shaped by the object and the cultural activity frame – in ways that sustain the interaction itself (in the sense of De Jaegher & Di Paolo, 2007). The interaction is asymmetric with the infant’s attention and gaze responding to and following the mother’s (object related) actions and the mother guiding the interaction, checking back with the infant and adapting her actions to the infant’s response. The interplay of actions has an overall smooth and orderly quality, even though the infant is slightly lagging behind in time; still the order of events in the activity is retained and meaningful for the participants, as the actions of each of them effectively serve as an affordance to the other’s next action (Zukow-Goldring, 2006, 2012; Rączaszek-Leonardi, Nomikou, & Rohlfing, 2013). The infant’s actions are also recognizable to the mother as turns in the context of a (culturally structured) conversation (Schegloff, 2007). The mother interprets and shapes the spontaneous behaviors of the infant to fit the cultural frame.

Like the earlier example interaction involving page turning (see Figure 3.6), this interaction is organized into action arcs, again clearly illustrated by the intonation curve (pitch and intensity). The relevant deictic “THIS” is placed at the peak of the arc; the infant shifts his gaze at that peak, as well as in the pause after the closure of the arc after “NO::::::::se.” It is well known from the literature on infant directed speech that the rise in pitch – approaching the peak of the arc – makes it more likely that infants shift their gaze and is often used as an invitation for turn-taking. (Ryan, 1978; Stern, Spieker, & MacKain, 1982; Ferrier, 1985; Papoušek, Papoušek, & Symmes, 1991). As infants and caregivers repeatedly move through action arcs together, they co-regulate and share arousal and excitement, as well as act out and experience the structure, shape, and dynamics of actions together.
3.4 ECOLOGIES IN TRANSFORMATION: SKETCHING A DEVELOPMENTAL TRAJECTORY OF BOOK SHARING OVER THE FIRST YEAR

Over the first year, the quality and dynamics of book sharing interactions underwent considerable change in tandem with motor development, amounting to transformations of the whole ecological setting including spatial configurations the strategies and behavior of the caregivers as well as the objects used. Some aspects of these changes have already been described in the first section, as they became manifest in gross measurements on the population level: book sharing episode durations slightly increased until 6 months, then sharply declined at 7 months. From around 6 months on, solitary interactions emerged and became the dominant type of book interactions at 8 months until social book sharing took over again at 10 months finally reaching a balance at 12 months (see Figure 3.1). These results closely match a series of qualitative changes observed in the course of the longitudinal study. This section will sketch a developmental trajectory of book sharing over the first year based on these changes. For this purpose, the data samples are pooled into four age groups in accordance with the newly observed interaction qualities in each period:

(1) 3–4 months: early coordinated interactions with infants actively engaged but following mothers’ lead cued by local dynamical events (described in previous parts).

(2) 5–6 months: richer interactions with increased infant participation and more fluent attention coordination, including (a) infants shifting their gaze back to the book without being cued, and (b) interspersed affective communicative exchanges related to the book.

(3) 6–9 months: social book sharing interactions turning largely into solitary book exploration with attention to own object actions, paralleling infants’ new autonomous object manipulation, posture, and locomotion.

(4) 9–12 months: reconstituted social book sharing: infants effectively integrate autonomous object actions – which become increasingly conventional – with the socially shared activity.
Each sub-section begins with a description of the newly observed interaction qualities in terms of the infant’s activities as well as the overall ecological setting. Selected example episodes are then described and analyzed in more detail to explore and discuss attention and action coordination processes. For an overview of the changing characteristics of book sharing over the first year of life see Figure 3.12.

3.4.1.1 5–6 months: an early peak at social book sharing interactions

From 5–6 months, the 2 months immediately following the early phase described in the previous sections, book sharing activities became richer, smoother, and more sophisticated in parallel with the infants’ developing motor and attention skills and the increasing routine and attunement between the partners. During active participation infants used manual manipulation more extensively, showed improved aim when grasping pages, and their page flipping became more fluent. The repertoire of book interactions was extended by the addition of newly emerging actions, motor schemes such as banging, rapid opening and closing of the fingers ("scratching") on the surface of the pages, and mouthing objects (which also began to have slightly disruptive effects on the otherwise smooth interaction). Still, these actions were largely shaped into the cultural frame by caregivers. Coordinating and switching attention between object and caregiver was performed more easily and effortlessly: infants now followed the caregiver's lead more fluently, with faster, better aimed gaze shifts from the object to the caregiver's hands or face – following his or her voice – and then looking back to the book again spontaneously, without necessarily being prompted by local, dynamical events created by the caregiver (see Figures 3.8-9 below).

In accordance with infants’ improving postural control and new ability to maintain a sitting position with only slight support, spatial configurations with the interaction partners facing each other at a 90° angle became more frequent. At the same time, mothers less frequently acted on the infants’ body (putting them through the motions of a specific action); rather, mothers used their own body and voice, especially their hands, to enact meaning and perform lively visual demonstrations (including the beginning
use of baby signs). In line with the increasing frequency and skill of infants’ object manipulations, books with touchy-feely textures and attached graspable objects became prominent, as did books made of real paper with audio-haptic crinkle.

**Figures 3.8-9** illustrate the new quality and range of book sharing interactions at 5, and especially 6 months with a focus on co-ordination of attention and of action.

In the first example (see **Figure 3.8**) the mother is sitting on the couch cross-legged with her 6-month-old daughter placed at a 90° angle in the hollow formed by the mother’s left leg with her back supported by the mother’s left thigh and a sofa cushion. They are both facing a small square paperback “Mr. Men and Miss Little” book with thin paper pages, which the mother is holding. Immediately after a sharp rise in the intonation curve (“er ist SO::stark” [“he is SO::strong”]), the infant turns her gaze upwards toward her mother’s face, who in turn responds with an eye-greeting and a more pronounced facial expression and affective intonation. They share and reinforce each other’s expression of surprise and amazement in voice and facial expression before first the infant and then the mother turn their gaze back to the book again.
Figure 3.8: ELAN analysis detail of book sharing interaction with 6-month-old infant sitting at a 90° angle on the mother's lap. Infant and mother looking at book together (A). Infant looking up at mother's face in conjunction with salient vocal event at (B). Affective communicative exchange with mutual reinforcement (C,D). The infant's gaze spontaneously returns the book (E), before mother's gaze returns there as well (F).
In the second example (see Figure 3.9), the mother and her 6-month-old son sitting on her lap at a 90° angle are sharing a book about animal noises and have just arrived at the last page. After setting the scene by “Who’s your favorite?” the mother starts curving her right hand with the fingertips pressed together through the air toward the infant – accompanied by “a bzzzzzz bzzzzy bee” – with her eyes fixated on the infant, who is still involved with the book, his left hand reaching for and touching the animal picture on the upper right corner of the right page. When the mother’s hand finally touches the infant’s belly, he turns his gaze and head to her hand and begins tracking her hand as she starts moving it with her fingers joined side by side in up and down waves acting out “... or a ssSSSSSSSSsssssssssssnake.” As the mother concludes her enactment of the snake, the infant looks up first at the mother’s mouth and then at her eyes, beginning to smile. He then turns his gaze to the book again, his smile broadening, shortly after being followed by the mother returning her gaze to the book.

Infants’ attention coordination becoming more fluent and guided by routine. In both examples the infant is responding to an aspect of the mother’s behavior related to the book, e.g., the intonation curve going up as part of the mother’s interpretation conjunction with salient vocal event at (B). Affective communicative exchange with mutual reinforcement (C,D). The infant’s gaze spontaneously returns the book (E), before mother’s gaze returns there as well (F) of the narrative. In a previous example at 4 months (Figure 3.7), the infant was responding to and following the mother’s salient actions but kept lagging slightly behind and so the mother’s gaze had already moved back to the book by the time the infant had shifted his gaze to his mother’s face. In contrast, this time the eyes of mother and infant meet, facilitated by the 90° configuration and the infant’s more fluent movement. The infant thus elicits a communicative exchange of affect, including mutual acknowledgement and reinforcement. Also in contrast to the previous interactions, in both these cases it is now the infant who first turns his/her gaze back to the book again, before the mother does....
While infants, despite their growing motor skills, are still unable to autonomously move in and explore the world of objects, they are now turning their gaze and head more fluently from book to the caregiver’s hand or face and back again. They do so spontaneously, without necessarily being cued by dynamical movements, but arguably guided by routine, at times even arriving back at the book first, taking the lead in coordinating attention. Thus, within these interactions, infants demonstrate a basic understanding of the activity as shared and of the spatiotemporal structure and format of the book sharing activity at hand. The examples at 6 months also invite us to consider how small changes in the temporal dynamics of the interaction can lead to profound qualitative shifts as infants’ more fluent gaze coordination enables episodes of affective communicative emotional exchanges, and thus increase the infants’ ability to effectively shape the interaction dynamics of the book sharing activity.
Figure 3.9: ELAN analysis detail of book sharing interaction with 6-month-old infant sitting at a 90° angle on mother’s lap. Mother using extensive voice and hand acting to illustrate animals and animal sounds (“a bzzzzz bzzzzy bee,” “a ssSSSSSSSSsssssssssssssnake,” A–C). Infant gaze alternating between book, hand, mouth, and eyes (gaze targets inscribed on still images). After communicative affective exchange (D,E) spontaneously looking back to the book (F) before mother shifts her gaze back there (G).
**Interspersed affective communicative exchanges related to the book.**

Whereas at 3–4 months, infants showed “serious intent” when engaging in book sharing interactions, along with these novel communicative exchanges, infants now show pronounced affective exchanges.

While the mother narrates the story, in the short span of 5 min the infant displays and moves through a whole range of emotions in rapid succession, in concordance with the mother’s tone of voice, her gestures and movements: from surprise and amazement to amusement, and from being “staggered” to concern and sadness (see Figure 3.10). The emotions build up and develop in the flow of the interaction. In response to the mother’s voice and actions the infant looks up to her face with an expression of surprise, for example after an abrupt rise in pitch contour in “SO::strong,” the mother takes up her daughter’s expression and responds to it with widely opened eyes, raised eye-brows, and a sharp intake of breath indicating surprise (.h). She then repeats the passage that drew her daughter’s attention to her “SO::strong,” again with exaggerated pitch contour, reinforcing and further shaping her daughter’s emotion, thus acknowledging and reinforcing each other (compare Stern, 1985; Jensen, 2014 this issue).

So they were moving through the emotions together without however seeming to be seriously upset or sad. Importantly, these communicative exchanges are situated in the book sharing context, immediately following and leading back into attentional engagement with the book. Thus, the exchanging of emotions appears clearly linked to the book, and even to constitute a jointly relating to and negotiating “about” the book (see general discussion below).

3.4.1.2 6–9 months: shifting attention to object exploration

During the next few months, however, roughly in the period between 6 and 9 months of age, the interaction dynamics of infant-caregiver-object interactions underwent a significant transformation and the course of the developmental trajectory took a sharp turn: infant-object-caregiver interactions decreased in number relative to solitary book exploration, and
book sharing interactions showed a considerable decrease in duration and appeared generally less smooth compared to the period before, in spite of the infants further developing their capacity to sustain attention (see Figures 3.11B,E).

These changes occurred in a period when the infants’ developing strength and postural control allowed them to adopt and maintain a stable sitting position for longer periods of time, enabling them to reach and grasp and bimanually manipulate objects without falling over. Also, many infants at this age started locomoting by rolling and (“army”) crawling, and actively initiated interactions in a clearly visible way. The 7-month-old girl in Figure 3.11A for example, noticing a book sharing interaction taking place between her mother and sister, glances over her shoulder, rolls over from back to belly, and crawls across the room toward the book (still held by her mother but abandoned by now by her older sibling), thereby prompting her mother – albeit without explicit social signals – to start a book sharing interaction. Infants were also better able to focus and maintain their attention – see the 6-month-old boy in Figure 3.11B intently watching his mother's stroking a texture and closing in to see better. However, they were also more likely to quickly terminate interactions as their newly developed autonomous object exploration and locomotion activities drew them into new attentional engagements. In Figure 3.11C the same 6-month-old, after sitting back up again, accidentally touches a toy ring, subsequently grasps it and – with his eyes still on the book – brings it to his mouth, at which point his gaze is finally distracted away from the book and he becomes pre-occupied with exploring the ring, bringing the book sharing activity to a halt.
Figure 3.10: Book sharing interaction with 6-month-old infant sitting at a 90° angle on mother’s lap including extensive voice and hand acting.
Still images showing sequence of emotional exchanges: going in rapid succession and hand in hand with the mother’s tone of voice and movement when narrating the story, the infant moves from surprise, amazement, to amusement, and from being staggered to concern and sadness.
In this period, facilitated by the now stable sitting posture, infants got at times deeply involved with objects, e.g., banging, mouthing and manipulating books or other objects in solitary play to the extent of seemingly ignoring people: having escaped from a book sharing interaction after barely 2 min the boy in Figure 3.11E engages in manipulating a single object for nearly 6 min without interruption immediately afterwards. Infants did, however, from time to time look up at people’s faces, e.g., when introduced to an object, or in what might be early forms of instrumental looking: after having pushed a book out of reach, a 6-month-old girl lying on her belly turned her head up to her mother’s face and vocalized.

These changes were also reflected in the caregiver’s behavior: they were now often content to leave the infants to their solitary play. When they did try to engage them in book sharing, their efforts of directing attention became more vigorous: for example, they called their infant’s name repeatedly with increasing intensity to get the infant’s attention and resorted to acting on the infant’s body again, but now in an exaggerated fashion to keep the infant entertained. Caregivers also adapted by changing the situational context: for example, they tried to engage infants in book sharing interactions before bedtime, when infants are already tired, or changed the spatial configuration by placing infants on their lap, thereby actively constraining their action possibilities.

Books chosen by caregivers during this period had more interactive elements: in addition to the touchy-feely textures, flaps, and small graspable objects, they now included buttons producing various animal noises and moveable parts set on massive plastic pages eliciting blinking lights and nursery rhymes when operated correctly (Figure 3.11D). Thus, books are designed to invite manual exploration and multimodal interaction, drawing in infants now able to approach and engage with books on their own. On their part, caregivers included these highly salient object interaction opportunities in their social interactions to make them more interesting again to their infants with mixed results (Figure 3.11E).
3.4.1.3 9–12 months: putting books, caregivers and world back together

At 9–12 months, infants continued to engage in many solitary book interactions, but in contrast to the previous months, when they had primarily been exercising various motor schemes, banging, scratching, mouthing the book, as well as bimanually exploring books, they now started showing many more behaviors associated with conventional book interactions such as sitting still and looking at the pictures, turning pages, opening flaps, pointing at pictures, touching textures, and vocalizing.

Also in contrast to the previous period, the proportion of social book sharing episodes in relation to solitary ones increased again. Both solitary and social book interactions showed considerable variations in duration. Although the majority of the interactions were short, at times infants engaged in book interactions for extended periods lasting up to 7 min, as well as chained several episodes together into much longer lasting book activities. For example, they would ask for another round of looking at a specific book several times in a row, or, according to the mothers’ reports, entertain themselves during car journeys by looking at books and turning pages for extended periods of time.
Figure 3.11: (A) 7-month-old infant initiating book sharing by crawling toward the book. (B) 6-month-old, sitting freely, focusing on mother’s dynamical pointing and further closing in. (C) the 6-month-old in the same interaction getting distracted after accidentally touching and subsequently grasping and mouthing a toy ring. (D) 7-month-old absorbed in solitary play: correctly operating interaction device resulting in music and blinking. (E) 7-month-old and 9-month-old escaping from the book sharing activity despite their mother’s attempts to engage them. (F) 11-month-old proactively performing appropriate actions for “Pat the bunny”: putting his finger through the ring, sharing affect with his mother while making dolly’s ball squeak by banging on it, and “waving bye-bye” directed at the researcher, thus connecting the book sharing context with the visitor context. (G) Mother naming, pointing at, and signing “bird,” 12-month-old infant turning head looking out of the window while mother is still involved with the book, before mother turns her head recounting how they saw a bird out there the day before.
Book sharing episodes, even short ones, encompassed an increased number of action turns and showed a new quality and a larger degree of integration between interactions with the caregiver and with objects, between book and world and across time and space. Infants now more actively integrated manual object actions into their social engagements (e.g., approaching the mother with a book, laughing) and, when engaged with objects, now integrated social interactions (pointers, requests...), which may or may not include gaze alternations. Moreover, they were now actively bidding for and directing others’ attention.

Infants now moved pro-actively in the spatiotemporal attention-action framework of an activity: spontaneously performing appropriate actions in a specific context independent of temporal order, e.g. performing an action corresponding to a specific book page (“pat the bunny,” “put the finger through mommy's ring,” “wave goodbye” – see Figure 3.11F), and were also able to anticipate what came next. The infants’ actions extended much further over space and time, between the book and the world, while still being part of and coming back to the shared activity. For example, a 12-month-old boy interrupted his immediate engagement with the book, ran off and found the object depicted in the picture book and returned to mother and book. Or when the mother in Figure 3.11G is pointing out and signing “bird” referring to the picture in the book the 12-month-old infant is turning and looking out of the window. Not realizing this, the mother first finishes her signing, and then herself turns to look to the window recounting how they had encountered a bird there on the previous day.

### 3.5 CONCLUSIONS, GENERAL DISCUSSION, AND OUTLOOK

Our 3 main findings were:

1. Infant-caregiver-object interactions occurred from as early as 3 months. They unfolded as joint, mutually coordinated activities depending on the active contribution of all participants, and involved different kinds and degrees of attention as well as action coordination between co-participants and object.
(2) Over the course of the first year the quality and dynamics of book sharing interactions underwent considerable change in tandem with motor development, amounting to transformations of the whole ecological setting: book sharing episodes became more fluent and sophisticated until 6 months, after which there was a marked decrease in duration whereas solitary interactions became dominant, as infants developed novel postural, manipulation and locomotion skills and their attention shifted to learning to effectively act on the object world. Subsequently, social book sharing interactions resurfaced in the period from 9 to 12 month, showing novel qualities, as infants began to effectively integrate manual object actions – which also became increasingly conventional – within the socially shared activity.

(3) Our understanding of the emergence and development of triadic interactions and co-ordination and sharing of attention and action can be enhanced by looking at the larger ecological context, especially at the hitherto overlooked early period from 3 to 6 months and how shared spaces of meaning and action are structured together in and through interaction, creating the foundation for cooperation and cultural learning.

3.5.1.1 Development of triadic interactions

With regard to various theoretical accounts concerning the development of triadic interactions our observations suggest that:

Interactions with objects and interactions with people are not separated during the first year as often suggested in the literature (Bakeman & Adamson, 1984; Tomasello et al., 2005). On the contrary, at around 3 months when infants’ interests start to reach beyond the dyad but they lack the means to effectively interact with the material world on their own yet, objects are introduced by their caregivers in the context of social interactions.

Instead of a late, sudden appearance of triadic interactions at the end of the first year, we report a much more gradual development (compare
Striano & Reid, 2009; De Barbaro et al., 2013) albeit following a non-linear trajectory, characterized by an apparent dip after around 6 months followed by a recovery starting from 9 months; this would also explain why the earlier interactions have been largely overlooked in the literature.

The qualitative changes in the period between 9 and 12 months need a more differentiated conceptual framework as many of the criteria for triadicity – active contribution of the infant, coordination of attention and action between caregiver and object, etc. – already seem to be met by earlier interactions. Key notions need to be clarified and re-conceptualized, including: the nature of the infant's active contribution, infants' coordination of attention/orientation actions in relation to their coordination of manual actions and in particular the concept of joint attention.

**3–4 months.** At 3–4 months the infants showed active interest in the activity. They were responsive, amenable to and following the caregiver's lead, effectively co-ordinating their engagement between caregiver and object, their attention being drawn by local dynamical cues created by the caregiver (though following with slight delay) and their (rudimentary) manual actions were shaped into cultural frames by the caregiver. Thus the interaction was coordinated but asymmetric, smooth and orderly but slightly off-set (see Figure 3.12).

Accounts of infants' (lack of) triadic behavior at this early age do not begin to capture these intricacies revealed through the qualitative micro-analysis. For example, in Bakeman and Adamson's (1984) notion of passive joint engagement, the caregiver establishes and sustains the (passive) triadic interaction essentially all by herself. By turning to whatever the infant is engaged with or directing the infant's attention to a specific target, she ensures that infant and caregiver are "actively involved in the same object, but the baby evidences little awareness of the other's involvement or even presence." (p. 1281) In early book sharing, however, the infants were clearly not oblivious to the caregivers' presence, as evidenced by e.g., their regular gaze shifts between caregiver and object, drawn by the caregiver's voice and movements. Rather, early book sharing already comes close to their description of coordinated joint engagement characterized by the
infant being “actively involved with and coordinating his or her attention to both another person and the object that person is involved with.”

While it is arguable whether the responsive nature of the 3–4 month infant’s engagement completely matches this set of criteria introduced to describe the behavior of infants 9 months and older, by 5–6 months, infants’ active involvement was pronounced, especially with respect to their attention coordination.

5–6 months. At 5–6 months infants now coordinated their engagement between caregiver and object more fluently, and shifted their gaze back to the book by themselves without the need for a prompt arguably guided by routine. Their gaze often arrived back at the book first, thus at times leading the interaction. As faster gaze shifts led to meeting the caregiver’s eyes, infants now entered into affective exchanges and sequentially coordinated these exchanges with periods of shared object involvement. Despite their improved motor skills, infants were still unable to move in and explore the world of objects on their own. In book sharing, their range of manual contributions has expanded, including both helpful and disruptive actions, which were still mostly shaped into the cultural frame by their caregivers. Thus the interaction is co-ordinated and more symmetric with regard to attention, but asymmetric in terms of action, and overall orderly and fluent (see Figure 3.12).
<table>
<thead>
<tr>
<th>Skill</th>
<th>Infant</th>
<th>Caregiver</th>
<th>Books Used</th>
<th>Interaction on Level of Triad</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motor Skills</strong></td>
<td>attention: responsive coordination cued by dynamic, local events</td>
<td>guiding attention: carving out interaction units dynamic, multimodal, salient</td>
<td>sensory stimulative books: black &amp; white patterns, thin paper pages</td>
<td><strong>Constellation</strong>&lt;br&gt;infant on caregiver's lap both facing book&lt;br&gt;<strong>Resulting Interaction</strong>&lt;br&gt;sustained, co-ordinated, smooth orderly, slightly off-set (inf delay) asymmetric inf attention responsive (slight delay) inf action malleable/guided</td>
</tr>
<tr>
<td><strong>Motor Skills</strong></td>
<td>attention: fluent coordination guided by routine (a cultural frame)</td>
<td>guiding attention: carving out interaction units dynamic multimodal salient</td>
<td>sensory stimulative books: coloured pics&lt;br&gt;touch &amp; feel textures&lt;br&gt;hard pages</td>
<td><strong>Constellation</strong>&lt;br&gt;inf supine on floor&lt;br&gt;<strong>Resulting Interaction</strong>&lt;br&gt;sustained, co-ordinated, smooth orderly, fluent symmetric/asymmetric attention I symmetric action asymmetric: inf action guided</td>
</tr>
<tr>
<td><strong>Motor Skills</strong></td>
<td>attention: to own action drawn into own object actions, switching to others: sequential</td>
<td>intensifying attention getting: calling infant's name&lt;br&gt;acting on infant's body&lt;br&gt;changing context: before bed when infant is tired&lt;br&gt;on lap to restrain action&lt;br&gt;using special effect books</td>
<td>sensory stimulative books: coloured pics&lt;br&gt;touch &amp; feel textures&lt;br&gt;crinking &quot;real&quot; paper/books</td>
<td><strong>Constellation</strong>&lt;br&gt;free sitting on floor&lt;br&gt;<strong>Resulting Interaction</strong>&lt;br&gt;shorter, dis-co-ordinated, cross-purpose I symmetric attention I symmetric action I symmetric</td>
</tr>
</tbody>
</table>

**Note:**<br>Infant: 3-4m, 5-6m, 6-9m, 9-12m
Figure 3.12: Ecologies in transformation. The table gives an overview of book sharing as it changes over the first year. The columns list relevant characteristics for the respective participants: infant (inf): motor skills and book sharing actions sorted in attentional, manual and affective; caregiver (cg): book sharing actions in terms of function and modalities they are implemented in; books: type of book used; and for the interaction as a whole: the spatial configuration of the participants and the quality of the resulting interaction. The rows list the pooled age groups (3–4, 5–6, 6–9, 9–12 months).

Due to the interspersed affective exchanges, the interaction already resembles Hubley and Trevarthen's concept of secondary intersubjectivity, characterized by integrating “acts of joint praxis” around objects with “interpersonal communicative acts” (Hubley and Trevarthen, 1979). On the other hand, infants may not show enough manual object actions yet, and alternating back and forth between shared book involvement and communicative affective exchanges sequentially (see Figure 3.9) may not be “integrated” enough to match the criteria again set to describe the behavior of infants around 9 month and above.

Whatever the verdict on its “triadic” status, this alternation between engagements may constitute a basic form of “joint aboutness”—jointly communicating about something—which plays an important role in secondary intersubjectivity. It is also reminiscent of a crucial notion in Liebal and Carpenter's account of joint attention: one of its central features, “knowledge of knowing together,” is held to be established via what they call “sharing looks.” These looks close the triangle of the triad, turning “not-yet-shared attention into truly joint, shared attention,” confirming that attention is shared, with the goal of bringing about “an alignment of attitudes” (Carpenter and Liebal, 2011; compare Hobson, 2005). Their account again refers to infants at around 9 months and older and was not intended to capture the behavior of younger infants. Notably, social book sharing interactions at 6 months seem to already constitute a basic comment structure, in Bruner's terms (1975), in that infant and caregiver exchange affect in relation to, or even “jointly negotiate about” the book. Thus the affective exchanges in conjunction with the joint involvement with the book, its pictures, and vocal narrative might constitute a basic form of “content”
and the succession of emotional exchanges may build up toward a basic form of “emotional narrative.”

6–9 months. At 6–9 months, infants were actively seeking out and autonomously manipulating books, mostly engaging in solitary book exploration, with their attention primarily drawn to their own manual object actions, only at times looking up at their caregivers. Thus the social book sharing episodes were shorter, as the infants failed to keep up their engagement with the caregiver long enough to sustain the interaction. Though the interactions were now more symmetric, due to the infants’ more autonomous object manipulation, they were also less coordinated, at times dis-coordinated: when their caregivers attempted to guide them, infants were frequently already involved in an action, putting them at cross purposes (compare De Barbaro et al., 2013), and their manual actions could no longer easily be shaped into the cultural frame of book sharing (see Figure 3.12).

Looking at the period between 6 and 9 months revealed that the configuration commonly described in the literature for most of the first year does indeed occur: there was little joint or shared action as infants were drawn into deep object involvement to the point of seemingly “ignoring people” (e.g. Tomasello, 1999). However, when looked at more closely in the bigger ecological context, the apparent dip in triadic interactions at this point is not the beginning of the story but rather is only temporary, following a period of already well coordinated infant-caregiver-object interactions.

Rather than reflecting an enduring lack of cognitive capacities, the relative paucity of triadic interactions compared to solitary book sharing interactions between 6 and 9 months can hence be understood as a change of interaction dynamics due to new achievements (developing object manipulation, posture and mobility) and accordingly shifting interests. This shift of interest toward objects has long been known in the literature (Trevarthen & Hubley, 1978; Bakeman & Adamson, 1984). To characterize it (beyond noting basic correlations with infant postural and motor development) further investigations are required at the micro-
developmental level (see De Barbaro et al., 2013). The primary focus in the literature on the development of triadic interactions in terms of underlying cognitive capacities “coming on line” only later on explains why the diminished and discoordinated social object interactions at this age range are ignored and why the significance of early triadic interactions has been so often neglected and even overlooked (Tomasello et al., 2005; compare Reid and Striano, 2007).

9–12 months. At 9–12 months infants’ attention and action were guided not only through dynamical cues and routines but also by indirect and conventional means (words, instructions, demonstrations). Infants’ fluent coordination at this age incorporated manual object actions into social actions and social actions into manual object actions across different cultural activity frameworks, across time and space. Infants increasingly shaped and adapted their now versatile locomotion and object manipulation actions according to the conventional frame and to communicative exchanges, and were themselves actively directing others’ attention and action. The episodes were of varying duration, with a high frequency of action turns, and often chained together. The interactions were mostly coordinated and symmetric, orderly and fluent (see Figure 3.12).

This period clearly encompasses significant qualitative changes in the interactions. Rather than appearing suddenly supposedly mediated by a newly emerging capacity of joint attention, these changes can be seen as part of a gradual development (compare De Barbaro et al., 2013) coming out of the interplay of multiple strands of development in interaction with the social and cultural environment and the entire ecology of the activity.

In order to further explore and better understand the interplay of these multiples strands of development we need to reframe, refine, and expand key notions such as (visual) joint attention to create conceptual frameworks, which likewise allow for an interplay of multiple concepts capturing different aspects of the interactions, cultural activities, and their ecologies. For example, whereas the concept of joint attention, which developed in the context of experiments on gaze following and gaze checking (Scaife & Bruner, 1975), is primarily focused on the visual domain, processes such as
sharing of experience, attention coordination, mutual orienting can rely on multiple modalities bound together in structured actions. The role of gaze within this interplay of modalities is only beginning to be explored in more detail (e.g., social gaze to eye-hand-coordination in caregiver-infant-object interactions, Yu & Smith, 2013).

3.5.1.2 Jointly structuring shared spaces of meaning and action

The richness of early infant-caregiver-object interactions in naturalistic contexts invites an expansion of focus from the supposedly late emerging triadic interactions primarily associated with visual (joint) attention to studying how shared spaces of meaning and action are multi-modally structured together from early on.

The infants’ situation at 3–6 months (showing interest in their surroundings but not yet being able to explore the object world on their own) makes this age window particularly interesting for learning socially (including learning “about objects and the world”), as the infants readily engage in the highly structured and experientially rich joint activities offered by their caregivers.

Book sharing is such an activity. It serves as a “container” holding infant, caregiver, and world together in a small confined space opening up possibilities for shared experience and action and fostering learning (Wood, Bruner, & Ross, 1976; Vygotsky, 1978a). In pointing actions, for example, rather than having to follow a pointing finger to a distant target, the close encounters of early book sharing allow the finger pointing and the object pointed at to meet in immediate vicinity and within the infant’s reach, often accompanied by salient, dynamical gestures and actual, audible contact events. The container offers a rich reservoir of – and substrate for creating – interaction structures which are easily accessible to learn from and act upon together (Shotter, 1983; Goodwin, 2013). Part of this (spatial as well as temporal) structuring is provided by the cultural book sharing framework created around and manifested in the artifact book. Not only does the book invite the infants to physically engage with it (scaffolding their manual actions), it also embodies and reliably reproduces a stable, recognizable and
predictable sequence of actions. What makes the activity come alive is the caregivers’ active moment-to-moment structuring as they dynamically enact and carve out “building blocks” of interaction, pattern actions, and shape actions into action arcs in dialog with the infants.

The wealth of information available in infants’ natural environments has been emphasized by computational approaches in order to explain the impressive early achievements of infant learners, focusing primarily on the problem of word-reference learning (L. B. Smith, Suanda, & Yu, 2014). Also the statistical validity of social cues (caregivers’ action and gaze directions) for finding and disambiguating meaning in the complex cluttered streams of objects, actions, events – and words – has been shown using statistical learning models (Frank, Tenenbaum, & Fernald, 2012). Caregivers in real world activities actively select and structure their infant-directed speech, performing “auditory packaging” closely coupled to the relevant actions, creating crossmodal invariances, thus simplifying learning by highlighting relevant aspects within the interaction (Nomikou & Rohlfing, 2011; Bahrick & Lickliter, 2012; see also Leavens et al., 2014, this issue).

The present study invites us to take a step beyond the structuring of “perceptual input,” and consider the infant’s active, embodied participation and engagement in joint practices. Infants experience the activity first hand, actively seeking out and probing their environment through active vision and active touch. They are fully immersed and emotionally invested in coordinated interactions with their caregivers and the book, actively structuring shared spaces of meaning and action together. To describe this structuring in more detail we used the notion of “action arcs.” The basic arc structure with a beginning, build up, climax, and resolution is ubiquitous in physiological processes, e.g., breathing, and is fundamental to action, with different actions following different dynamic trajectories (compare Stern, 2010; Trevarthen & Delafield-Butt, 2013).

As infants and caregivers repeatedly move through action arcs together, they co-regulate and share arousal and excitement, as well as act out and experience the structure, shape, and dynamics of actions together. These types of co-regulation could be regarded as merely coordination of behavior
with sharing of affect (Tomasello et al., 2005). However, in moving through these arcs together, sharing of affect goes hand in hand with, and is inseparable from, learning about the structure of the action: infants become familiar with the dynamic trajectories as they are led through the motions, providing an opportunity to learn about structure and dynamics of actions, about themselves, their partner, the object involved, and their relation. Moreover, they get to experience and learn about the effects their own actions have on the partner and the unfolding of the activity.

Through such immersion in participation, infants are able to learn specific routines and practices, and more generally, “ways of interacting,” following the implicit norms of their culture (Mauss, 1973; Rietveld, 2008). It also provides the opportunity to learn about other people as social agents, whose actions significantly shape the unfolding of the activity. Through being drawn repeatedly by cues and movements to the relevant locations – hands, faces, objects – “where the action takes place” – infants become accustomed to and learn to anticipate the specific sequences of action trajectories (e.g., Hunnius & Bekkering, 2010), and the interplay of gaze, hand actions, and object use – in short how people act.

Crucially, infants are learning how to learn: when to look, where to get important information, and when to join in with an appropriate action (e.g., after a rising action at the peak of an action arc). Once established as interpersonal routines, action structures lend themselves to be played with, e.g., introducing temporal variations that violate expectations (as in teasing), thus highlighting and making explicit mutual coupling and coregulation, potentially helping to develop action coordination skills and cooperation (Reddy, 2008; Reddy, Liebal, Hicks, Jonnalagadda, & Chintalapuri, 2013). As active participants even in early interactions, infants become familiar with how to jointly structure activities and begin to learn how to negotiate and modify this shared structuring of activities. This skill, developed further, may be characteristic of how infants coordinate triadic interactions at 9–12 months, and crucial for cultural learning and culture creation.
3.6 ACKNOWLEDGMENTS

We are grateful to all the participating families and infants for sharing this precious time. We would like to thank Tina Reichelt and Elisabeth Zimmermann for support in data collection, and Hanne De Jaegher, Valentina Fantasia, Alessandra Fasulo, Joanna Rączaszek-Leonardi, Thomas Wiben Jensen, and Michael Schmitz for discussion, as well as the participants of data sharing sessions and research seminars in Portsmouth and Vienna. Many thanks also to the 3 reviewers for their helpful comments and suggestions. This work is supported by the Marie-Curie Initial Training Network, “TESIS: Toward an Embodied Science of InterSubjectivity” (FP7-PEOPLE-2010-ITN, 264828).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Received: 15 May 2014; accepted: 13 November 2014; published online: 10 December 2014.


This article was submitted to Cognitive Science, a section of the journal Frontiers in Psychology. Copyright © 2014 Rossmanith, Costall, Reichelt, López and Reddy. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

3.7 GLOSSARY

In the micro-analytic descriptions and ELAN illustrations some transcription conventions from conversation analysis were used where appropriate. (See Zukow, 1982; Jefferson, 2004)

? Question mark: rising intonation
. Full stop/period: falling intonation
, Comma: continuing intonation
! Exclamation mark: animated tone
AIr Upper case: increased loudness relative to surrounding sound
.h Period preceding h: audible inhalation, in particular: sharp intake of breath
    indicating surprise
Bl::g Colons: lengthening of preceding sound, the more colons, the longer.
4 Varieties of triads: different patterns of joint attention-and-action-co-ordination in different ecological infant-caregiver-object activity contexts and their development over the first year

4.1 Introduction

4.1.1 Background

As laid out in the introduction and the first chapter, the standard narrative – that infants begin to properly participate in joint activities and cultural learning (e.g. language and tool use) only from 9-12 months on – largely rests on 1) a rather narrow, rational conception of participation and co-ordination with 2) an exclusive focus on (visual) joint attention conceived of as the single crucial form of co-ordination newly emerging around 9 months, and is based on 3) only a small number of research settings: mostly a) proto-conversation for early interactions and b) (staged) object play and labelling for later ones, also often guided by normative goals, i.e. word learning or imitation with objects. In everyday life, however, infants are from early on involved in a wide range of social activities, which differ in terms of their settings and activity structures, and hence may afford different forms of attention-and-action-co-ordination.

In this chapter we particularly focus on the final, third point stated above and from there work our way up to revisit and re-evaluate the other assumptions.

The range and richness of social activities infants are involved in from early on was well-reflected in the 300h+ video material of 3-12 month infant-caregiver everyday activities collected in the course of our longitudinal study. Sampling the video material which had been collected into a video library we preliminarily tagged the interaction episodes using culturally established activity categories as far as available, such as feeding, nappy change, soothing, (social/object) play, book sharing, sibling interaction,
etc., and coming up with new working categories where necessary, e.g. "witnessing" other people’s activities.

In the previous chapter we reported how, as a first step, we singled out one of them: **book sharing** (due to its prototypically cultural characteristics and early widespread occurrence) as a **first model activity** for culturally shaped infant-caregiver-object interactions to be explored in more detail. Now we seek to **extend** the **analysis** from the example of book-sharing – testing, generalizing and adapting the concepts, methods, results and hypotheses (e.g. whether there generally is a dip in infant-caregiver-object interactions around 7-9 months) – to a **wider range of infants’ everyday social object interactions**.

### 4.1.2 Aims for the next steps

For the next steps we aim to: 1) describe and compare different forms of co-ordination and sense-making between infants, caregivers, and objects across different ecological activity contexts, and 2) document their variety (this chapter) and 3) sketch their development over infants’ first year of life (in chapter 5).

To do so we first need to develop a tentative conceptual framework, method, and vocabulary allowing us to address and distinctively characterize: a) different ecological activity contexts, as well as b) different forms of co-ordination.

### 4.1.3 Relevance and applications

Extending our knowledge about early social co-ordination beyond (visual) joint attention and beyond the small number of known contexts by a broader sampling and comparative analysis of social co-ordination in different ecological activity settings will contribute to both 1) basic research as well as 2) practical applications.

1) **Basic research**: Analysis across contexts may contribute to a) an ecologically more valid estimation of the range of everyday object activities which infants participate in, helping us to better situate and evaluate existing research as well as inform future research, as well as b) a better,
activity-context-specific understanding of different forms of social co-
ordination, how multiple modalities and strands of actions play together in
coordination and how particular forms of co-ordination are linked to
particular contexts, i.e. how contexts “shape” interaction.

2) **Practical applications**: Here, a) a better understanding of what
different contexts afford, and, conversely, which kinds of co-ordination work
well in a particular context with its particular affordances, may help to both
recognize specific communication and interaction difficulties and to provide
support e.g. through developing individual and situation tailored training
(e.g. for children on the autism spectrum or infants at risk). In addition, b)
investigating in more detail, how learning is shaped by and how it differs
across different contexts, may help us to be more aware and make diligent
use of (already existing) learning opportunities, as well as to design (new)
learning environments, particularly beneficial for specific forms of learning.

4.1.4 **Chapter overview:**

As the steps sketched above constitute a larger ongoing research project
exceeding the scope of the current thesis, we here will focus on:

1) Sketching the **development of a tentative prototype of a**
   **methodology for analysing co-ordination across different**
   **activity contexts** building on video analysis at both a macro- and
   micro-level

2) Introducing a **tentative framework of different ecological activity**
   **contexts** based on the **general results of macro-analysis** and
   summarizing those general results by giving an overview of the
   activity contexts’ respective **distinctive structural characteristics**.

3) Documenting the **process and results of macro-analysis in more**
   **detail for one examplar ecological activity context**: infant-
   directed caregiving tasks involving objects

4) Conducting a **detailed micro-analysis for nappy change** as an
   example of a particularly rich type of infant-directed caregiving task,
   analysing and illustrating the **complex dynamic patterns of multi-
   modal, multi-strand co-ordination** with distinctive functional
infant participation, which is already seen in specific joint object routines from 3 months.

5) Briefly discussing some of the implications from macro- and microanalysis: how particular activity-specific characteristics challenge currently established theory and conceptions.

4.2 Developing a conceptual framework and methodology: 1) A macro-level perspective

4.2.1 Developing a macro-level analysis

Proceeding from the youngest ages we now more systematically compared interactions within and across the culturally predefined activity categories they had preliminarily been assigned to. We approached each new video episode first looking at the situation as it presented itself to us as a whole searching for basic similarities and differences which – already at a global, macro-level - prove helpful to distinguish and distinctly characterize different activity contexts. Systematizing our observation notes about particular distinguishing features, we came up with the following macro-level factors to look at, which allow us to carve out general, distinctive structural characteristics of an activity as a whole:

In macro-analysis we specify: 1) more stable, slower changing structures of an ecological activity context: a) material ones: people and objects present, and their spatial configurations, and b)“ideal”, "procedural" ones: culturally shaped (normative) activity structures, less tangible yet recognizable to researchers as members of the target culture. To further differentiate within these activity structures, we look at what we tentatively called structuredness, governedness, directedness, and social framing.

Structuredness is used to describe the degree to which an activity is structured into more or less invariant subparts, and the extent to which variability is allowed or constrained. We further differentiate two aspects of structuredness: First, with respect to the outcome of an activity, i.e. to what extent the activity is a task, i.e. completion oriented, aimed at a specific endpoint and second, with respect to the process itself, to what extent there is a fixed sequence of events. Either or both of these two aspects may
display structuredness, with both of them contributing to the general “degree of structuredness”, which we conceived of as a gradient along which we distinguish 3 points: open, semi-structured, or highly structured activities.

*Governedness* seeks to capture what (primarily) shapes the process, the unfolding of the activity overall. Factors which regulate the interactions may include a) persons, i.e. needs, interests, inclinations of one or both participants, b) the object, whose structure may suggest a specific sequence of events, as well as c) practicalities of the activity, and cultural routines, norms, and rituals (e.g. “conversation” format or a specific nursery rhyme).

*Directedness & mediation:* who or what is an activity – or specific actions within an activity – functionally directed at, and, who or what is actually (directly) acted on (potentially serving as a mediator or instrument), thus directly or indirectly affected by the participants’ actions in the course of the activity.

*Social framing:* here this is meant to refer specifically to the way an activity is conceived of (both normatively in the respective culture and/or by the caregivers as revealed by their utterances and actions) in terms of its social structure: the designated participants and roles and relations (see nodes and targets of action next section), the appropriate degree of respective participation, specific contributions etc.

In this macro analysis we also specify 2) the more variable, faster changing functional roles and relations of the above structures with regard to the co-ordination of interaction in terms of a network of relating: *nodes* involved in interaction (*sources and targets of action*), their *relations* and *modes of relating/engagement*.
**Nodes of interaction/engagement:** what are the **sources and targets of action** involved in a particular interaction and how many nodes are there (dimensionality)? Typically the main **sources** from which actions originate are **people participating** in an interaction, occasionally however, objects too can act as sources of action, in particular automata such as "wind-up-toys", or when they are mediated by people as e.g. in the case of soft toys in a soft-toy conversation. **Targets of action** are people and/or objects, which the sources of action engage with in a particular activity, be it in an attentive-communicative way (e.g. *looked at as a focus of attention, verbally addressed as a conversation partner, spoken about as a topic of an utterance*) or in a physically-effective way (e.g. *grasped as a target of a manual action, tickled, wiped, kicked, kissed, etc.*). These roles are not mutually exclusive and can also change during an interaction.

**Connections between nodes:**

Which sources and targets of action enter into a **relation**, that is, co-regulate, affect and/or are affected by the other node during a particular interaction or activity? What is the **direction of engagement** in these relations, and are they uni- or bilateral? Are there relations that are **particularly emphasised** and **predominant** in a particular interaction or activity?

Examples: While e.g. *peekaboo* is all about the relation between infant and caregiver with objects mediating and modulating it, placing an *infant in a baby-gym* puts the relation between infant and object centre stage; whereas in...
witnessing a caregiver cook, the intrinsically predominant relation is that between caregiver, foods and kitchen utensils.

**Modes of engagement**

Further zooming in on the relations between particular nodes: how do they engage with each other, in particular which modalities and action strands are utilized and in which combination?

The following macro level distinction has proved useful to roughly capture and distinguish key differences of co-ordination in different activity contexts and over the course of development: a) **attentive-communicative engagement**, and b) **physically-effective engagement**:

a) **Attentive-communicative engagement** (indicated by dashed green arrows, sound by grey waves), used primarily for (distal) **orienting, control and monitoring of actions**, as well as for **communicating**, involving gaze, vocalizations, facial expressions, and to some extent whole-body movements.

b) **Physically-effective engagement** (indicated by solid red arrows), used to bring about changes in the material world, involving body and limb movements, up to fine manual manipulation, whether concerning objects and/or people.

This distinction runs perpendicular or even askew to the usual target based distinction between people- vs. object engagement (e.g. Trevarthen): Objects are indeed often addressed in an attentive-communicative way, just as people were found to be engaged in a physically-effective way, too. Therefore our distinction instead picks up on the different effector systems and ways of relating and co-ordination involved, thus also reflecting different functional roles in co-ordination (also showing aspects of but not being congruent with Kirsh’s distinction between **epistemic and pragmatic action** (Kirsh & Maglio, 1994) or, perhaps more so, the distinction between **acquisitory vs. effective action** (Mallot, 1997). As
these forms require different sensorimotor and co-ordination skills which develop at different points in time, this distinction is particularly useful to capture key changes in co-ordination over development (see in particular chapter 6).

4.2.2 Macro analysis general results: towards a framework characterizing activity contexts

Based on similarities and differences of interactions as assessed through macro analysis, 3 clusters of ecological activity contexts emerged for the earliest months: 1) infant directed caregiving tasks, 2) engaged witnessing of other people’s object interactions, and 3) establishing contact between infants and objects. Over time activity context 3) developed and spread into a wide range of social object-interactions, which we channelled into the categories: 4) object-mediated social activities, 5) socially-accompanied object activities, and 6) supported participation in cultural (object) practices. See figure 4.4 for illustration. The following paragraphs give a short overview of the activity contexts’ respective distinctive structural characteristics and sketch how particular activity-specific characteristics made visible by macro and micro analyses challenge currently established theory and conceptions.
4.2.2.1 Infant-directed primary caregiving tasks (Figure 4.4(1), see also section 4.2.3)

Not yet being able to tend to their basic bodily needs on their own, infants were from the beginning of the study involved in infant-directed primary caregiving tasks (such as nappy change, feeding, bathing), where they were acted on with objects by the caregivers. However, even in these completion oriented tasks mostly governed by the requirements necessary for completion, infants, when circumstances allowed it, were also invited to and indeed did actively participate in and even functionally contribute to the task with anticipatory whole body movements adapted to the particular context co-defined by the object, in particular to the mother’s actions on them and the object ("helpful" bottom lift in nappy change). See section 4.3 for more details.
Implications from macro/micro analyses challenging current conceptions:

With extensive infant-caregiver face-to-face engagement these infant-directed caregiving tasks do resemble dyadic interactions in some aspects, but through the object as a third pole new forms of coordination dynamics emerge which are not yet captured by - and not quite fitting into - current conceptions of the development of triadic interaction. In particular, the examples from the caregiving activity context, in particular nappy change, invite us a) to recognize/consider the crucial roles and interplay of multiple modalities and strands of action to create a “joint activity”. They also invite us b) to revisit established notions of participation, joint action and cooperation, most of which require complex object-related object acts (e.g. inserting wooden figures in a truck) and rework them into a range of concise, more differentiated notions suitable for addressing development (e.g. including whole-body movements appropriately placed in a complex context as forms of participation).

4.2.2.2 Involved witnessing of other people’s object actions (figure 4.4(2))

Given that 1) there were numerous family activities not directed at or directly concerning infants, and 2) infants - now awake for longer periods of time and increasingly interested in their surroundings - were not able to stay on their own and “entertain” themselves yet, infants often came to “witness”, that is, observe and experience other people’s (object) activities (e.g. household chores, family meals, or games and play activities of older siblings), which they were taken along to. While persistently pursuing the activity at hand, the requirements of which primarily governed the interactions, the caregivers actively maintained contact with the infants throughout the activity and (to varying degrees) adapted their object-actions to the infant’s presence (e.g. commenting their actions in strongly modulated infant-directed speech, showing objects, etc.).
Implications from macro/micro analyses challenging current conceptions:

What is interesting about this activity context is that infants in their early months spent considerable amounts of time in a social activity which differed significantly from dyadic face-to-face interaction typically associated with early social interactions. With regard to its configuration infant “witnessing” rather resembled “3rd person distanced observation”, except that, in contrast to how observing others from some distance is usually conceived of, it showed high degrees of involvement.

The witnessing context invites discussing: a) How even an activity, which does not seem particularly “joint” at first sight, is turned into a shared activity and jointly co-ordinated, and b) What learning opportunities such an – yet largely overlooked – activity context provides for infants, and, in turn, what role it might play in their learning to understand people, their actions, as well as objects, including their functions and names. This may happen through salient events inherent in the activity itself – object contact events or communicative events between other people – as well as through caregivers’ active patterning for the infant.

4.2.2.3 Establishing contact between infants and objects (figure 4.4(3))

Three to four month-old infants, getting interested in, but not yet able to effectively act on the material world, not only 1) were acted on with objects and actively participated in infant-directed primary caregiving tasks, and 2) engaged in witnessing other people’s (object) actions, but also 3) experienced a substantial number of interactions which primarily centred on establishing contact between infants and objects (toys and everyday objects), actively and persistently pursued by caregivers without any other visible goal. In contrast to the first two activity contexts, this one is largely open, governed primarily by the infant’s interest and affect (+ motor skills), as well as the characteristics of the material structures: the properties of the object(s) of interest and the form of scaffolding provided, constituting a specific spatial configuration affording specific
modes of engagement (distal, proximal), as well as going hand in hand with the **social framing** (here in particular the degree of caregiver involvement). These parameters of the setting showed substantial variation, mostly clustered in three sub-types each illustrated here with one example:

a) **jointly observing a mobile:**
distal object engagement with bodily+mechanical scaffold, and medium social involvement, indirectly affecting the infant via directly acting on object.

b) **(socially accompanied) object engagement in baby-gym:**
mechanically scaffolded, transient object engagement at hitting distance, autonomous but socially accompanied (low physical involvement by caregivers but: affect & action attunement responding - and structurally corresponding - to the infant's actions)

c) **embodying the (conventional) use of a rattle:** bodily scaffolded sustained proximal object engagement with high social involvement, where the mother acts on the infants arm holding the rattle moving it in synchrony with the culturally scripted utterance “shake, shake, shake.”

**Implications from macro/micro analyses challenging current conceptions:**

While the narrative assumes that infants’ interactions before 9 months are exclusively dyadic, or, even more stringently, that “objects” start to enter social interactions only from 6 months on (Bakeman & Adamson, 1984), and that true “sharing” of aspects of the world only starts from 9 months, the interactions reported here show that **infants already from 3 months on were actively introduced to objects by caregivers in socially scaffolded interactions.**

A particularly interesting aspect of this context inviting a more detailed look and discussion is that the **modes and patterns of engagement** within this activity contexts **differ widely, depending on** the respective **setting**, which might be particularly **relevant for** (investigating) **learning processes**, as well as for supporting, and designing environments for, learning processes. Notably these differences also include social framing (e.g. mothers regarded infants’ baby-gym object-engagements as
autonomous, physically interfered little, but did a large amount of affect-and-action attunement, whereas when introducing a single object they physically acted on and guided the infant frequently.)

Furthermore, even a very calm, rather passive example such as observing a mobile together – which does not seem very “joint”, “participatory”, nor active on the part of the infant when applying the classic criterion of visual joint attention (as gaze remains on the mobile) – demonstrates that we need to take into account other, in this case more proximal modes of engagement (grasping, touch) as well, to get a more complete understanding of engagement, co-ordination, and creating a “shared situation”.

4-5 months – objects getting into reach to be explored

Over the next weeks, as infants’ motor (reaching and grasping) and attention skills rapidly increased, activity context 3) “Establishing context between infants and objects” developed and spread out into a wide range of social object-interactions which then constituted a large part of the daily activities. Again based on macro analysis, in particular (depending on) the different role distributions among the nodes – which nodes/relations seemed at the centre of gravity, which nodes/relations were mediating – we further differentiated these activities channelling them into 3 categories: 4) object-mediated social activities, 5) socially-accompanied object activities, and 6) supported participation in cultural (object) practices. See figure 4.4 for illustration.

4.2.2.4 Object-mediated social interactions (figure 4.4(4))

In this new form of social object activity – starting from 3 months but emerging as a qualitatively new form from 4 month – the interaction itself and the relatedness of the participants form the “topic” and primary concern at the centre of the activity. In this they are similar to dyadic interactions, however, as in “establishing contact”, caregivers introduced objects (dolls, soft toys, as well as everyday objects such as cloth napkins, blankets, strings or paper rolls) into the interactions, and these objects were acted on by the participants in such a way that the quality of the
interaction and the relatedness of the participants was modified. The activities were governed by the participants (including objects, e.g. serving as participant in soft toy conversation) and to varying degrees by cultural scripts. Accordingly their structure(dness) ranged from largely open activities, where the participants playfully acted towards each other via the object (e.g. peeking and speaking through a paper roll) and let themselves be surprised by - and adaptively responded to – the novel quality of object mediated actions, to more or less structured ones, such as multi-person conversation, body part games, or peekaboo following and practicing cultural scripts.

Implications from macro/micro analyses challenging current conceptions:

Objects did play a prominent role in primarily social interactions from the beginning of the study and particularly from 4 months on, modifying and mediating them. In some cases objects were undermining the usual strict people-object distinction (e.g. soft-toys and body parts). The more structured interactions again invite to discuss how caregivers set up training containers which factually “embody” the essential normative features of the activity to be learned, naturally constraining and channelling infants’ actions towards normatively “correct” contributions, facilitating learning through participation: e.g. as the mother turned away from her infant entering into exclusive engagement with the soft-toy, the infant – not being responded to - “holds” her phrase, turns to the soft-toy, and, after “being attended to” by the soft-toy, finished her phrase in the direction of the soft-toy.

With infants’ participation now including new forms of object involvement this activity context also contributes to further differentiating the notions of infant action, in particular object engagement, and participation: in peekaboo the infant moves the blanket in relation to her own body, modifying visual engagement with her mother (body relative object act). The alternating division of labour between the participants shows how turn taking can gradually develop into role taking.
4.2.2.5 Socially-accompanied object interactions (figure 4.4(5))

As the most direct extension and further development of 3) “establishing contact” and in clear contrast to 4) object-mediated social interactions, this activity entirely revolves around infant-object engagement. While previously infants’ object interest crucially depended on caregivers’ active “establishing and sustaining contact” to become effective, now infant-object engagement – e.g. hitting and manipulating objects suspended in a baby-gym, handling and mouthing detached objects lying on the ground – has become more pronounced and autonomous with infants’ motor skills increasing and caregivers’ (physically effective) contributions mostly winding down to establishing initial contact or re-establishing lost contact; but even in such cases, we now observed infants actively showing pronounced object bids. Infants’ object interactions constituted a very open activity, primarily governed by their own interest and motor skills as well as the material properties of the objects which infants encountered and adapted to in their increasingly persistent and patterned object-directed actions. Yet most of the time this was not a solitary activity but socially accompanied with attentive-communicative engagement: caregivers, letting themselves be “informed” by the infant’s actions, responded with a similarly structured action in a different modality, reflecting back infants’ actions with slight variation in “affect and action” attunement.

Implications from macro/micro analyses challenging current conceptions:

Notable here is that even infants’ autonomous object explorations were in the early months to a large extent socially accompanied by caregivers. And here the infants’ exclusive visual focus on the object does not automatically render this into a solitary activity with infants not even being aware of other people, as it has often been raised as an argument for infants’s inability to co-ordinate engagement between people and objects
before 9 months (Bakeman & Adamson, 1984; Tomasello, 1999). Instead, caregivers’ verbally attuning to and reflecting-back infants’ affect and action signals connection without the need for visual checking, thus leaving the infant’s visual modality free for object exploration – powerfully demonstrating the value of division of labour between different modalities with their respective functional characteristics. That infants – while seemingly ignoring people – instead are indeed aware of their companions is indicated by examples such as a 5 month old crying in protest upon the “seemingly ignored” mother leaving the room. Moreover, the caregiver’s affect and action attunement and varied mirroring also marks and consolidates particular actions of the infant as shared action structures, thus creating and structuring a shared action space. This opens up the question of what this affect and action attunement contributes to infants’ development and learning of actions, self, relating, action-coordination, shared experience and knowledge.

4.2.2.6 Participating in cultural (object) practices (figure 4.4(6))

While all activities infants are involved in are of course part of culture, the practices described here e.g. book sharing, building a tower together and not merely feeding but enacting a shared meal with distributed roles, go beyond infancy in their relevance, as they were forms of cultural activity practiced by more mature members of a culture worthwhile to be trained in. Object and people engagement seemed to be balanced and of equal importance in this semi-open activities and subsumed under and governed by larger general cultural scripts, which included both social as well as object engagement as integral parts of the script.

Implications from macro/micro analysis challenging current conceptions:

Notably infants were immersed in and hence experienced the rich, full-fledged forms of these complex cultural activities. Fully enacting those activities went well beyond the current understanding and abilities of the
infant and far beyond what was needed for the interaction to function on a basic level.

Crucially, the other participants explicitly dedicated an active role to the infant: e.g. inventing a simple role fitted to the abilities of the infant e.g. throwing over the tower after building it, or helping the infant to “fill” an advanced role by scaffolding and/or enacting those parts of the activity, which the infant is not able to do yet in the name of the infant (e.g. a mother role-playing her infant and adding to the infant’s actions additional layers of conventional talk and action when having a meal together with the father (“hmm, that’s good, mummy, you’re a great cook.”)

Being in the middle of and fully immersed in these meaningful activities provides a rich source of information and a dense scaffold, enabling e.g. a 5 month old to intently monitor her sisters’ tower building including following the mother’s pointing across a distance far beyond the infants’ action space.

Outlook: 6-7 months – infants getting more autonomous, sitting, and starting to locomote

The developments during the next weeks resembled the “dip” we had noticed in social book sharing and can help to put this “dip” into a larger context: we now observed a “split” between infants enjoying exuberant social play (including object-mediated social interactions) whereas object interactions seemed to become increasingly solitary: Infants now initiated or re-elicited object-mediated social interactions (such as peekaboo and body-part games) on their own (even with relative strangers such as the researcher) demonstrating bigger independence from their immediate context and caregiver cues. Conversely, by the time they could sit on their own, infants’ autonomous object interactions – while still often socially accompanied, have increasingly gained the character of “solitary interactions” with a pronounced exclusive visual focus on the infant’s own bi-manual object exploration, apparently “ignoring” people. Yet instances such as the beginnings of instrumental looking when an object fell out of
reach, or a 7 month-old’s checking a stranger’s reaction to her object actions showed that infants did take into account other people besides their object interactions. Moreover – moving to participation in cultural object practices – infants did enjoy being introduced to and practicing particular more complex conventional object actions (albeit not understanding them as conventional in the full meaning of the term, of course) such as turning and removing a key from a cupboard, opening a door and checking who’s behind it, or assisting in watering the flowers. In general, infants were by now well orientated within the respective spatio-temporal structures of the different activity contexts as indicated by their anticipatory gaze and actions.

4.2.3 Macro analysis: more detailed results illustrated by the example of infant-directed caregiving routines involving objects

4.2.3.1 Introducing joint infant-directed caregiving routines involving objects

From the beginning of the study infants experienced (as they likely have from birth on a daily basis) being at the centre of caregiving activities involving objects: such as for example, changing the nappy, feeding, bathing, and dressing the infant. These activities are similar to dyadic interactions in that they were primarily directed at infants, but with objects entering in the interaction, at least one additional third pole or centre of interaction is introduced, opening up quite different forms of coordination dynamics.

Caregivers used a wide variety of objects - ranging from cloth napkins, bottles, sponges, nappies, clothes to changing mats and cushions, to mobiles, soft toys and music boxes - as instruments to help tending to infants’ basic bodily needs which they could not yet take care of themselves. Infants often encountered these objects through being acted upon with them: [e.g.] having dribble wiped off with a napkin, being gently brushed with a sponge or rubbed dry with a towel, having a bottle or spoon touching their lips or put into their mouth, having nappies or clothes attached to them, being put into sacks, swaddling cloths, or under blankets, etc...
As these activities were performed in service of meeting infants’ basic needs, they typically took the form of a **completion oriented task** with a **specific functional goal**, often reflected in a more or less **invariant activity structure**, such as a sequence of steps necessary to achieve the goal. However, neither needs nor goals were necessarily always clear, transparent, or stable for the participants. Caregiving routinely included figuring out why an infant was fussy and initiating caregiving activities based only at a guess of infants’ needs. Conversely, neither goals nor the structure of the activity were necessarily clear and transparent to infants, who still may engage in these activities for their own motivations.

Infants’ active participation in shared activities has been documented from an early age for **dyadic interactions**, where – utilizing their already well established vocalization, facial expression and whole-body-movement skills – they coordinate their actions with an interaction partner appropriately adjusting their actions in response to or anticipation of their partner’s actions. The examples in the literature can be subsumed under two different forms of co-ordination: 1) responding with complementary actions or generally performing specific context appropriate actions (turn taking in proto-conversation (Bateson, 1975, 1979; Tiffany Field & Fogel, 1982; Stern, 1971; Trevarthen, 1977), or adjusting to being picked up (Reddy, Markova, & Wallot, 2013; Negayama et al., 2015) and 2) responding with identical or similar actions (imitation, such as shaping (one’s) vocalizations, facial or manual actions in accordance to a model).

However relatively little is known yet about young infants’ **participation in early joint activities involving objects** including such common activities as infants’ and caregivers’ shared caregiving routines even though these constitute a substantial part of everyday living (but see Newson & Newson, 1975; Rączaszek-Leonardi et al., 2013; Valsiner, 1987).

Given infants’ still limited knowledge of goals and activity structures, limited postural and (fine) motor capabilities – let alone their supposed inability to co-ordinate engagement between people and objects – how do these caregiving activities unfold, how is joint praxis around objects co-
ordinated, and to what extent do infants actively participate, engage with the objects involved, and co-ordinate their engagement?

Working definition: joint infant-directed object routines

For this chapter we consider all activities where caregivers perform actions 1) concerning and directed at infants and 2) in which objects are a) functionally involved as instruments to help tending to infants’ (bodily) needs, and b) typically used to act on the infant.

We will investigate to what degree these activities can be regarded as joint, which requires active participation on the part of the infant (beyond being acted on and engaging in dyadic communication), that is, adjusting their actions in response to or anticipation of their partner’s actions as well as to the object. Also, to what extent do infants effectively contribute to the activity?

4.2.4 The common characteristics across caregiving routines in terms of sources and targets of action, their relations, spatial configuration, and modes of engagement

What these activities had in common was that infants typically experienced them on a one-on-one basis with their mothers (sometimes fathers, occasionally grandmothers, aunts or family friends) as primary actors and interaction partners. At times a second caregiver or an older sibling was present as well. Being directed at the infants these joint practices took place at a close (intimate) range, with the interaction partners in (direct) proximal contact with each other, as well as with most the objects involved. These included: 1) large objects used for postural support (and containment) providing infants with particular, context-specific experiences of large area surface contact and a supportive “ground” to act from such as soft (“hugging”) breast feeding cushions, smooth elastic changing mats, swinging baby rockers, smooth plastic bathing tubs filled
with water providing buoyancy but resisting movement, yielding bed mattresses, etc., 2) small detached movable objects, which infants were primarily acted on with, such as cloth napkins for wiping away dribble or playing peekaboo with, sponges, wet wipes, towels, nappies and clothes, blankets, bottles and spoons, and teething rings. Some of these were also among the first objects infants actively engaged with, either in a) individual engagement in parallel to an on-going caregiving activity (ranging from task-preserving distraction to task-disrupting interference), where particularly objects with a large soft surface area that are easy to get hold of were used, such as cloth napkins, clothes, towels and blankets; b) in joint engagement in between or after the main caregiving task (e.g. playing peekaboo with a napkin), or c) functionally contributing to the task, either by indirect engagement using whole body movements (e.g. bottom lift in nappy change) or, in rare cases, even manual engagement (e.g. engaging with teething ring after the mother has put it in the infant’s mouth, autonomously holding drinking bottle at three months, later, from 5 months, spoons, etc.), in particular using objects functionally designed for infant handling such as special drinking bottles.

In addition, a third group of objects was utilized: 3) movable objects used for distraction (mobiles, paper lanterns, music box, soft toy, radio, etc.), which formed a partial exception as they were mostly experienced from a distance, particularly in the early ages up to 5 months.

4.2.5 Differences within the context of caregiving in terms of settings/spatial configuration, task characteristics, and modes of engagement

Analysing and comparing the interactions within the main category of infant-directed object routines, we distinguished and pooled episodes into distinct caregiving contexts illustrated in figure 4.5: feeding and winding [row 1]; bathing, wiping dry and skincare [row 2]; soothing and putting infants to bed [row 3], nappy change and dressing [row 4]. Below we briefly describe these contexts in terms of a) setting and b) task characteristics, and c) ways of engagement between the sources and targets of action.
Figure 4.5: Illustrations of caregiving activities in (from top): feeding, bathing, soothing, and nappy change. Details see text.
Figure 4.5: extended caption:

Row 1: feeding, from left to right: breast, bottle, weaning (spoon, bottle)

1a: breastfeeding: mother and infant lying on the bed on their sides facing each other, the infant is drinking and pressing his hands against the mother’s breast, while his posture is supported by the mattress and the mother’s hands, 1b: breastfeeding: sitting cross-legged, the mother has placed the infant on her slanted right thigh, slightly elevating it and using an additional cushion to further support the infant’s back. While drinking the infant’s right hand alternatingly engages with the breast and the present cloth napkin.

1c: bottle feeding: the girl resting reclined on her mother’s lap, touches and grasps the bottle during bottle-feeding, while a is close by, which is used for wiping off dribble, sometimes manually engaged with, and occasionally used for a round of peekaboo.

1d: bottle feeding: rare case of 3 month old autonomously holding the bottle between her hands, as mother holds her supine on her lap.

1e: winding: infant engages with cloth napkin during winding

1f: spoon feeding reminiscent of dyadic interaction: 4 month old infant opens mouth in response to the mother’s exaggerated, affectionate mouth-opening-demonstrations, at which point the mother while gently holding down the infants hands, carefully directs the spoon into his mouth, thus together managing feeding with the infant staying primarily attentive-communicatively engaged with the mother, hardly looking at the spoon.

1g: weaning: a 5 month old girl presented with water in a bottle responds (in contrast to previous undifferentiated expressions of discontent) with a specific (when presented with water not milk) and persistent gesture of disgust and rejection (averting her face, frowning in disgust, crying, and pushing the bottle away with her arm).

1h: cultural activity of eating: the 5 month old infant is participating in the complex cultural activity of eating, staged as a role play by the parents,
where the mother enacts the infant’s role, translating her daughter’s non-verbal behaviour into possible verbal comments about the interaction.

**Row 2: bathing, wiping dry, skincare** 2a-b, caregivers’ manual resources invested in balancing their 3 month old sons in bathing tub (body submersed in, head above water) and washing infant’s body, while at the same time in a) communicating (vocals + facial expression) with the infant, who stays engaged in face-to-face communication throughout the activity, in b) monitoring and communicating with the infant, who keeps visually oriented straight ahead, as well as communicating with the older sibling and restraining his interference; 2c: Rare case of infant observing sponge actions in situation where both parents are present, hence more resources are available: while the mother is supporting the 5 month old infant (already sitting more upright, needing less support), the father is explicitly displaying the sponge, squeezing out water etc., drawing the infant’s gaze.

2d: 3 month old infant in sparsely water filled adult bathing tub gleefully thrashing arms and legs, laughing towards the mother who gives her time and space for thrashing before gently washing her

2e-h: wiping dry/skin care: as 5 min later the same infant from d), now supine on a changing mat on the floor, is wrapped in towels, rubbed dry, and wriggled, she is e) engaging in face-to-face communications with her mother, exchanging laughs, f) grasping and mouthing the towel, g) getting involved in and responding with laughter to a round of peekaboo. This illustrates how a relaxed situation after a task is often used as opportunity for social games, e.g. – in presence of any kind of cloth such as here – for a round of peekaboo.

**Row 3: soothing, putting infants to bed – changing or reducing highly involved engagement**

3a: soothing: the mother successfully uses a book with crinkly pages to distract her 3 month old daughter out of crying, after which the mother immediately starts to engage her in the cultural activity of book sharing.
3b: soothing: a teddy bear is animated to distract the 3 month old boy out of crying.

3e-f: going to bed: using mobiles, dim light, white noise and a movement restricting swaddling cloth (f) to change the 3 month old infants’ more intense ways of engaging with the world into calmer ones, facilitating falling asleep.

3g-h, 3 month old infant is covered with a blanket by the mother for going to sleep lying under the baby gym she had been engaged with before, the 4 month old infant autonomously pulls up the blanket over her face.

Row 4: nappy changing strategies: 4a-b: engaging infants in parallel distraction activities: using musical toys (including light effects) for distal audio-visual engagement, in addition infants manually engage with the towel (a) and the musical sheep toy (b) respectively.

4c-d: sequential, intermittent engagement in play breaks: mothers pause during nappy change to engage in play with their infants: animating a soft toy and touching her 3 month old daughter with it (c) or in case of the older 5 month old infant: the mother offers her own face for body part play.

4e-f: Task relevant person & object engagement: the mother invites the infants to actively contribute, leading to infants lifting their bottom in “a helpful way”.

4.2.5.1 Feeding and winding – close vicinity and/or coinciding of sources and targets of action, attentive-communicative and physical engagement coincide.

Functionally, the task of feeding primarily consists in enabling the infant (who is yet unable to do so autonomously) to engage in the fundamental life-sustaining activity of feeding by (1) providing digestible food (for the yet immature digestive tract), (2) establishing and sustaining contact between the infant and a food-source (with the infant’s postural, reaching, and locomotion skills for acquiring food still in statu nascendi), as well as (3) (mechanically) assisting with digestion afterwards (winding). Crucially,
feeding at the same time constituted a highly intimate social interaction (tending to the partners’ social and psychological needs and wellbeing) – especially in the early months – and over time, as the infant’s actions became more autonomous and their action repertoire grew, this activity increasingly became a container for transmitting cultural norms, as well.

In the first months of the study feeding was practiced as breast feeding (figure 4.5, row 1, a, b) or bottle feeding (figure 4.5, row 1, c-e). What characterised both forms and made feeding in the early months a very special interaction context is the closeness of the sources and targets of interaction, the proximal engagement through direct bodily contact and tight coupling via multiple channels and modalities: smell and touch, as well as sound and vision.

The infant was held on, or close to, the caregiver’s body in a stable position, reclined or lying on the side, supported by the caregiver’s body and sometimes by breastfeeding cushions or by a mattress (figure 4.5, row 1, a). Either the infant was placed and stabilized near the breast (figure 4.5, row 1, a, b) or a bottle was steadily held near the infant’s mouth (figure 4.5, row 1, c-e). While orally engaging with the food source, infants frequently brought their hands to it as well, touching and pushing the breast, or touching the bottle and/or the caregiver’s hands (figure 4.5, row 1, c,d) in rare cases even from 3 months on autonomously holding the bottle between their hands (figure 4.5, row 1, e), thus functionally contributing to the task.

Infants also frequently engaged with the ubiquitous cloth napkin, kept around to wipe off dribble and protect caregivers’ clothing during winding, grasping and mouthing it, caregivers also occasionally used it to engage infants in a round of peekaboo.

In the case of breastfeeding, engaging with the partner and engaging with the “object” (the source of food) coincided – and so did attentive-communicative and physically effective (bodily/manual) engagement to an extent: compared to the typical, distal visuo-audio-movement based forms of attentive-communicative engagement
characteristic of other activities such as proto-conversation, here proximal forms of engagement involving touch, close range smell etc., were prevalent. The breast\textsuperscript{16} was among the first targets of action infants engaged with – via mouth and hands – in a not only communicative but directly physically effective way, experiencing the result in receiving milk. At the same time, the infants’ actions were communicative as well, as the mothers were directly feeling and responding to them, verbally and by touch.

With the participants close and tightly coupled by proximal engagement, and the sources and targets of attentional-communicative and physically effective acting (at least partially) coinciding, the situation comes very close to Werner and Kaplan’s “primordial sharing situation” (1963). Thus to study how the partners co-ordinate and attune to each other in this situation would be of great interest in itself as well as with regard to development, however would require additional physiological measurements beyond camera recordings.

Of all caregiving situations tending to the infants’ bodily needs, feeding was arguably the situation where infants’ and caregivers’ interests and “goals” converged and overlapped the most, and where the situation is relatively transparent to the infant. And in this special situation the infant was able – by acting on the breast – to directly and autonomously (with just a little postural support) control and fulfil their need.

At the same time in the feeding situation (particularly later when feeding solid foods), conflicting forces and actions of resistance and refusal became particularly apparent, indicating differences and discrepancies in immediate interests, motives, and overarching functional goals between infant and mother, also partially related to conflicts with practical constraints and cultural norms.

\textsuperscript{16} For a detailed discussion of the special role and potential significance of the breast as a “transitory object” for the self-other differentiation and regulation see Winnicott (1971)
In the feeding situation it becomes visible how the infant's particular wants and not-wants become more focused, specific, and persistent over time, as illustrated in figure 4.5.r1g, where at 5 months the girl's undifferentiated expression of discontent from previous months has turned into a specific and persistent gesture of disgust and rejection whenever and only if presented with water (averting her face, frowning in disgust, crying, and pushing the bottle away with her arm). As infants’ motor actions grew more autonomous, caregivers on the one hand increasingly constrained the infant's actions to prevent disrupting interference (see the mother in figure 4.5.r1f gently holding down the 4 month old’s hands under the bib) (also compare Valsiner, 1987), and on the other hand actively shaped feeding interactions into collaborative activities of mutual co-ordination, with different configurations of engagement: The 4 month old infant in figure 4.5.r1f joyously opens his mouth in response to his mother's exaggerated, affective mouth opening demonstrations, at which point the mother, while gently holding down his hands, carefully directs the spoon into his mouth, thus successfully managing (the first spoon) feeding (interactions) together in a manner reminiscent of and sharing many features with dyadic interaction, particular in that the infant keeps primarily focused on and affectively communicating with the mother, hardly ever looking at the spoon.

With the slightly older 5 month old infant in figure 4.5.r1h the engagement configuration and dynamics is a different one: the infant mostly keeps her gaze steadily fixed on the spoon, which the father moves through the air in large swiping movements varying in speed, keeping his daughter's attention and highlighting and training the co-ordinative aspects of the interaction. While at the same time the parents stage a role play enacting the cultural activity of eating, with the mother playing the infant’s part and “translating” her non-verbal behaviour into verbal comments about the situation.
4.2.5.2 Bathing, wiping dry, skincare – Near-absence of infant-object engagement in a delicate situation while room for social and object engagement afterwards

The task of washing the baby was typically tended to by caregivers - bent over a baby bathing tub – carefully holding and stabilizing their infant’s head above the water with one hand and gently acting on the infant’s body with a sponge with the other, thus engaging simultaneously with and effectively bringing together infant and object. With most of their manual action resources invested in this delicate process, the mothers were additionally verbally inquiring into and commenting on the infant’s general experience or comforting and soothing the infant when needed, but did not particularly mark or point out the objects to the infants, keeping their object actions low-profile and close to the infant’s body (figure 4.5, row 2, a, b, d). The infants – floating in the water with their heads stabilized – mostly looked straight ahead (figure 4.5, row 2, b) or at their caregiver’s face (figure 4.5, row 2, a, d) while receiving the gentle touches of hands and sponges, but – apart from thrashing their legs in the water (figure 4.5, row 2, d) – hardly engaged with the objects involved, looking at the sponge only rarely (mostly when two caregivers were around) but not attempting to grasp it (figure 4.5, row 2, c). When laid down wrapped in and wiped dry with towels, however, they were communicating with their caregivers (figure 4.5, row 2, e) at times letting themselves be drawn into in a round of towel-peekaboo (figure 4.5, row 2, g-h), and readily grasped and mouthed (the corners of) towels, which together with napkins were among the earliest objects encountered by infants which they actively could get hold of. However, in doing so, infants were not functionally contributing to the core task.
4.2.5.3 Soothing and putting infants to bed – reducing highly involved engagement

The functional task at the centre of soothing and putting an infant to bed may be summarized as “disengagement”, that is to say a change or reduction of engagement, respectively. While in soothing caregivers seek to calm the infant and direct attention away from (engagement with) discomfort, when putting an infant to bed they reduce the infant's active and highly involved world engagements of being wide awake to a more passive and less involved state receptive for sleep. In addition to (dyadic) direct bodily regulation lowering arousal, such as holding, rocking, gently stroking and verbally soothing the infant, a range of objects is utilized in various different functional roles: for soothing small detached, movable objects are presented in a salient manner – rustling a book with crinkly pages (figure 4.5, row 3, a-c), vividly animating a soft-toy in large swiping movements (figure 4.5, row 3, d), or touching, tickling the infant with it – capturing the infant's attention and directing it away from discomfort, instead often encouraging sustained engagement with the object. In order to reduce infant’s intense engagements to a more passive state receptive for sleep when bringing infants to bed, objects are used to substitute highly involved intense engagement considered non beneficial with engagement of a slower, less intense kind, presenting the infant with a slow moving mobile (figure 4.5, row 3, e-f, a radio or musical toy playing soothing music, white noise (figure 4.5, row 3, f), or simulating heartbeat. Blankets and swaddling cloths were used to provide warmth and sustained touch and pressure, supposedly recreating the atmosphere in the womb (figure 4.5, row 3, f). The 4 month old infant in figure 4.5, row 3, g-h had even learned to cover herself with a blanket for going to sleep (after having been covered by the mother during the previous months, figure 4.5, row 3, g-h ), pulling it over her head (body-relative-object-translation), thus autonomously regulating her level of “input” and engagement.
4.2.5.4 Nappy change and dressing – a variety of configurations and dynamics of engagement: from constraining and engaging the infant in parallel distraction activity to active functional contribution through whole body movements.

Pursuing the task of cleaning the infant and providing him/her with a clean nappy, the caregivers are challenged by a number of further practical constraints: i.e. to prevent injury and contamination, that is, not having the infant roll off the changing table and not having the content of the nappy contaminating clothes or surroundings.

These task characteristics and constraints were also reflected on a cultural-material level in the situational setting, in particular as they were literally “built into” a number of artifacts especially designed to facilitate the process: Infants were placed (in a 180° or 90° orientation) on changing mats with soft, easily wipeable surfaces and elevated rims on washing machines (figure 4.5, row 4, a, f) or changing tables (figure 4.5, row 4, b, c) (with ready-for-use moist cleaning tissues close by), and often under or next to mobiles, lampions, musical toys (figure 4.5, row 4, a, b) (used to “keep the infant’s attention occupied” and contribute to creating a “recognizable nappy-changing context” - as explicitly stated by several parents.

Alternatively, infants were flexibly placed (at an 180° angle) on changing mats or mobile foldable changing bags laid out on the carpeted or tiled floor (figure 4.5, row 4, e, g) bed or couch, adjusted to situational need and spatial availability.

The caregivers in the study met the challenge of smoothly completing the task of nappy change with its specific constraints together with an infant who was active and mobile, but whose knowledge and motor skills are very limited and who is largely unaware of the task and its constraints, minimizing complications by employing different strategies of setting up and co-ordinating strands of engagement. These different strategies varied with particular circumstances – time available, location, mood of the infant, inclination and interaction history of the dyad, etc. – and resulted in
different configurations and modes of engagement, as well as different patterns and dynamics of co-ordination:

0) First of all, what caregivers did NOT do was focussing exclusively on the functional aspect of the task and acting on objects and infants exclusively instrumentally to quickly get the job done, rather caregivers acknowledged and engaged with the infants during the task as communication (and sometimes action) partners. The question is, however, how (exactly) these two or more strands of engagement were co-ordinated.

1) **Engagement in parallel distraction activity: the caregiver takes over sole control and handling of the task** by acting on infant and nappy, while preventing the infant from engaging and interfering with the task and its objects (nappy, clothes) by gently constraining his/her actions (e.g. holding the feet up and together) while engaging the infant in a **parallel distraction activity through** attentional-communicative or physically-effective (manual) engagement with an object and/or in pronounced distracting communication (figure 4.5, row 4, a, b).

2) **Sequential, intermittent engagement in play breaks: the caregiver intermittently takes over control and handling of the task for large stretches of time**, acting on infant and object, and gently constraining the infant’s action, while giving space at **interspersed intervals for the infant to actively and freely engage** in dyadic face to face and whole body communication, tickling, and body-part or object play (figure 4.5, row 4, c, d).

3) **Task relevant person and (indirect) object engagement.** The caregiver **partially shares control and task** by **inviting the infant to actively contribute in a functional way to the nappy changing activity,** by letting the infant engage actively with the caregiver, her actions, and – indirectly – the objects involved in nappy change (figure 4.5, row 4, e, f).
Context box: infant & caregiver’s everyday caregiving activities, culture society and economics

Concerning, in principle, all everyday infant caregiver activities, we will use the example of nappy change to once briefly illustrate 1) how aspects of culture permeate and impact on a young infant’s everyday life experience down to his/her physiological processes like the cycle of digestion, 2) how they shape, channel and mold infants’ and parent’s actions and activities into a cultural frame and 3) how vice versa the infant’s everyday life activities impact (back) on the larger society, global economical networks, and ecology.

The practice of nappy change indeed closely reflects the western way of living with its specific constraints: homes with carpeted floors, highly structured timetables etc. Through the disposable nappy, the infant’s needs are managed from the outside and subsumed under these constraints, at the same time illustrating a certain predilection for technological solutions, as the nappy with its absorbent polymer fabric is a relatively recent, and costly, high-tech design product. With between 3000 and 4000 nappies per child, disposable nappy sales are a considerable economic factor, and are even used as an indicator for the overall health of a country’s economy. Also they represent a considerable consumption of resources and result in 3.8 million tons (7.6 billion pounds) of nappy waste per year in the US alone, making up 2-30 percent of the landfill in Europe (Environmental Protection Agency (Ireland), 2009). Thus the product at the centre of the nappy-change activity is increasingly becoming a topic of growing ecological concern, illustrating the close interpenetration of global technological, economic, and ecological dynamics, cultural ways of living, and the everyday routines a young infant is involved in.
4.3 Micro Analysis: Bottom lift in nappy changing as an example for infants’ active participation in caregiving tasks

We chose nappy change as a model activity to explore in more detail how caregivers and infants organize and co-ordinate their early shared caregiving routines involving objects, as infants experienced nappy change from birth several times a day as a highly structured routine (see also Nomikou & Rohlfing 2011), also often used as a context for intimate dyadic interaction such as proto-conversation or social play (compare Bateson 1975). We selected the following bottom-lift episode - where an only 3 month-old infant shows a high level of active participation - to explore the range of and capabilities and opportunities for participation even of very young infants. However, bottom lift in nappy change was by no means an exceptional case, but was observed in 5 out of 16 infant-caregiver pairs.

4.3.1 Description and transcripts
Figure 4.6: Nappy changing activity, details see text.
The mother is kneeling on the floor in front of her 3-month-old daughter lying supine on a mobile, foldable changing mat (on the floor) and looking at her mother (see figure 4.6). Marking the beginning of the action with a sharp intake of breath “.h!” the mother opens the left nappy strap (to her right) with a loud tearing noise. As the infant keeps banging her right leg rhythmically on the floor she pauses, drawing back her left hand, (briefly) looking at the infant’s leg before facing the baby again, and asks: “What are you DOing?” (while reestablishing eye contact and looking at the infant’s face again). The mother then goes on - briefly glancing at the nappy - to open the second, right strap, before pausing again, making eye contact with and addressing the infant: “You’re ready? Are you REAdy?”, twice repeating the phrase with a rising intonation contour and with growing intensity. Now, as the mother is folding down the front part of the nappy with her right hand - marking the action with another “.h!” - the infant tucks up her legs sideways (with her head being pulled back, chin moving up a little as the legs are lifted), and then pulls them closer to her chest - right hand on her right knee - thus lifting her bottom off the mat. The mother pulls out the nappy from under the infant with her right hand while placing her left hand under the infant’s tucked up feet as a precaution, before beginning to fold the soiled nappy with both hands. Again looking into the infant’s eyes and holding mutual gaze she emphatically responds: “thank you! thank you!” accompanied by two pronounced nods with their "(up)-down-up” motions matching the intonation contours of the “thank you"s, while folding and putting away the nappy. The infant, meanwhile, has lowered her feet again, smiling.

This “bottom lift” as a convenient, “helpful” action coordination facilitating nappy change was observed in at least 4 different mother-infant-pairs from 3 months on, who had each developed their own idiosyncratic interaction patterns: not only did the mothers use different ways of structuring and vocally marking actions (from labeling the object about to be used “ein Windi [a nappy]”, to asking the infant for readiness or counting in “one, two three - <pause> -!“), also different styles and techniques of
bottom lifts were employed by the infants: e.g. the 3 month old boy in figure 4.5.r4f pushed his heels into the mat and arched his back thus effectively lifting his bottom from the mat. The same technique was observed in another 4 month-old girl, whose heel however came to rest on the nappy sliding away with it pulling it out under her again - this instance of a bottom lift thus ending up being disruptive to the activity rather than helpful.

**Later development**

At 4 months – despite being distracted by and continuously looking at the camera – as soon as the mother had opened the tapes with a tearing noise and folded down the middle part of the nappy the infant pulled up legs and bottom in one swift sweeping movement – apparently guided by auditory and tactile stimuli alone – again being immediately thanked by the mother.

At 6 months the same infant, in the meanwhile grown considerably, was again looking away, her gaze following the trajectory of her older brother running around. With the front part of the nappy already completely folded back and down, it was only upon her mother’s “ta”, by now established as a general cue in many situations, that the infant tucked up her legs (which also had considerably changed in proportion) by 90 degrees, however leaving the final bottom lift to her mother.

Moreover, by 6 months the nappy and in particular the closing straps for some babies had turned into something “graspable” and “openable”: one mother-infant pair, for example, developed a pattern in which the mother’s efforts to close the straps were repeatedly responded to by the infant’s re-opening them with a chuckle - despite the mother’s emphatic, though not quite serious, “no”s - considerably impeding nappy change but at the same time turning into a game of its own (see figure 4.5.r4g).
4.3.2 Microanalysis results: modes of engagement and dynamic patterns of multi-modal multi-strand co-ordination

4.3.2.1 The role and contributions of the caregiver

In line with the functional goal of the task (nappy change) centred on the relationship between object and infant, the mother regulates this relationship by manually acting directly on the nappy and on the infant, while the infant does not directly engage with the object herself (see figure 4.6). Also, the interaction shows many features typical of dyadic interactions: the partners are oriented to each other, and engage with each other through mutual gaze, vocalization, facial expressions, and whole body movements. The mother thus engages in multiple strands of action towards the infant and the object, requiring considerable amounts of co-ordination, which is also clearly reflected in her gaze behaviour: performing 8 rapid gaze shifts in this 13 sec excerpt, monitoring, acknowledging connection and communicating with her infant, as well as guiding her own instrumental actions, (which) she performs on the nappy, thus on her part displaying typical hallmark features of triadic engagement.

These surface characteristics already reflect how caregivers dealt with the challenge of performing a delicate task with specific constrains with an infant who is active but has limited motor skills, knowledge and awareness of the task. As mentioned above, rather than focussing exclusively on the functional aspect of the task and acting on objects and infants instrumentally to quickly get the job done, caregivers acknowledged and engaged with the infants during the task as interaction partners.

But how does the mother in this case co-ordinate these strands of engagement and set up the activity for the infant? And to what extent does the infant actively participate and/or engage in triadic co-ordination?

In this example the mother does not take over full control and handling of the task, rendering the infant rather passive (as the constraints and practicalities of the task in combination with the asymmetry in knowledge, skill and motivation might suggest, see strategies 1-2), but rather enacts
strategy 3 stated above: partially sharing the control and performance sub-actions of the task.

She explicitly acts to co-ordinate her own actions with her daughter as she would with an active co-participant – thus acknowledging her daughter as one – in such a way as competent (adult) members of her culture readily recognize as a cultural script/format of a co-operative interchange: while performing actions on the nappy herself, she is taking her daughter's actions (e.g. her kicking) into account, acknowledging them by vocally marking and inquiring about them “What are you doing?”, waiting for her daughter until she is ready, inviting her “Are you REAdy?” and giving her space to participate, and thanking her after her contribution.

What, however, is the situation like for the infant? What can the 3 months old unfamiliar with language as well as specific cultural pragmatic formats draw from all this? How can she understand any of this and what can this possibly mean to her?

Crucially, there are additional features inherent in the mother’s actions that go beyond what we readily perceive as the verbal surface (and action script) layer discussed above, which make the situation accessible to the infant as a structured shared activity she is part of and can actively take part in.

Indeed the mother creates a shared container or action landscape for participation using multimodal, analogue, affective means, helping the infant to orient and channeling the infant’s actions towards performing contextually beneficial actions:

First, she further highlights the structure of the activity, making it easier to parse and participate in (see also Nomikou & Rohlfing, 2011): she enhances the patterning already inherent in her manual actions by a) pausing and making eye contact in between steps and b) by verbally marking sub-actions, e.g. through adding an exclamation of surprise “...h!”.

The latter not only increases the salience of the respective sub-action, but also – by introducing an additional strand of action and letting the resulting multiple strands of action spanning different modalities rhythmically
coincide – makes each sub-action stand out as a separately perceivable multimodal invariance. Recurring across episodes of nappy change, these can over time be recognized and learned.

Second, the mother not only patterns the actions locally making events stand out, but also makes the course and dynamics of the actions, as well as the global structure of the task, perceivable by shaping the actions into dynamic action arcs (both at the level of sub-actions as well as that of the overarching activity) with beginning, build up, climax, and resolution (as first described for book-sharing activities by Rossmanith et al., 2014, chapter 3, see figures 3.6. and 4.7)

The mother opens the arc by undoing the first strap of the nappy with a tearing noise vocally marked by a sharp intake of breath “.h!”. After a short detour “What are you doing?”, she starts building up the arc undoing the second strap, inviting the infant in a whisper “you’re READY?”, then repeating the invitation a second time with increased intensity “Are you READY?”, further creating tension with its flat drawn out pitch contour and a sudden slight increase in pitch at the end. At this point of heightened tension the infant joins in: their co-ordinated actions – the infant lifting her bottom and the mother folding down the nappy and marking the co-ordinated action with another “.h!” – occur at and constitute the peak of the action arc. After that, the mother responds to and marks the successful completion of the infant’s part in the activity with two emphatic “Thank you!”s with pitch contours initially falling, as would be expected at the closure of an action arc, but then the contours rebound and move all the way to a high pitched “you”, each matching the “down-up” movement of the two exaggerated accompanying nods, to which the infant responds with a smile.

These particular glissando-like wide-range pitch movements, embodying excitement, in concert with the exaggerated nods, seem to 1) mark the “you”, that is, highlight the infant’s contribution with particular
saliency specifically attuned to a very young infant\textsuperscript{17}, and 2) reflect the temporary closure of the current action while already segueing into the next phases of the multi-step nappy-change activity, i.e. the infant still will be cleaned, provided with a new nappy and dressed again. Only then the mother will finally conclude the overarching action arc running through the activity with an emphatic “ALL done”, spoken twice while maintaining eye-contact with the infant, each with a marked fall across the whole pitch range, and accompanied by two synchronous gentle chest rubs – thus again creating specific cross-modal invariances with a specific contour and specific dynamic characteristics constituting a memorable “closure event”.

Thus moving together through this action arc, the format “nappy change” (itself an instance of a more general “co-operative exchange”) does not simply appear as a sequence of distinct events, but rather as an co-created arc, a specific trajectory of shared arousal, affect and action, jointly enacted, which hence becomes predictable (to some extent) and meaningful and over time gets established as a part of a shared action landscape.

This sharedness of action arcs is also already hinting at the third aspect of the mother’s active shaping of the interaction:

\textsuperscript{17} For the nappy change activities of this mother-infant dyad these extreme wide-range pitch contours were only observed at 3 months, while from 4 months on the mother’s “thank you”-contours matched the conventional ones much more closely. Similar exaggerated wide-range pitch contours were observed for other activities, other caregivers, even other languages (German), when acting with very young infants.
Third – while first patterning the activity through actions, vocalisations and eye-contact and second dynamically creating action arcs – she not only structures the activity so it is easier to parse and understand, but does it in such a way that lets the infant experience that the action concerns her, that she is involved in a shared activity she can actively participate in, addressing her verbally and by gaze and orchestrating various modalities and strands of action to binding herself, infant and object together in a shared stream of experience:

(By) manually acting on the object, visually alternating sequentially between object and infant, and vocally marking and patterning both her own actions on infant and object, as well as the infant’s actions, the mother splits up various modalities and strands of action and distributes them among different targets, but combines them into one rhythmical stream of action which effectively binds together infant, object, and herself into what the
infant can experience as one coherently patterned shared action process, so that the infant can experience herself as actively part of this shared process and - recurring across nappy changes over time - learn to understand being addressed and drawn into actions arcs at specific points as invitations and cues to participate with specific actions.

Moving through action arcs together might contribute to the “sharedness” of the activity and hence the involvement of the infant as a participant by 1) creating not only a similar experience of performing the action itself, but also similar arousal and affect in both partners in parallel through the dynamics of the action itself (changes in speed (action, sound), intensity (action impact and volume) and elevation (motion height and pitch), which is 2) probably mutually reflected and reinforced by the partners.

4.3.2.2 The role and contributions of the infant

But what about the infant? To what extent is she co-ordinating her engagement with the mother and the object? Or should we interpret the infant’s actions, who does not seem to manually engage with the object, as engaging in a dyadic interaction with the mother, the minimal requirement of which would be reciprocity, that is each one responding to the actions of the other? To what extent is the infant actively participating in or contributing to the task?

The infant is active from the beginning: she is (and throughout the activity primarily stays) visually engaged and in communicative connection with her mother, and (at the same time) performs unilateral kicking movements. As the kicking - waited out and commented on by the mother as not being part of the nappy change repertoire - recedes, the mother proceeds with the nappy change, and the infant - in response to her mother’s cues - moves into performing a whole-body-coordinated bottom lift. Rather than performing an action on the object directly by physical contact, manipulating or translating it (which she is not able to do yet), the infant is interacting with the object indirectly by changing her body posture (moving, translating, shifting parts of her body in relation to the ground and in relation to other parts of her body, (requiring whole-body-coordination)
in response to cues provided by her mother acting on the object thus anticipating and enabling the next action performed on the object by the mother, and thus she actively contributes to the task.

4.3.2.3 The role and contributions of the objects: nappies, mobiles, etc.

The nappy is an object which the infant is primarily acted on with, whose presence - as something “worn”, closely attached to the infant’s body - is experienced in terms of weight, temperature, texture, movement constraint etc., but which the infant – at least in the early months – is not actively engaging with. However - together with a number of other characteristic objects (e.g. the supporting soft surface lied upon, as well as distally (visually and/or auditory) perceived ones - like a mobile, lampion or radio, used to keep infants occupied and still - the nappy contributes to defining the specific “nappy-changing-context” and - as it is handled by the caregiver - provides cues (visual, haptic and auditory) for the infant’s actions.

So while not directly grasping or manipulating the nappy herself the infant nevertheless “interacts” with the object by appropriately responding to the cues arising from - and adapting his/her actions (whole body movements) to - the caregiver’s actions on the nappy. Thus the infant is changing the situation in such a way that allows the mother to perform the next object manipulation thereby propelling forward the interaction along the typical sequence of events.

As infants got older the nappy change context got enriched by more objects coming into reach as graspable ones, first and foremost body parts of the infants themselves, knees, hands, a little later feet that can be clasped, as well as the caregiver’s hands, face and hair, objects at hand (e.g. a shiny crinkling packet of wiping clothes or a spare nappy to be manipulated, often handed to infants to entertain them and keep them occupied), or nappies or clothes themselves (the latter sometimes with disruptive effects, see above).
Part III: Integration of empirical and theoretical parts

This part of the thesis seeks to integrate results from conceptual and empirical analyses and reconceptualize the problems and challenges considered at the heart triadic interactions and their development in terms of co-ordination. Chapters 5 and 6 explore through two examples how this approach can be applied longitudinally, what it can contribute to our understanding of the development of “self” and “other”, and participation, respectively.

5 Structure and Openness in the Development of Self in Infancy

Nicole Rossmanith and Vasudevi Reddy

Abstract: From early infancy, structures are created in engaging with the world. Increasingly complex forms of self, other, and world emerge with shared rhythms, affective patterns and interpersonal routines, cultural norms, concepts and symbols, and so on. These open up an increasing number of possibilities for new kinds and levels of engagement and for further developing a world together. However, these same structures, becoming more rigid, salient, and perhaps reified with time, may obscure or obstruct engagement and constrain development. We explore this paradoxical relationship between structure and openness to engagement and attempt to understand the process of formation and change of structures in self and its relations to the world.

This paper explores what we can learn from engagements in infancy and particularly about the role of structure and openness in the process of being

and developing as a self. Thinking about self in preverbal infants challenges us. It immediately invites us to explore nonconceptual aspects of self (in infants as well as adults) and offers connections to Buddhist notions about relation, process, and the self as an illusion. We take a relational and process view of self and explore how aspects of self arise in relation to the world and to others in the process of living, interacting, and separating. In such a view, self appears at times more pronounced and solid, at times more fleeting and permeable, having different boundaries or facets in different relations while becoming increasingly differentiated over time. We focus in particular on the intriguing tension between structure and openness in development: how patterns and structures are created in engagement and foster further engagement and change, but through becoming entrenched may also get in the way of engagement and change.

5.1 Patterns, Structure, and Openness

Patterns of action and interaction from everyday life may become established over time – even reified – as stable and relatively invariant ‘structures’. We use the notion of structure here because it vividly suggests a certain palpability: while patterns come and go, change, cancel each other out, and vanish, structure suggests something more stable. It carries with it a higher degree of persistence, of continued existence beyond the original interaction context, and a more perceivable effect on its surroundings: obstructing, guiding, and channelling processes. Importantly, the word also suggests that it can be an ‘object’ in itself to be engaged with and inviting further exploration and change (see also Goodwin, 2013; Shotter, 1983).

Structures might become incorporated as habits or skills or, on a larger scale, as personalities. They might become established as shared routines which are publicly accessible or exist as implicit norms only noticed from the reactions to breaking them. They might exist as material artefacts, as systems of patterns scaffolding each other such as concepts, language, beliefs, theories, and value systems, or as explicit conventionally established laws and institutions (Deacon, 1998). Although differing widely in kind, complexity, and composition, what all these structures have in common is
that they regulate the activities of everyday living (be it cooking, washing, working, playing, sleeping, having conversations, or moving in public spaces), they channel and constrain our acting and sense-making in specific ways and, as we enact them, they also contribute to constituting our selves and our multiple identities.

5.2 Methodological Challenges and Theoretical Stances on Self in Infancy

Enquiring into the ‘self’ in preverbal infants and trying to relate such an enquiry to other approaches studying the self poses a challenge: not only has it often been questioned whether preverbal infants even have a self, but, if they do, whether we can understand what it might be like. Their lack of speech can constrain psychologists’ and phenomenologists’ explorations of infancy.

Theories about the emergence of self-awareness or self–other awareness vary enormously in their emphases, their methods, and their conceptions of selfhood. Some theoretical positions in developmental psychology posit that objective self-awareness is necessary for an awareness of self as an object to others, and that this arises late in infancy (between the middle and the end of the second year). The implication is that prior to this awareness neither is the experience of being experienced as a self by another person possible, nor are the self-conscious emotions such as embarrassment, shame, and pride possible. This objective self-awareness is argued to depend on the prior emergence of a concept of self or ‘idea of me’ (Lewis, 1995), which is tested by infants’ recognition of the visual self in the mirror self-recognition task (Gallup Jr, 1968; see also Amsterdam, 1972). However, other theories argue for a multifaceted and relational approach to selfhood, positing, for example, that an initial ecological self emerges in the foetus’s and neonate’s interactions with the material world, followed by an interpersonal self in face-to-face interactions with people in the first months, then followed eventually by a conceptual self and a verbal and moral self (Neisser, 1993) after infancy. Others argue that the self is largely an affective entity developing within emotional exchanges from the earliest days of life (e.g.
The emphasis on affect as core to the self has been pursued by many theorists such as Daniel Stern (1985), challenging the dominant conceptual emphasis and arguing that self-conscious affectivity emerges prior to concepts and is either independent of a concept of self or actually enables its development (Izard & Hyson, 1986; R. P. Hobson, 1990; P. R. Hobson, Chidambi, Lee, & Meyer, 2006; Reddy, 2003). A focus on the linguistic aspects of self — the use of the personal pronouns ‘I’ and ‘You’ in particular, and the difficulty which toddlers have in reversing them appropriately — has also led to links with the particular difficulty children with autism have both in relation to pronoun reversal and with a typical sense of self and other (R. P. Hobson, 1990; R. P. Hobson & Meyer, 2005).

5.3 The Emergence of Self in Infant Engagement

Self–non-self differentiation: Although writers such as Freud, Piaget, and James proposed a process of selective differentiation of self, they largely operated under the assumption that infants initially have a completely undifferentiated sense of self. This assumption has since been challenged by a considerable number of empirical studies. For example, when touched on their cheek infants turn their head towards the touch (a rooting response facilitating breastfeeding), but, if the touch happens to come from their own moving hand, they turn their heads significantly less (Rochat & Hespos, 1997), thus showing a differentiation between sensory events which are related to their own actions (in this case involving a coinciding double touch on face and hand) and those which are not. Indeed infants are intensely interested in the consequences of their own actions and strive to maintain closed action-perception loops. When 4-week-old infants, lying in the dark, encounter a narrow beam of light which the researchers have shone either above their chest or their face, infants move their arms in the appropriate area to keep their hands in the light, exploring the relationship between moving their hand and seeing their hand move (van der Meer, 1997). Recent studies suggest that even foetuses not only anticipate the sensory consequences of their actions, but selectively guide their actions according to their particular expectations. 4D ultrasound recordings of the kinematics of arm movements of twin foetuses have demonstrated prospective control
of action from the fourteenth week of gestation. Not only were self-directed movements of foetuses towards their sensitive eye region slower with longer spans of deceleration compared to movements towards their mouth, movements directed towards their twin – but not movements directed towards the uterine wall – showed equally long or even longer periods of deceleration before contact (Castiello et al., 2010).

These examples show that engaging in prospective control of actions as a living, acting being is closely connected with a basic distinction between self and non-self: infants come to respond differently to events estimated to be related to self-generated actions and events which are not, and infants encountering the different consequences of actions directed towards self, world, or other come to anticipate these consequences and guide their actions appropriately. Thus they effectively enact a pattern of systematic distinctions between self vs. world or other. This basic ‘self–world distinction’ is only one among many instances of the already highly structured organization exhibited by newborns, including a range of discriminatory abilities, sensitivities and preferences, and action tendencies enabling them to selectively engage with the world.

*Neonatal preferences:* Human neonates (and the neonates of some other species too) can be, within minutes of birth if not too stressed by the process or by maternal drugs, absurdly interested in the world around them. They look intently at faces and objects positioned roughly a foot or so away from them where their vision is most clear, and they turn their heads to follow things – especially faces, but also other objects and voices – which have got their attention. Controlled studies have demonstrated a range of sensory preferences: for face-like patterns (Johnson, Posner, & Rothbart, 1991) and especially for faces with eyes directed towards them (Farroni, Csibra, Simion, & Johnson, 2002), for sounds in the human voice range, for female over male voices, and for sweet (milky) odours over others. For objects positioned within what is called their ‘reach space’, they tend to swipe out with their arms, with evident tension and orientation towards the object in their entire body (see von Hofsten, 1984). In this responsiveness the neonate both reveals embodied structures — including a basic self–world
distinction and selective interest in specific aspects of the world — and an enormous openness for engagement.

5.4 Infant Selves in Early Dialogues

As young infants – whose perception and action is organized in such a way that it enables them to engage with the world – meet an interaction partner – who is orientated towards the infant, eager to interact, and experienced in interaction – infants will, from the very beginning, enter into sustained interactions and communicative exchanges. Whether in neonatal imitation, a little later in protoconversations, in smiling and coyness, in nappy change routines, or in responses to being picked up, they are addressed, respond, and are responded to, each partner taking turns and co-regulating the engagement. These young infants – though they might show interest in and direct actions towards objects – are not very capable of effectively acting on the object world. Yet, in a social context – when they are attended to, addressed, and invited to act – even small or ‘fuzzy’ actions on the part of the infants may be (affectively) responded to by an (attentive) partner and hence may have a big effect on their experience and their action-perception loops. And, since actions successful in creating an interchange are likely to be repeated, this in turn may spur a process of jointly creating action patterns and structures.

Although neonatal imitation is still a hotly debated phenomenon, many studies have shown that neonates respond to gestures such as tongue protrusion and mouth opening by an adult model by doing the same themselves. Exploring this phenomenon with your own newborn can be an exciting thing to do (as tried by one of us). Sometimes, if the tongue protruding engagement continues over the first days, the exchange can become a startlingly clear ‘game’ or strong interactive structure. Newborns not only respond to these actions by the adult, but may also try to elicit them, initiating a round of actions themselves if the adult does not do it first (Meltzoff & Moore, 1994; Nagy & Molnar, 2004). The openness to engaging can thus lead very quickly to a structure of engagement and a desire for re-experiencing that structure even shortly after birth.
Proto-conversations: Early interactions not only involve infant responses with actions similar to their partner’s but also complementary actions. In proto-conversations (from around the second month) infant and caregiver engage in preverbal dialogue through rhythmical vocalizations and movements co-regulating affect and arousal and taking turns in co-producing complex phrase patterns (Bateson, 1975; Snow, 1977; Trevarthen, 1977). They also jointly perform (culturally shaped) everyday routines (sometimes involving objects), with the infants actively participating, anticipating, and responding to actions directed to them including being picked up, having their nappy changed, playing peek-a-boo, or book-sharing.

What all these interactions have in common is that infant and partner are intimately engaged in second-person I–You relations, where the partners directly address each other with their bodies and multiple modalities (gaze, vocalization, bodily movement, sometimes touch) and co-regulate each other’s affect, arousal, and actions. Violations of this contingent coupling of address and response – as shown by ‘stillface’ experiments or delayed video communication (Tronick et al., 1978; Cohn & Tronick, 1983) – can be enormously upsetting to 2month-old infants, leading to frowning, closed mouth expressions, and actual crying. In the typical flow of interaction with its cycles of engagement and disengagement, infants and caregivers repeatedly move through and shape affect-imbued action arcs together: starting a specific action, building it up to a climax before bringing it to conclusion in resolution.

Self-other awareness: ‘I–you and we’ experienced in jointly created actions: How do the participants in these kinds of engagements experience self and other? Jointly enacted actions are and feel different from solitary actions. They cannot be experienced alone. They define social context, allowing each participant to distinguish the ‘I’ from the ‘I-with-Other’ (see Stern, 1985), with the participants experiencing ‘being addressed’ and a mutuality of ensuing actions. Being addressed could affect the infant through a wide range of sensory modalities: most prominently, feeling someone’s gaze directed to oneself, accompanied perhaps by a pronounced facial
expression, rhythmical dynamically modulated vocalizations and movements, being touched, acted upon, or being moved. All these modalities and strands of action may contribute to the more global processes of living through waves of arousal and affect and moving through action arcs together: whether going through a wide-ranging fast-paced dynamic flow of excitement and joy in a co-enacted tickling game, or being soothed and calmed faster and more easily with than without the other, or conversely, experiencing potential friction when being pulled through the motions of being dressed. Daniel Stern, focusing mostly on affective co-regulation, talked about ‘I-with-self-regulating-other’, and ‘I-resonating-with-other’ in this context (ibid.). Taking a closer look at these situations might even provide the resources for a more differentiated experience of self– other and the experience of a basic form of ‘we’.

‘Me’: being the object of attention and the target of action: In the infant’s awareness and understanding of being addressed, we may already see the roots of ‘me’-awareness in the sense of feeling oneself as the target or ‘object’ of another’s attention and actions. Infants demonstrate an awareness that ‘they are meant’ as well as an understanding of ‘what to do’ by responding appropriately when being addressed, whether through vocal turn-taking in proto-conversations or through anticipatory adjustments in everyday activities such as nappy changing or being picked up (Reddy, Markova, et al., 2013; Rossmanith, Reichelt, Costall, López, & Reddy, in preparation see chapter 4; Rączaszek-Leonardi et al., 2013). When approached by their mother with her arms outstretched, the majority of 2- and 3-month-olds, in anticipation of being picked up, start adjusting their bodies in ways that enhance a smoother and less risky pick-up, stretching out and stiffening their legs or tucking them up tight, raising or opening out their arms, turning their heads sideways or raising their chins, while intently watching the mother’s face (Reddy, Markova, et al., 2013). In a simple self-directed context they have learned about others’ actions and intentions, and have developed structured responses appropriate to the impending physical change. Perhaps the most striking manifestation of a ‘me’-self-awareness and indeed of affective self-consciousness is when 2–3-
month-old infants who have just started to respond to others with a smile in social communication (Wolff, 1987) suddenly turn away from their communication partner with a coy smile. Like adults, who, when given a sudden or intimate compliment, briefly avert gaze or raise a hand to cover their smiles, infants too, when greeted or addressed smilingly by another, may smile intensely and avert gaze or head briefly, and may even raise an arm in a whole body coy reaction, revealing not only pleasure at the other’s greeting but a sense of being overwhelmed by the emotion (Reddy, 2000). Such emotional reactions to direct gaze not only reveal (more) complex emotional structures in the infant, but suggest an early form of self-conscious affectivity arising within a relation in which they are aware of the other’s awareness of them.

These cases suggest that a basic form of the ‘me’-aspect of the self (Mead, 1934) is constituted here by multiple aspects coming together: first, emphasizing the infant’s body, there is a change of affective experience involving both proprio- and intero-ception: the ‘me’ as affected by the other; second, emphasizing the other, there is the experience of the other’s actions via exteroception: the ‘me’ as being the object of the other’s attention and actions. However, in these infant–caregiver dialogues the ‘me’-aspect of self does not become completely objectified, the interactions do not move to an extreme ‘I– it-pole’ of relating (Buber, 1923/1983), where the ‘me’ is merely the object of a relation but does not take part in regulating that relation, nor in shaping this ‘me’-structure.19

19 That might, for example, be the case when the actions of one interaction partner are strongly constrained and his or her autonomy is significantly reduced resulting in a merely instrumental way of relating (compare De Jaegher, Di Paolo, & Gallagher, 2010). Or when the ‘me’ – as in older children and adults – becomes (to some degree) identified with ideal entities or loci within a larger system of ideas (e.g. a ‘role’) and is implicitly or explicitly judged against a larger framework of values or norms. Not only is this ‘me’ objectified in the sense of being looked at from some distance as an ‘entity’ in relation to other entities, but also partially eludes the reach of
'I': enacting agency - being effective and being responded to: On the contrary, the dialogical I–You relation which infant and caregiver are engaged in is characterized by an interplay of being affected by and affecting another, where the 'me' who is looked at is also acknowledged as a 'you' (rather than merely being treated as an object) by the other. Invited and given space to act from early on (in proto-conversations as well as in joint object routines) the infant can experience the 'self' as agency, as an 'I', enacting itself into the available action structures. Also, when putting forth its own actions, the 'I' can experience itself as having an effect, being responded to and thus being acknowledged. This occurs particularly through affect- and action-attunement, where the mother responds to the infant’s action by mirroring back an aspect of the action with cross-modal variation (e.g. verbally commenting on an infant’s manual action, matching it in rhythm and level of arousal) thus carving out and reifying the action as well as acknowledging the actor in its current state of arousal and affect (Stern, 1985).

Structure, openness, and self in early dialogue: In these early infant–caregiver dialogues, processes of self-other awareness arise in an interplay between jointly enacted action structures and openness to the other and to change. Co-created structures provide a framework for actions in which and as which a ‘self’ may experience itself. With the help of the adult and through the ensuing interaction dynamics and repetitions, stable patterns emerge providing a medium for stable and consistent experience of ‘self’ and ‘other’. Conversely, in being open to and acknowledging the other’s presence, in responding, and letting oneself be affected and led by the other, the equally vital role of openness becomes manifest for interactions to continue and ‘self’ and ‘other’ to be experienced.

its owner as norms and value systems are typically created collectively and exceed the control of an individual.
5.5 Selves Enacted through Joint Participation in Shared Cultural Routines

Infants and caregivers jointly practise cultural routines from early on, which are not entirely dyadic (solely revolving around the interaction partners and their affective exchanges) but where ‘something else’, a ‘third thing’, plays a major role: an object, a specific cultural form of action, a specific goal to be reached, or a societal rule to be followed. Here the participants – though staying in contact with each other – are typically not completely (but are at least partially) focused on each other, with some modalities decoupled and involved with something else; the dyad is opening up to the world.

Joint structuring of actions involving objects: From early on infants actively participate in joint cultural practices involving objects, e.g. looking at a book together from 3 months on, ‘helpfully’ lifting their bottom in nappy changing routines at 3 months, or covering and recovering their face with a blanket in peek-a-boo even at 4 months. Caregivers facilitate this participation by structuring the interaction in various ways: multimodally carving out objects to engage with, e.g. dynamically pointing up and down the picture of an elephant while rhythmically voicing ‘e-le-phant’, or creating multimodal invariances which stand out as a ‘unit’ from a background and thus parse the scene. And they rhythmically pattern actions, ostensively shaping them into exciting action arcs with a beginning, build-up, climax, and resolution. In this way they not only co-regulate the infant’s arousal, for example by placing relevant action events or objects at prominent points in the arc and marking them verbally, but also create a practical and intuitive action framework helping the infant to orient and participate. Repeatedly moving together through these affect-imbued action arcs provides the opportunity for infants to experience themselves acting into a space opened up by the caregiver, and, in relation with the other, enacting an activity which is bigger than that which could be accomplished alone (Rossmanith et al., in preparation).

At 3–4 months infants are mostly alert and responsive, and — with their attention drawn to relevant objects and events by the local cues provided by
the caregiver — follow their caregiver’s lead through the activity, even if with slight delay. In some frequently practised and clearly structured routines, infants already show great fluency at 3 or 4 months. Over the next months, interaction with their caregivers becomes more finely attuned and subtle, some of the routines are ritualized: that is, a partially performed or subtly indicated action becomes sufficient for the well-practised partner to respond appropriately. In the nappy change interactions of a well attuned mother–infant dyad at 6 months, for example, a ‘Ta!’ (i.e. thank you) — which in previous months was used by the mother to thank the infant at the end of the interaction — now could be deliberately used at any time as a cue to prompt the infant to lift her bottom.

Directives and societal rules — being drawn into explicit cultural structures: From around 6 or 7 months of age parents (differently in different societies) increasingly use directives as they draw infants into cultural practices and actions, by setting up playful as well as serious and repeated routines of engagement around them. They issue many directives to infants everyday — to look at something, to wave, to clap, to come to them, to go to someone, to fold their hands in prayer before an idol, to roll a ball to them, to give, to take, to not spit, to wait, to stand, to sit, and so on — before infants can quite understand the words involved, and they repeat these directives, turning them into routines. Tentative infant compliance with these routines enables the construction of structures at higher levels and the practice of participation itself (which is a different level of structure, a metastructural practice) (Reddy, Liebal, et al., 2013).

Self-other awareness between structure and openness in cultural routines: As a first repertoire of stable shared action routines is established, the infant may experience itself in and across specific routines as a co-participant performing specific actions with another person’s specific action, the whole embedded within a larger framework. The I-aspect might now arise more pronounced and determined within these routines and infants start initiating social actions on objects. The me-aspect might be felt more defined as being the person of whom specific actions in specific contexts are expected. The possibilities for and instances of openness to engage become
more differentiated as well as more specific: one partner can invite and wait for the other to participate with specific actions at a specific point in time. The other can be willing and choose to participate at a specific time with a specific action — or not.

These kinds of interactions and the ways in which self and other might arise in them might indeed form a bridge between self–other awareness in dyadic interactions (with its potential I–You and We seeds), and full-blown triadic interactions (with more ‘objectified’ forms of self, where the self can see itself as seen by others, as an ‘entity’ in relation to other ‘entities’). As interactions get known and established as stable routines — hence ‘reified’ — the ‘I–You and We’ from dyadic interactions find themselves not only in relation to each other but in relation to ‘something else’ (to specific bigger action structures), which over time get elaborated into a highly structured shared action space, within which both participants can direct one another’s attention and actions.

5.6 Jointly Modifying and Negotiating Shared Routines

Social games: Once an interaction pattern is established it can also be modified, re-combined, and played with, as is evident in social games from 3–4 months. In one example at 7 months, a playing mother– infant pair quickly move from ‘blowing-belly-button’ to ‘peek-a-boo’ to ‘smelly feet’ with the mother briefly pausing at suspense and decision points from where she could move into a different game depending on the infant’s response (Rossmanith et al., in preparation).

Clowning: Also from this age, around the middle of the first year infants play a much more obvious role in setting up new interactive structures and violating structures that they have just mastered or accepted. Sensitive to others’ emotional reactions to things, from around 7 months infants start to play with these reactions — often performing extreme and absurd actions — in order to evoke and elicit laughter and amusement in others. Such clowning might involve shaking the head repeatedly, making funny sounds, odd facial expressions, doing absurd things with objects, or fake coughing. The actions themselves rely clearly on the relation with the other — if they
amuse they survive, if they receive no response they die out. These relational structures confirm not only the infant in a new role as clown, but also the relationship itself as one constituted by amusement.

**Teasing:** From around 9 months of age infants also start to provoke. They playfully violate established understandings and gestures or newly learned ‘rules’. Having just learned to give objects in response to open palm requests, they offer and cheekily withdraw the object before it is taken. Having grasped that the plug socket is a no-go area, they reach out a hand and almost touch it, watching for reactions. They tease others by provocatively playing with newly established structures. In teasing, infants use structure (as do adults in even more complex ways) to take the relationship further; the violations cause surprise, alarm, amusement, and then a denouement, a coming together into a deeper level of resolution and intimacy. A new openness has been enabled between self and other (Reddy, 1991; Reddy & Mireault, 2015).

**Self–other awareness between structure and openness:** Here against the background of already established shared action structures, ‘self’ and ‘other’ can be felt acutely in varying, exaggerating, violating these structures. Deviating and breaking out of established patterns as well as profoundly affecting the other as a consequence may contribute to a new form of experiencing agency and ‘self’ as the author of specific, original actions (the ‘I’-aspect of self). At the same time being seen performing specific actions and being acknowledged and admired for particular ones affirms the ‘me’-aspect of self. On the receiving end of such variations in action structures, when one’s expectations are playfully violated, the sense of ‘me’ may deepen as one feels oneself addressed with intention, with a deliberate effort to affect ‘me’ building on a history of such engagements. Thus, what is implicit in these playful variations is ‘I affect you in a way in which I am acknowledging our relationship’.

These variations may contribute to further develop a sense of ‘we-awareness’ not only by adding to a unique shared history but also by inviting the participants — rather than to merely know and follow the steps of the routine — to pay close attention to the other’s (unexpected) actions.
and to diligently coordinate one's own actions with them, thus enhancing mutual attunement and fostering cooperation. Across the repertoire of established action routines and their respective variations, a particular interaction structure comes to stand out even clearer and becomes increasingly graspable as an ‘object’ which can be jointly related to, modified, and negotiated about.

5.7 Opening a Conversation with Spiritual Traditions

So what can we learn from infants’ engagements in their everyday worlds about the self? And how might this connect to Buddhist and other spiritual traditions?

We identify here several points of linkage which may serve as the start of a conversation between these different traditions.

First, very young infants do not have a concept of self and lack the linguistic and narrative wherewithal which for adults is a major source for maintaining a self as a reified, separate, and enduring entity. In Buddhist dharma, this reified self is pointed out as illusory and a source of suffering. However, contrary to earlier accounts about the lack of self–non-self differentiation at birth, numerous studies now suggest that even before birth, and certainly immediately after birth, infants distinguish between consequences of actions generated by self and other, respectively, and guide their actions appropriately with respect to different targets (self, world, or other). This set of basic self–non-self distinctions does not map onto adult conceptual and linguistic distinctions between self, other, and world. Rather, it can be seen as a basic pattern enacted by living beings, in the course of striving to sustain themselves – approaching pleasure and avoiding pain – relating to and participating in a world of things and others. This resonates with the Buddhist notion of co-dependent arising or origination of self and world (compare Varela et al., 1991), and may be regarded as the beginning of a gradual process of self-reification, coming in different forms and degrees.

Second, the interpersonal self emerges within interpersonal engagements, with each participant feeling the other through their own
actions and emotional responses. In an infant’s coy smile to an adult greeting, for example, the expression of self–other awareness involves a dynamic changing shape (intensifying, turning away, returning), regulating the interaction and leaving both partners affected and changed. In such delicate encounters both ‘self’ and ‘other’ are structures, at times more pronounced and solid, at times more fleeting and permeable, with different boundaries or facets in different relations. This corresponds to images of the self as an ‘eddy in the social stream’ (Mead, 1934), drawing from James’s ‘stream of consciousness’ (James, 1892), which are strikingly similar to the Buddhist simile of the self as a wave in the ocean; each highlights the dynamic, fleeting character of a self embodied in a larger action or process that cannot be understood as an isolated entity.

Third, we suggest that the self develops within and through jointly created structures:

First, in very early dyadic interactions, interaction histories provide patterns or structures which serve as a medium within which engagements can occur and self and other can be consistently experienced in direct contact. The powerful effects of patterns of contingent relating between self and other can be seen both in infant coy smiles to greetings and in distress to its violation. This is shown in the still-face studies (Cohn & Tronick, 1983) where infant distress is very evident (and can affect the partner too), and in the studies of infants of depressed mothers (T. Field et al., 1988) who had adapted to the different response patterns of their mothers and subsequently affected new interaction partners with their adaptation.

Second, structures can also act as containers within which infants can experience agency in particular contexts. For example, jointly performed routines such as nappy changing involve a series of steps — lying on the changing mat, feeling the nappy open, hearing the mother speak, feeling the nappy removed, etc., with the infant herself performing some actions — such as lifting the bottom at the right moment. The infant’s participation is enabled by the structure of the routine. These repeated experiences, allowing action parsing and recognition of specific contingencies, may lead
to an ‘objectified’ understanding of self as a **co-participant** in complex actions.

With the establishment of these kinds of interactions, the basic processes that continually shape us as ‘**personal and cultural selves**’ are already in place in the first year of life. To the extent that we are what we do, our skills, habits, daily routines, interaction dynamics all constitute a ‘**procedural self**’, a circumscribed space of actions and lived experience. Interpersonal and cultural interactions, customs and (implicit) norms sanction — encourage or impede — our actions, defining and delimiting this **self-in-social-context**.

Third, structures can also be the **stable background against which infants can experience themselves in new ways**. They can now be ‘**owners**’ of actions as in ‘clowning’ or ‘showing-off’, showing pride at being acknowledged and recognized by others’ responses to the actions, or shame at the absence of such recognition. They can be ‘**authors**’ of new actions, varying and deliberately deviating from established routines as in teasing. These interactions may contribute to a **solidifying sense of self extended over longer periods of time** which owns and performs a multitude of actions.

Fourth, as **structures (joint action patterns)** become increasingly fluent and familiar they also become **increasingly reified** and thus **available as objects** which can be jointly attended to, modified, or negotiated. Early forms of such reification are already evident at around 6 months, e.g. in book sharing, when infants and caregiver engage in affective exchanges as ‘comments’ on the ongoing activity. Joint relating to such action routines as ‘objects’ becomes more frequent over the next months as caregivers and infants jointly modify their games, particularly clear from around 9 months in deliberate teasing.

This way of jointly relating to previously established action patterns as objects of shared concern could be argued to be a basic form of **reflection**, a **reflective movement directing attention (and actions) to the jointly created action**. By enhancing the understanding of the action structure as well as the role of the self, this may over time lead to a **more objectified**
sense of self, as being seen from a distance, from another’s perspective, as one ‘entity’ in relation to other ‘entities’.

Re-flection or reflexivity is perhaps the most characteristic, or at least the most frequently discussed, aspect of ‘self’. Mead defined the self as ‘that which can be an object to itself’ (Mead, 1934, p. 140). Adyashanti describes the self as a reflective movement of consciousness looking at its own experiences (which e.g. might mean looking at the ‘ego’ which itself is founded in reflection) (Adyashanti, this issue). The reflexive aspect of self-awareness may have different forms and levels (Almaas, this issue). Inherent or pre-reflective reflexivity may be distinguished from explicit reflection and finally the movement of reflection alone. Inherent reflexivity of consciousness is the awareness of being aware which accompanies every experience. Explicit reflection may be seen as the reflection that occurs through rational thought and language, and can be placed as occurring in the middle of the second year with the emergence of a concept of self (Lewis, 1995). Finally, in the context of spiritual practice, when awareness is deliberately directed onto itself, the self may be experienced as this reflecting awareness witnessing experiences as they arise. Adyashanti (this issue) describes a point where there is the movement of reflection but nothing to reflect upon, no conflict, distinction, division (and hence experience) so all that remains is awareness aware of itself and the process of reflection and hence the self ceases altogether.

Returning to self-development in infancy, even the basic feedback loops of prospective actions could already be conceived of as a basic form of ‘re-flection’. In early dialogues, specific actions are reflected back to the infant by the adult via affect attunement. When infants and caregivers finally start jointly relating to – that is reflecting on – shared actions, this may contribute to more complex ways of creating new structures together, including imitation and symbolic activities such as labelling. As infants over the next years come to participate and live in increasingly complex networks of structure and structure creation, they increasingly treat objects in conventional ways – but also engage in symbol play – and at around 3 years they participate in setting up explicit rules, and stick to and
feel bound by established rules (Rakoczy et al., 2008). All this again goes hand in hand with (ever) more complex ways of shaping and negotiating the self, in particular the 'conceptual' and 'narrative' self (Gallagher & Hutto, 2008). Thus, for developmental psychology, the development of a conceptual self is seen as an important milestone on the road to a fully functioning self – while Buddhist dharma cautions us about self-reification as an illusion and source of suffering. Exploring how these perspectives relate, we enquire into how the self and indeed jointly created action structures both inherently enhance engagement as well as lead to separation.

5.8 Structures Enhancing and Hindering Engagement

Interaction patterns established as shared action routines provide opportunities for further engagement on a cultural as well as an interpersonal level. At the interpersonal level the establishment of shared action patterns by repeated performance of the same actions between interaction partners allows the patterns to become familiar and hence predictable. This not only fosters a feeling of security but also allows more active, meaningful (and anticipatory) participation by the infant, thus serving as a ‘container’ for feeling self and other in engagement, joined in a larger activity. As the component actions become established and reified, they can serve as ‘objects’ for further reference, modification, and expansion, that is, further engagement. Thus, structure seems to open the door for the emergence of further structures to engage with. In the interpersonal situation, these emerging and enlarging structures contribute to a unique history of the particular relationship, increasing intimacy and what might be called a ‘we-ness’ between them. The joint shaping of what might be seen as ‘an action landscape’ inevitably excludes other paths and channels, with other potentials remaining unrealized. But this selectivity too maintains the structures and identity of the relationship.

These consequences also hold at the cultural level. Neither caregivers nor infants are isolated agents, but are always participants in and co-creators of sociocultural niches. As illustrated in previous sections, infants
participate in cultural practices from early on and are trained in the implicit norms and conventions as well as the more explicit societal rules of the respective social group. Practices and conventions range from differentiation of language-specific phonemes and prosody (Werker & Tees, 1984), as well as culture-specific ways of moving, sitting, and walking (Adolph, Karasik, & Tamis-LeMonda, 2010; Mauss, 1973), to cultural formats such as greetings, multi-party conversations, book sharing, helping, or specific rules of etiquette. For example, infant burps and bowel movements, outside the immediate feeding context (within which they tend to be praised), may be responded to with a culturally appropriate ‘Excuse you!’, which might surprise foreign visitors, but is a common feature in British middle-class households. Familiarity with and skills in these practices allow the infant to interact and participate beyond the interpersonal and to belong to a specific sociocultural group. The infant’s familiarity with the book as an artefact and with the sequence of activities in book sharing allows her to meaningfully interact beyond the close circle of daily interaction partners, with other members of the culture.

On the flip-side of the coin, however, the very same jointly practised and ingrained action patterns which enable joint engagement, intimacy, and belonging also lead to separation from others and the world (and even to separation within the self). The 7- or 8-month old’s anxious withdrawal from the stranger who is unfamiliar with her games and action patterns (compare also Trevarthen, 2004) could be seen as an early example of separation and an incipient ‘us’ and ‘them’ divide. Sticking to established action patterns can also contribute to separations within the self. The caregiver’s interests, inclinations, and adherence to sociocultural norms might lead to some of the infants’ actions being routinely picked up and ‘affirmed’ while others remain unresponded to and dropped. Daniel Stern (1985) talks about ‘selective affect attunement’ as a powerful contributor to the development of personality types or one-sided coping strategies for interacting: think of the case of an infant only consistently attuned to by the caregivers when acting enthusiastically but not when in a sad mood. Learning that one of these states is shareable while the other is
not, she will likely live and relate to these respective emotions differently (perhaps ignoring or denying one altogether), thus creating a line of separation within her life, experience, and self.

5.9 Structure and Openness

The self along with other structures enhances engagement and at the same time hinders it, creating both separation and connection, stability and change. Even within each interaction, there is openness in the form of variability. The sources of this variability are multiple: they lie in the separate agency and spontaneity of the participants, in the encounter between present actions with previously established structures, and in the ability of the participants to be affected and to change in the present encounter between them. Each self meets the other as both familiar and surprising or transcendent, as a particular, bounded self as well as in ways in which these boundaries recede. This dialectic of knowing and not knowing the other, and knowing and not knowing what is going to happen in the encounter, is common to participants. Within this shared unknowing lies connection (see also Buber, 1923/1983). And within this connection lies both form and emptiness.

5.10 Acknowledgments

Parts of the research contributing to this paper were supported by the Marie-Curie ITN ‘TESIS: Toward an Embodied Science of InterSubjectivity’ (FP7-PEOPLE-2010-ITN, 264828).
6 The development of participation in infant-caregiver-object practices: changes in attention-and-action coordination

In this chapter, building on conceptual analysis and the results from our empirical analysis from chapters 3 and 4, we seek to reframe the conception of participation, joint action, and co-operation, usually conceived of in cognitive and rational terms, to a more embodied and situated version. We then explore how such a framework illustrated with empirical examples can be used to describe and account for development of participation.

6.1 Setting the stage

6.1.1 Challenges of joint action, participation, and cooperation

How do joint acting, participation, and co-operation in joint object activities develop over the first year of life? What does it take for an infant to meaningfully engage in these forms of interaction? Most definitions of participation, co-operation and joint action typically involve jointly working towards a shared goal (Bratman, 1992; Hubley, 1983; Rohlfing, Wrede, Vollmer, & Oudeyer, 2016; Sebanz et al., 2006; Tomasello & Carpenter, 2007; but see Fantasia, De Jaegher, & Fasulo, 2014), which constitutes a third pole of interaction to relate to. This makes joint actions – even in the absence of objects – equivalent to triadic interactions, requiring the infant to simultaneously take into account and coordinate engagement between the partner and the shared goal. Following through with this frame, in order to properly participate the infant must have mastered triadic interaction. Thus the developmental challenges which need to be mastered in order to participate are the same ones which are typically considered crucial for triadic interaction: coordinating engagement between different nodes of the triad – as in (visual) joint attention, operating specific kinds of complex knowledge – i.e. being able to parse actions and understand their instrumentality (distinguishing between means and ends), as well as to understand other people as intentional agents (having specific goals, perceptions, selective attention and action strategies, see Trevarthen &
Hubley 1978, Tomasello, 2005). Thus proper participation, cooperation, and joint action can only be expected from 9 months on.

6.1.2 Framing and analysing the challenges from an embodied and situated cognitive science perspective

However, adopting an embodied and situated perspective, we suggest a shift in framing joint acting, participation, and cooperation. From this perspective, the main challenges infants are facing present themselves as: 1) Controlling their actions, 2) contributing them at the right time in an appropriate way in a 3) complex task, which they may need to understand, and, in addition, 4) coordinating their actions with a partner and his/her own actions.

Drawing from data from our naturalistic study (see part II) we will suggest a new framework briefly sketching the development of participation along the following lines:

1) What were the modes and strands of engagement infants utilized? How did infants within a specific mode of engagement co-ordinate engagement and disengagement with one or multiple targets in the course of an unfolding activity? As modes of engagement we here differentiate – based on previous observation and analysis – a) attentive-communicative engagement, used primarily for (distal) orienting, control and monitoring of actions, as well as for communicating, involving gaze, vocalizations, facial expressions, and to some extent whole-body movements; and b) physically-effective engagement, used to bring about changes in the material world, involving body and limb movements, up to fine manual manipulation, whether concerning objects and/or people.

2) How did infants co-ordinate the different modes and strands of engagement with each other as an activity was unfolding?

3) To what extent and how did infants participate, and in particular functionally contribute, to the joint activity?
Figure 6.1: The development of participation. Details see text.
A tentative trajectory of the development of participation in terms of coordination of attention and action over the first year of life

6.2 3-4 months: infants participate in co-ordinated affect imbued arc-shaped social object routines.

6.2.1 Responsive co-ordination of attentive-communicative actions

At 3-4 months infants became increasingly interested in objects in their surroundings and thus, when in an alert and contented state, also attended to aspects of the object activities which they were directly involved in as well as those which they witnessed other people perform. Their co-ordination of attentive-communicative actions was responsive, that is, they tracked salient object actions with their eyes, and had their gaze drawn (with slight delay) to distinctively marked events such as excited vocal comments and exclamations of surprise, and to especially prominent multimodal events such as audio-visual contact events created when objects made contact with the ground or each other as the caregiver handled them.

Consequently, infants in this period were also easily distracted, frequently having their attention – while being engaged with one target or activity – abruptly drawn to something else. In one example (see Figure 6.1.c1.b), the infant was lying in the baby gym focussing on and trying to touch and grasp objects hanging at middle height right in front of the infant’s face above his chest. The infant’s gaze, and the infant’s engagement in general, was drawn to and fro between the swinging object and the mother’s affect-attuned cheering face, temporarily disrupting his efforts at autonomous object engagement. Sometimes, however, the infant became less distracted even within a single episode; as indeed in this example the infant looked less and less at his mother who, on her part, realizing the disruptive effect, reduced the intensity of her cheers.

For most joint object activities, however, (both for guided ones as well as for witnessing situations) the infants’ responsiveness made them amenable to and inclined to follow the caregiver’s lead. The latter, by multi-modally marking particular events, carving out relevant aspects or building blocks of actions,
patterning actions, as well as shaping them into affect-imbued, arousing action arcs, guided infants' attention through the activity and towards its relevant aspects and shape.

Generally, as a result of their responsiveness in this period, infants at this early age tended to end up looking at “where the action was” in their surroundings. They looked at crucial key events of an on-going activity (whether they were specifically created for them by caregivers (figure 6.1.c1.a), or occurred as an inherent part of the interactions unfolding around them (figure 6.1.c1.c); that is, at points in space and time where potentially important information could be picked up, providing plenty of opportunity to learn about objects, people, and action structures.

Witnessing her mother emptying the dishwasher, the girl in figure 6.1.c1 tracks her mother's hand reaching for and grasping plates in the dishwasher and transporting them to the overhead cupboard. After a few iterations the infant’s attention gets drawn to her own hands, which have found and are now clasping each other. The clattering noise of the plate making contact with the dishes in the cupboard, however, draws her gaze back to the plate, and – after her gaze dwells at the location of the noise a little longer - she continues monitoring her mother transporting plates with her hands.

6.2.2 Co-ordination of whole body and manual actions

6.2.2.1 Scaffolded manual physically effective object acts

As the infants at 3-4 months were not able to support their posture and properly perform visually guided reaching and grasping on their own yet, their manual object actions were still limited.

They did, however, perform – and in particular answer action invitations with – simple manual object acts scaffolded by artefacts and/or caregivers in a co-ordinated way, such as well-aimed hitting of objects suspended from a toy bar after the mother had set it in motion by tilting the baby-seat (see figure 6.1.c1.e). Often these early object acts even fit or were moulded into more conventional cultural formats: autonomously holding their drinking bottle (see figure 4.5.r1.d), or turning the page of a book held steady by the mother, whose rigid pages facilitate grasping and channel the infants into performing
the conventionally “correct” action by reducing the degrees of freedom (see figure 6.1,c1d, as well as book sharing: section 3.3.1.3).

6.2.2.2 Early autonomous physically effective manual object engagement frequently diverging from attentive-communicative engagement

When acting on objects autonomously (e.g. scaffolded by a baby gym or handling an easily graspable and manipulable object), infants’ physically-effective manual engagement and their attentive-communicative engagement did not necessarily converge on the same target. Indeed, on the contrary, they frequently diverged, or were sequentially directed at the same object. For example, manual object acts were often not performed under sustained focal visual control. This was probably - at least in part - due to postural challenges. When lying in a baby-gym, infants typically (as illustrated in figure 6.1.c1.e) – after first sequentially alternating their gaze between the objects suspended high above them and performing a series of whole body movements, waving hands and feet in a diagonally coupled way – turned their head to the side or centred it orienting towards a point between the suspended objects, before directing a smooth swiping movement unilaterally towards one or bilaterally towards two objects, often successfully hitting them with their fist or back of the hand. When, however, their posture was supported, e.g. when reaching for a low hanging object with their arm remaining supported by the floor, or sitting reclined in a baby rocker with their back supported and hitting an object suspending from the toy bar, which the mother has made prominent by setting it in motion through tilting the baby rocker (see figure 6.1.c1.e), infants typically kept their visual focus on the object.

Similarly, when engaged in grasping and sustained manipulation or mouthing a proximal readily and constantly available object, they often were simultaneously attentively-communicatively involved with their caregivers, looking at them as well as communicating with them (via facial expression and vocalisation), again performing the manual actions largely without focal visual control.
6.2.2.3 Anticipatory (and co-operative?) whole body movements functionally contributing to a shared task

While 1) infants’ manual skills – in particular unilateral reaching and grasping movements especially under challenging postural conditions – were still limited but in this period undergoing intense practice (with artefact and caregiver support), and 2) their manual object action and visual attention often diverged, infants at this age have already been performing whole body movements routinely for several weeks, communicatively engaging with their caregivers. By now, they were even able to perform whole body movements within a specific caregiving task such as nappy change at the appropriate moment in an anticipatory way facilitated by object action cues, such as lifting their bottom in nappy change at the sight and sound of the nappy being opened, and the mothers’ “Are you ready?” vocalizations, thereby actively contributing to the task, allowing the mother to perform the next step in the routine to remove the nappy and thus arguably engaging in the first rudimentary forms of object related co-operation (see figure 6.1.c1.f, as well as section 4, in particular figure 4.6)

6.2.3 Summary: infants’ participation at 3-4 months

Even at the early age of three to four months, when their “autonomous” object engagement was still very limited, required scaffolding, and was often performed with diverging attentive-communicative engagement, infants were able to participate in joint object routines by following the caregivers’ lead (thanks to their responsiveness) and letting their scaffolded object actions be moulded into cultural forms. Further, they were already able to participate in an anticipatory and arguably cooperative way. This cooperation was possible as long as they could contribute with an action already in their repertoires such as whole body movements, which they performed while staying engaged with the caregiver throughout the activity, timing and adapting their actions in response to and in anticipation of the caregivers’ object actions. And as long as the actions needed were not as challenging as uni-manually translating an object in relation to other objects
under focal vision control as typically used in experimental cooperation tasks (see e.g. Hubley & Trevarthen, 1979; Moro & Rodríguez, 2004)

6.3 4-6 months: a range of social object routines was fluently established, and their shapes, i.e. space-time-intensity contours, playfully varied

In the period between 4-6 months a number of joint object practices became established as shared routines with action structures which were by now familiar to both interaction partners. These included nappy change, peekaboo, book sharing, feeding, social body-part games, soft-toy conversations and shared conventional-object-action games e.g., turning and pulling out a key, or supported opening and looking behind a door.

6.3.1 Pro-active, fluent and co-ordination of attentive-communicative engagement, including affective exchanges

While infants, despite their growing motor skills, were still unable to autonomously move in and explore the world of objects on their own, their co-ordination of attentive-communicative actions had become more fluent (they were now faster in alternating their gaze between caregiver and objects). They were more autonomous (that is less distractable and drawn by immediate cues) e.g., they were not distracted anymore during object engagement in the baby gym by the mother’s cheering and affect attunement vocalizations. They were also often anticipatory, in particular in familiar routines: e.g., after an emotional exchange with the mother during book sharing, the girl in figure 6.1.c2.a returns her gaze to the book on her own, without being dynamically cued, but arguably guided by routine (compare book sharing: section 3.4.1.1).

Generally infants’ more pro-active and controlled co-ordination of attentive-communicative actions now clearly indicates what could be called “knowledge” of the activity structure, but what we would rather seek to characterize more precisely as infants’ now having established a stable
(spatiotemporal) orientation within several specific activity frameworks.

Infants had also become more adept at getting oriented in a situation, as well as at controlled switching between engagements, that is, at disengaging and re-engaging without losing track of an activity and its respective spatiotemporal framework. This held for switching between targets of engagement within a joint activity as well as for switching engagement between activities taking place in parallel, e.g. infants’ own increasing autonomous object engagement and someone else’s object engagement (figure 6.1.c.2.b), or between their autonomous object engagement and attentive-communicative engagement with his mother accompanying him. The boy in figure 6.1.c.2.b is sitting in his baby seat engaged with a mechanical toy. After his gaze had initially been drawn by noise and movement to his mother working at her laptop, as well as to the cup she is grasping, placed some distance away on the bookshelf – he now from time to time looks up from his toy to look at the mother – sometimes drawn by slight movement but more often spontaneously (without noticeable cues) – and then all the way to the cup on the shelf, as if to check whether they were still there or if there had been a change in the situation. His sequential switching and coordination of engaging in his own object activity and visually engaging with the mother working on the laptop and the cup on

20 Karl Bühler, when taking a more general comparative view at living organisms in general, considered even more basic forms of “being oriented”, such as a cat turning to where a noise came from, as particularly important phenomena clearly qualifying for and marking the beginning of the subject area of psychology, as for him such orientation was a hallmark of/marker for basic forms of cognition, for an essentially cognitive way of relating to the world, since they require some basic – and as we today would say “dynamically embodied and situated” – form of systemic “representation”. Interestingly he thus actually defined “cognition” in terms of “being oriented”, or to put it more actively and less statically: as “being able to act, or acting in an oriented way”. (e.g. Bühler 1936)
the shelf, indicates his awareness of and interest in the mother’s current and potential object engagements.

The same boy, lying on the floor on his side in an episode of accompanied object play, is arching his body and stretching his arms overhead and back to reach his toy just out of reach and mostly out of sight. As the mother acknowledges his efforts and cheers him on with an affect and action-attuned “strE::::::tch”, and pulls the toy a little closer to him, he briefly pauses his efforts, relaxes his body, and looks at the mother, before returning to his bent-over-backward arms-over-head stretch, directly aiming for, and finally successfully retrieving, the overhead toy.

6.3.1.1 Playful variation of space-time-intensity contours of action structures

When enacting (some of) the jointly established familiar action routines (e.g. in peekaboo, see figures 6.1.c1f, 6.2), caregivers and infants in this period frequently engaged in and visibly enjoyed variation of the respective action structure, in particular of the shape of the action arc, its space-time-intensity contour, mostly led by the caregivers. For example, by playfully delaying or rushing particular sub-actions and exaggerating actions in scope and intensity, tailored and responsive to the actions and affective expressions of the infant. Thus they were arguably not only playing with the jointly established structure – letting the fixed and the variable parts stand out even clearer – but also directing attention away from the fixed structure toward each other’s varying actions and mutual co-ordination, thus refining co-ordination itself, by making the partners more sensitive to each other's actions and making each other’s actions, intentions, and expectations felt more distinctly.

6.3.2 Interspersed affective exchanges/comments and joint reflective engagement

Another example of the fluent shifts of attention and controlled sequential co-ordination between different targets of attention are the newly observed affective exchanges interspersed within joint object activities with infants now oriented within jointly established routines. In peekaboo, an example of
object mediated social interaction where the partners move the object between their mutual lines of sight to modulate their attentive-communicative engagement, affective exchanges were already observed at 4 months, as the infant, upon her mother bursting forth from behind the blanket with a “boo!” and re-establishing eye-contact, briefly giggles towards her (see below, figure 6.1.c2.d). Conversely, in book sharing - a joint cultural activity primarily oriented towards and focused on the object book - this was only observed from 6 months on, when infants, upon the mother’s particularly emotional verbal enactments of the story in the book, turned their gaze to the mother’s face fast enough for their gaze to meet (also facilitated by infants now sitting more stable at a 90° configuration with shorter gaze-shift distance). This allowed them to exchange and mutually reinforce an expression of surprise, before first the infant – apparently oriented within this jointly established routine – and then the mother returned their gaze to the book (see figure 6.1.c2.a for an example, as well as section 3.4.1.1).

This affective exchange is directly emerging out of a jointly established routine around an object, which the partners, disengaging from the object, briefly put on hold to engage with each other, and soon – not having lost track of the joint object activity – return to. Thus this affective exchange can be seen as a basic form of “comment” “about” notable parts of this jointly established routine or the object involved (compare Bruner’s comment structure, Bruner, 1975; also see book sharing: section 3.5.1.1). More generally, it can be regarded as a basic sequential form of joint reflective engagement with previously established shared action structures, explicitly marking a specific part of these now to some extent reified structures as a “shared focus”, a “topic of shared concern”

Instances of joint reflective engagement highlighting a shared focus, arguably constitute instances of triadic engagement, and play an important role in the development of triadic engagement, as well as the development of joint reference, where the partners orient each other towards, and/or modify
each other’s relation to, a specific location within, a “building block” of, an already established shared action framework (see below).

6.3.3 Co-ordination of physically effective manual object actions

Along with their growing postural control, infants became increasingly more fluent in simple manual object acts, reaching, grasping, mouthing, and manipulating objects. While they still at times manipulated and mouthed an object while simultaneously attentionally-communicatively engaging with the caregiver (e.g., intently observing their mothers) they increasingly manipulated objects under focal visual control and no longer let themselves be easily distracted. Their range of action contributions, in particular manual actions, to joint object routines expanded. These included actions facilitating as well disrupting the joint activity (e.g. opening the straps of the nappy). Infants’ actions generally still let themselves be shaped into a cultural frame.

With slight initial support infants were now able to perform simple body-relative manual object acts, that is, actions in which the object is moved in a specific relation to the body (rather than in relation to another object), in an anticipatory manner within a joint object routine: After having been invited by the mother “What’s this? Do you wanna play Peekaboo?” and offered a blanket right over her chest within reaching distance, the 4 month old girl in figure 6.1.c2.d, pulls the blanket over her face, covering and then uncovering it again, thereby not only effectively enacting attentional-communicative but also the manual steps of peekaboo, until she finally accidentally loses hold of the blanket.

6.3.4 Summary: participation at 4-6 months

With their increasing postural, motor, and attentive and cognitive skills, the infants’ autonomous object engagement became more extensive (though still mostly scaffolded and socially accompanied by their caregivers) and was increasingly performed under focal visual control. That is, there was an increasing convergence between manual physically effective engagement and attentive-communicative engagement, along with a decreasing distractedness (prompted, e.g., by the mother’s regular pronounced affect-and-
action attunement), but an increasing controlled switching of attentive-communicative engagement between different targets within or between activities. These involved sometimes continuing manual-physically effective engagements not requiring visual focal control, such as mouthing, or putting on hold those engagements done under visual control – without completely losing track of the activity.

Infants’ participation also became more extensive and fluent. With their increased skills and with a number of joint object activities now robustly established as shared routines, infants moved within this shared action structures flexibly:

First, infants co-ordinated their attentive-communicative engagement fluently, pro-actively shifting back their gaze to the object at the centre of the joint activity without necessarily being cued or guided.

Second, regarding their physically effective manual engagement, infants were now able to functionally contribute to a (object mediated) joint activity such as peekaboo with anticipatory body-relative object acts, that is translating an object relative to their body (e.g. covering and uncovering the face with the blanket), co-ordinated with and regulating their attentive-communicative engagement with the mother, with the object mediating and modifying their connection. Thus infants gradually increase the amount and complexity of object engagement – and hence the degree of triadicity – they bring to participation. This kind of body-relative object act though may still be considered as relatively simple (at least once the infant has successfully got hold of the blanket with the support of the mother), using one’s own body as a reference system rather than precisely targeting some object-relative goal under focal visual control. And from a Piagetian perspective they could still be regarded as holistic circular reactions, where the object is not necessarily separated out from the action as a unit moved in relation to other units.

Last but not least infants start – together with their caregivers – to jointly relate to, and reflectively engage with, the previously established object routines. They mark parts of the jointly established structures as a currently “shared focus” of negotiation: mother and infant are now engaged in interspersed affective exchanges, which – originating from and leading back to a
shared object routine – can be considered as **basic forms of “commenting on” or “communicating about” the activity**. And they also enjoy modifying the space-time-intensity contours of these structures in joint play.

Thus infants were engaging in a **basic form of reflective joint negotiation** and their gradual development reached qualitatively new levels of triadic engagement and participation as the participants now gradually started to **effectively mark and acknowledge joint action structures as a third pole in the activity** and **actively negotiate them**.

### 6.4 Intermission: Initiating and re-elicitating established routines from 6 months on

From around **6 months** on infants started to **initiate** and **re-elicit** established social (object –mediated) routines, (as of yet) even with strangers, and at the same time we observed the first forms of **instrumental actions** and **social instrumental actions**.

**Initiating social object-mediated routines:** While the mother had turned away to fetch the infant's food, the by now 6 month old girl in figure 6.1.i1.a, who had already actively participated in mother-initiated peekaboo from 4 months on, looked at and made eye contact with me, smiling; she pulled the blanket over her head, pulled it down again and looked at me expectantly. I could not resist chiming in with “there she is again” albeit in a low voice, still somewhat reluctant to get involved while filming. She pulled the blanket over her head again and we went through two more rounds of smilingly covering “where/now she(‘s) gone? “ – and uncovering “THERE she is again!”.

**Re-elicting social object(body-part)-mediated routines:** After we had just finished playing a feet-clapping game in the getting-to-know-each-other-again phase before proper filming, the 7 month old girl in figure 6.1.i1.b looked up to my face, raising her arms towards me, upon which I repeated the game. After finishing, she again looked up to my face, performed another arm raise – this time in reduced (already ritualized?) form – and looked back at her feet even before I started moving to grasp her feet and repeat the game once more.
6.4.1 Summary: initiating joint object mediated routines from 6 months on.

Infants’ actions in these examples again demonstrate that a shared object mediated routine has indeed been established to some degree, either robustly, having been established over the course of many weeks, or loosely, within a few minutes over just a few iterations of a jointly performed body-part game. They demonstrate that infants are oriented in these structures, moving within them, or at least relating to them, quite flexibly and proactively, less bound by or dependent on their immediate context to act: while the blanket is still around as an action opportunity and affordance, the infant does not depend anymore on the mother’s first move to activate, initiate, and enact the peekaboo routine, even with a stranger. Thus, initiating specific joint activities themselves, infants come to \textbf{actively direct their partner’s attention and action via or towards an object} and \textbf{gradually expand their abilities} (having started early on with turn-taking in proto-conversation) \textbf{to participate by enacting increasingly differentiated reversible roles}.

6.5 6-9 months: combinatorial variation of (object mediated) social play routines versus solitary object play

During this period a \textbf{pronounced split} became increasingly apparent between object-mediated social games and solitary manual object engagement.

6.5.1 (Object mediated) social games

On the one hand infants enjoyed engaging in extended exuberant social play interactions (at times mediated by objects), centred around and fostering mutual co-ordination, (by) playing with each others’ expectations and intentions. This not only involved time and intensity variation but also \textbf{combinatorial switching between different subroutines}, where – while the mother would take the primary lead – the infant showed swift anticipatory and flexible co-ordination of affective-communicative as well as whole-body and manual object actions. Infants not only showed anticipatory responses and laughter to temporarily varied actions, but also laughter and quick adaptation to unexpected switches to the move of a different game by, on her part, varying the response. These varying moves between mother and infant can be considered as an \textbf{increasingly complex reflective joint engagement with already jointly
established action structures. They took the form of affective comments on jointly created results in the form of laughter and vocalisations, as well as of interventions and modifications in the form of varying action responses to modify the jointly established routine.

In the over 7-minute play sequence in figure 6.1.c3.a, the mother – kneeling in front of her daughter in a ready-to-go position, continuously keeping eye-contact – spontaneously moved into one or the other game inspired by the momentary situation and actions of the infant: as the infant removing her pacifier incidentally produces a loud plopping noise, the mother, imitating it, starts a wiggly-approach lip-smacking game, met by her daughter’s laughter as she delays and varies each new approach, (even) making ritualized intent movements and false starts, waiting for her daughter’s anticipatory laughter before approaching. As the infant’s belly incidentally gets uncovered she switches into a belly kissing/blowing game re-initiated and reinforced by the infant pulling back her dress even more. As the infant pulls up her dress so high that it nearly blocks the line of sight, the mother hides her head behind the dress before swiftly rising up with a loud “boo!”, seamlessly switching into the peek-a-boo game.

6.5.2 Solitary manual object engagement and focal visual control – mostly ignoring people

On the other hand, however, infants engaged less often for shorter periods of time, and instead frequently sought to escape previously enjoyed cultural object practices such as book sharing. Rather, facilitated by the now stable sitting posture, infants got deeply involved with objects, with their attention drawn to and largely bound by their own object acts. That is their attentive-communicative and physically-effective manual engagements were convergent and tightly coupled with the object: e.g., banging, mouthing and manipulating one or more objects. They were performing counter-resistance-relative object acts under focal visual control (in solitary play) to the extent of seemingly ignoring people. Having escaped from a book sharing interaction after barely 2 minutes, the 7 month old boy in figure 6.1.c3.b immediately afterwards engaged in manipulating a single object for nearly 6
minutes without interruption, throughout apparently “ignoring” his mother, that is, not once looking at or responding to her [see book sharing section 3.4.1.3]. Even when involved with objects, infants did, however, from time to time look up at people’s faces, e.g. when there was unexpected noise or motion around, when the mother uttered attention directives such as “look!” even if addressing a sibling, when newly introduced to an object, when checking the adult’s reaction or attitude to their own actions. These might be early forms of instrumental looking, for instance, after having pushed a book out of reach, a 6-month-old girl lying on her belly turned her head up to her mother’s face and vocalized (figure 6.1.i1.c).

6.5.3 Summary: participation at 6-9m

The trends of the previous months’ gradual development continued, with object engagement getting more autonomous and focused, and joint activities emphasising social engagement increasingly tuning in to the coordination of widely varied action structures, now however leading to a qualitative change in global engagement patterns: an apparent split between social and object engagement.

Along with improved postural control allowing free and stable sitting and with beginning locomotion, infants’ object engagement became more autonomous, e.g. less dependent on adults with regard to acquiring and keeping objects, and their attention was largely drawn to and bound by their own object actions: They now performed extended bimanual manipulations of one or more objects under focal visual control and with serious intent, and their attentive-communicative and physically-effective manual engagement converged and were strongly coupled around the object – resulting in largely solitary object engagement even to the point of infants apparently “ignoring” people. While in this period preferring solitary engagement in object-centred activities, infants, however, also keenly participated in exaggerated, even boisterous, social (often object-mediated) play, responding flexibly, appropriately, and with laughter not only to playfully varying the space-time-intensity contours of actions, but also to combinatorial variations,
making their intentions clearly visible and contributing to and modulating the jointly created variations through their responses.

In these interactions, led by the caregivers but with infants’ contributions steering the direction, they together further parsed and delineated separate and now exchangeable sub-parts of the jointly established routines and highlighted transition or decision points as places for co-ordinated negotiation, thus jointly creating a unique play sequence from partly known and partly new components. This transcending of contexts and transferring of actions between contexts can thus be considered as a basis for e.g. the development of symbol use.

6.6 9-12 months: co-composing complex social action structures including manual object-relative object acts.

As the emphasis for data analysis has primarily been on the first 8 months until now, we only analysed in detail a limited number of interactions from the age period between 9 and 12 months. Below we present a tentative first sketch and brief outlook based on our observations so far.

In the period between 9 and 12 months infants smoothly integrated attentive-communicative and whole-body as well as manual object actions in service of persistently pursuing one or another action sequence directed towards and involving objects and/or people. They were able to relate and deal with multiple engagements in the form of complex (inter-)action sequences at once, splitting up and distributing different modalities and strands of actions between them, and then integrating them across, and thus bridging, different contexts, and now making some singled out aspects of one relevant for the other.

Homing device: For the whole course of the research visit at nine months, the girl in figure 6.1.c4.b, no matter in which activity she was otherwise engaged in, tended to end up commando-crawling towards the objects in front of the fireplace where she was not supposed to go. Approaching the fireplace she intermittently stopped, turned around, grinned, and looked at her mother’s face, who was kneeling on the floor at some distance behind her, and only proceeded after the mother responded. During one particular approach the mother’s
response at the first stop was a low-voiced twice-repeated “where are you going?”, the second an “I know where you’re going!” with a slightly threatening undertone, met by the infant with a bout of exuberant arm and leg thrashing and laughter, and finally “I’m gonna get you”, accompanied by a ritualized mock attack landing next to the infant, before actually picking her up with a “NO” and placing her at some distance and facing away from the fireplace upon which the girl unerringly aimed for the fireplace again (so that the merry game could begin again).

In this interaction the girl integrates and utilizes and co-ordinates, how one (inter-)action sequence directed towards – and constituting her engagement with – the objects in front of the fireplace, employing a specific set of perception-action resources, is related to and impacts on another (inter-)action sequence directed towards – and constituting her engagement with – the mother, employing a somewhat different set of perception-action resources, and vice versa – giving rise to a game. The participants are both oriented in a shared space of meaning and action: they are both aware of the point in space (and time), where the infant’s actions tend towards, of the action trajectories both participants follow, the specific roles both contribute to the interaction, they are aware that they follow opposing forces and action tendencies, and the form of “negotiation” which is going to take place. The girl’s stopping in front of the fireplace and turning to the mother and her response to the mother’s answer with exuberant thrashing and laughter shows, that the goal or topic or goal the interaction is not the action to get to the fireplace, which then is negotiated, but that negotiation of or about something has itself become the topic.

Symbol use: When the mother in Figure 6.1.c4.c, after the infant has handed her the book, is pointing to, vocalizing and signing “bird” referring to a picture in the book, the 12 month old infant is turning around and looking out of the window. Not realizing this, the mother first finishes her signing, and then – recounting how they had seen a seagull outside the window on the previous day – herself turns to the window – only then realizing that the infant was already looking there.

In this interaction the 12 month-old boy singles out specific aspects of his
book-sharing engagement with his mother (a specific cluster of actions, meanings, invariances spanning the mother’s vocalizations, signing, and acts on the book) as links to another, memorized action context which had occurred at a different time and space. Both partners’ actions impact on each other’s orientation within a shared space of meaning and action. They allowed re-orienting and directing of oneself and the partner towards specific points of reference within this shared sense- and action(land)scape, spanning space and time, partly “virtual” and partly rooted in the presently shared situation. Aspects of this shared situation, presently available for perception and action and thus for affecting each other, are used as landmarks and tools for the alignment of orientation and pointers to the virtual shared space.

6.6.1 Summary: participation at 9-12: negotiation of negotiation

Infants were now able to flexibly co-ordinate affective-communicative engagement with whole-body as well as physically-effective manual object actions (including complex person/object-relative object acts), either in convergence or split up and distributed between multiple action strands and objects of engagement, and integrated into one complex action sequence. Infants were now robustly oriented, and flexibly moving in, a number of highly structured shared meaning-and-action spaces, parts of which (e.g. certain actions, objects, “goals” etc.) could be singled out to direct the partner’s attention, or could be transferred between to link different contexts/action-spaces. The partners were now both navigating a network of intricately structured, partially overlapping and connected shared spaces of meaning and action constituting a basic shared sense-and-action(land)scape, providing joint affordances and constraints to act within, as well as “building blocks” to jointly orient towards, act on and transform.

Moving one step further still, continuing the development which went from playful variations of the space-time-intensity contours of joint actions to combinatorial variations of these joint structures to the point where the infant is not only actively directing the mother’s attention and actions, but also where the shared topic of the interaction itself is not simply a specific “part of the action”, or an “object” anymore but is now the “negotiation of the action”. In
this triadic participation infants were thus moving one step further towards full-fledged participation in training to negotiate negotiation.

Conclusions and Outlook

6.7 From engaging in simple (dyadic) interchanges to participating in, negotiating and co-creating complex (triadic) interactions: the role of jointly enacting affect-imbued action arcs.

How do infants arrive there? How do young infants manage to understand and master these complex action structures and to share, jointly negotiate, and modify them, let alone learn to relate another complex structure (such as language) to these complex structures to communicate about them?

Our exploratory investigation of everyday infant-caregiver-object interactions and social object practices over the first year suggests that in order to answer these questions we need to take a closer and more systematic look at the development of the infant’s action co-ordination skills. These meet various complex co-ordination challenges, consisting not simply in co-ordinating “engagement with people and objects”, but rather in co-ordinating one’s own multiple action sequences directed towards different targets of engagement by orchestrating multiple modalities and strands of action, as well as co-ordinating them with someone else’s. This in turn requires some basic understanding of (often complex) situations, as well as of the interaction partners and their actions.

6.7.1 The development of action control and co-ordination

6.7.1.1 Affective-communicative engagement

The action systems involved in attentive-communicative engagement were the first ones to work in a functionally effective way: Infants effectively used gaze, facial expression, vocalization, and whole body movements from early on in social engagement in proto-conversation and beyond. However affective communicative-engagement went through a process of development, where the control of the process shifted several times between being primarily externally versus internally driven, before a balance between inner and outer control was
reached (moreover, these actions also needed to be co-ordinated with infant’s physically-effective actions):

In the beginning, infants were highly responsive, e.g. have their gaze drawn to outside events, their attentive-communicative engagement was under strong outside influence and control. While generally functionally beneficial, resulting in infants looking at key events (“where the action was”), that is, key points for picking up information, this also led to distraction, in particular during infants’ first limited efforts to pursue their autonomous actions, and often resulting in a divergence between their attentive-communicative engagement and their physically-effective object engagement. Later on affective-communicative engagement was increasingly controlled from within, reducing distractions, and was drawn to and converging with infant’s own physically-effective actions with which it needed to be co-ordinated (see below).

6.7.1.2 Physically-effective engagement

Physically-effective engagement, in particular manual object actions, took longer to become functionally operating in the real world. While aimed (pre-)reaching movements were observed in controlled laboratory conditions with postural support from early on (von Hofsten, 1984), manual object actions in real life situations, even when scaffolded, were much more difficult, as simultaneous control of posture was needed, as well as coordination with outside events and influences, e.g. distracting sights, noises, or cheering mothers. These challenges of co-ordination have to be taken into account, in particular when considering young infants’ participation and cooperation in joint activities. These are usually defined as requiring physically effective engagement “bringing about a change in the world” (Sebanz et al., 2006), which in typical test situations for co-operation take the form of e.g. inserting a small object into a toy truck (Hubley & Trevarthen, 1979), which actually constitutes a rather complex and late-occurring object-relative object act. When considering infants’ participation, we thus need to further differentiate infant actions and take into account which kind of actions the infant is already able to perform and, conversely, how difficult a specific action might be in terms of infant motor
control, as well as how actions are coordinated with objects and other people’s activities.

6.7.1.3 The Development of co-ordination between affective-communicative and physically-effective engagement.

Performing complex actions in the real world often requires intricate co-ordination between attentive-communicative engagement and physically-effective engagement, and the action systems involved therein.

While infant’s body and limb actions and the working of sensory modalities have been shown to work in a coordinated way from the beginning – even in the womb (Castiello et al., 2010; van der Meer, 1997; van der Meer, van der Weel, & Lee, 1995), however, this was not always straightforward in complex, real world environments. When 3 months old infants e.g. attempted to hit objects, we often observed a divergence between physically-effective and attentive-communicative engagement, likely due to difficult postural control, as well as due to distraction through outside events while pursuing their own object actions. In the course of the next months (around 4-7 months), we see an increasing convergence between those two ways of engagement. By 6-7 months, infants’ attention was bound to monitoring their own object manipulation under focal visual control in sustained solitary object play, at the cost of largely ignoring external influences including other people and their efforts to engage them. Only later (around 8-9) infants were able to decouple different modalities and strands of action, splitting up and flexibly distributing them to different targets of engagement, and integrating them to into complex, multi-step, multi-target action sequences (compare de Barbaro, Johnson, & Deák, 2013).

6.7.1.4 The development of infant’s functional contributions to a shared object activity

Taking into account the challenges and developments of motor control and co-ordination described above, we can take a more differentiated look at the motor side of infant participation and their functional contributions to the activity. Looking at the data, we can differentiate different types of infant actions and describe their gradual sequential occurrence in participation:
Infants from the beginning (3 months) functionally contributed through anticipatory whole-body movements like the bottom lift in nappy change performed in adaptation to and anticipation of the mother’s object actions. From 4 months on, they contributed to the task using body-relative object acts, where the object is translated by the infants relative to their own body, which can be at least partially controlled by relying on proprioception rather than having to rely strongly on precise focal visual control: e.g. grasping to mouth an object, pulling a blanket over one’s head to play a round of peekaboo, or to go to sleep. Only later, from around 5-6 months on they engaged in counter-resistance-related object acts, e.g. banging objects together or onto surfaces (though mostly done in solitary play), and only around 8-9 months they started contributing by engaging in precise person- and object-relative object acts, where objects are translated or manipulated in relation with or regard to another person or object, e.g. placing or giving an object.

6.7.2 Understanding and orienting within a complex task in accordance with a partner and his/her actions

However, to successfully participate and functionally contribute to an activity, infants – in addition to the ability and skill of controlling and co-ordinating their own actions – also need to be oriented within, that is, need some basic understanding of the complex task, its sub-parts, their interaction partners and their actions, in order to contribute the appropriate action at the appropriate time and co-ordinate their own actions with those of the partner. How are infants able to develop such an understanding?

Our micro-analyses of early infant-caregiver-object routines suggest that one key factor enabling understanding and participation may be jointly moving through affect-imbued action arcs intricately shaped and patterned by caregivers (see figure 6.1).

6.7.3 Action arcs as a key organizational principle in action co-ordination

We first noticed action arcs – with their distinct structure of beginning, build-up, climax, and resolution – in the pitch and intensity contours of the caregivers’ verbal utterances where they are most easily recognizable. Taking a closer look at the transcripts of the various parallel action strands the action arc stood out
as a general organizing principle running through the entire action; they can be found across and co-ordinate various strands of action. Relevant events, such as important words, can be found at the peak of the interactions, infants’ contributions are likely to be found at the peak or in-between the arcs (compare Brazelton et al., 1974; narrative or shared project: Delafield-Butt & Gangopadhyay, 2013; vitality contours: Stern, 2010; Trevarthen & Delafield-Butt, 2013).

We suggest, that in these early highly structured affect-imbued forms of interactions – rather than being merely exchanges and attunements of affect and behaviour having nothing as yet to do with cognition or knowledge (Tomasello et al., 2005) – a lot can be learned both in terms of knowledge about the complex action structures of self and other as well as in terms of the social co-ordination and negotiation of these, which, we argue, go hand in hand.

Jointly moving through shared action arcs allows the infant to:

1) Experience the sharedness of acting together, by simultaneously moving through the experience of the specific action along with its specific contour of arousal and affect together

2) Experience qualitative aspects of sharing the action, such as the degree of smoothness or friction, of alignment or opposition of forces

3) Learn, (understand, and predict) a simple general action structure with a start, climax, and ending.

4) Learn to participate, that is, to take over a specific action part at a specific time, fulfilling a specific role, scaffolded by caregivers utilizing the natural arc dynamics: delaying the climax and thus creating a tension and an opening, instigating a tendency to move in and complete the arc.

5) Learn to distinguish, recognize, and to participate in different simple actions by their different specific action contours.

6) Learn to jointly relate to, reflect on, communicate about, and modify, that is, reflectively engage with jointly established action arcs.

7) Learn, understand, and (jointly) perform new complex object- and/or conventional actions by mapping their complex structure onto familiar shared, predictable and hence meaningful action arcs, which can be nested into each other, thus over time creating an intricately structured
continuously refined sense-and-action(land)scape, weaving the substrate of culture.

1) Moving through action arcs together may contribute to the experience of the "sharedness" of an action in two ways: a) by creating a - at least to some extent - similar experience of performing the action itself, along with similar trajectories of arousal and affect in both partners in parallel through the dynamics of the action itself: i.e. changes in velocity (action, sound), intensity (action impact and volume), and elevation (motion height and pitch); and b) by further mutually reflecting and reinforcing these action, arousal, and emotion processes which the participants go through, by addressing, responding to and co-regulating each other (also see section 4.3: Nappy Change).

2) Moving through action arcs together the partners not only experience a "sharedness" but also its specific quality: is the process unfolding smoothly or is there some friction, are they moving in straight alignment or is there a lot of negotiating going on pulling in different directions, and is the negotiating going softly or can opposing forces be felt? Thus a lot can be learned about self and other (see chapter 5), about will, interest, purpose, goals, and skills, as well as about the relationship, mutual understanding, and co-ordination of the partners (also compare Bühler, 1927: "Only through resistance of the partner [...] do we learn something important about his/her soul/mind").

3) Action arcs are already and intuitively familiar to infants as they are inherent in action quite generally, both cyclic and transitive, such as breathing, thrashing one's arms and legs, experiencing a surprise, or reaching and grasping an object (Churchland et al., 2012). Whether one is performing the action oneself or engaging in it by observing (Flanagan & Johansson, 2003), such actions share to some extent common time-space-intensity dynamics with a start, climax, and end, which each of these actions is tending towards. Thus these actions in a sense already have an inherent goal, a certain form of "in-tent-ionality" already built into them. Although young infants do not necessarily as yet understand "goals" as rationally set ahead of time, and persistently pursued via strategies planned in advance, and then flexibly implemented in accordance with the unfolding situation and dealing with potential obstacles, nonetheless
infants are familiar with the experience of action arcs, the focused and directed energy necessary to get them started and keep them moving along a specific trajectory, and finally “tending towards” a certain end point. Thus every process shaped into that action-arc structure can be experienced and understood in these simple structural terms: a motivation-fuelled start, climax, and the finding of a resolution in an endpoint.

4) Whether it was in nappy change (see chapter 4), peekaboo (see figure 6.2), or book sharing (see chapter 3), mothers utilized and enhanced the inherent dynamics of action arcs to draw infants into participation: using the specific actions of the respective task they opened and built up an action arc, then delayed the climax, creating tension and a specific opening (see also Mehus, 2011: “anticipatory contextualisation”), instigating a tendency for infants to move in and thus resolve the tension through contributing to completing the arc and leading it towards resolution. Repeatedly using this general arc-on-hold
pattern across the different task contexts and manifesting it with their specific action “material”, mothers not only train infants to participate appropriately in the respective task but also draw infants into participation and “readiness to co-operate” in general (compare Fantasia et al., 2014; Mehus, 2011; Reddy, Liebal, et al., 2013). Moreover, in the course of patterning and shaping the activity into action arcs, caregivers also used conventional invitations such as questions: “Are you ready?”, “Do you wanna play?”, “Can you turn the page?” as markers or signals at the participation-take-off-point, as well as conventional phrases such as “Thank you”, “There you go”, or “All done” to acknowledge the infant’s contribution and (formally) close the activity. Thus, in addition to training infants to appropriately participate in the specific task and drawing them towards a general “participatory” or “co-operative” stance, the caregivers trained them to enact the conventional script, interaction format (Bruner, 1985), or pragmatic frame (Rohlfing et al., 2016) of “co-operative interchange”. And we suggest that mapping this conventional format onto the shared affect-imbued action arc may contribute to learning the conventional format (see point 7 below).

In the nappy change example the training of this conventional interaction format became particularly evident. By 6 months it had become ritualised, generalised and to some point conventionalised: the mother employed “Ta”, a conventional short form of “Thank you!” which she previously had only used towards the end of the nappy change to acknowledge and thank the infant for her contribution, now at the lift-off point as a(n arbitrary) signal for the infant to act (in particular when the infant was not paying attention). Is was used as an arbitrary signal in the sense that it apparently neither functioned exclusively qua a rising and hence inviting intonation contour (which was much less pronounced than before), nor qua being a context specific index inherent in the necessary steps of the activity (such as the noise and sight of the nappy strap being opened), rather it was condensed from the “thank you!” common to the acknowledgment phase of “co-operative interchanges”, transposed to the front and used as a signal of request, invitation, or imperative in a generalised way for “co-operative interchanges” across different activity contexts.
5) While actions in general can be considered to share an over-arching structure at the most global level, see above, clearly specific actions vary extensively in their structure and arc-shaped contour, and can thus be distinguished from a different action having qualitatively different dynamics and accordingly play a different role in a larger context, which can be distinguished as a different “meaning”. This point seems obvious for the contours and trajectories of manual actions: throwing an object has a distinctively different action- and control contour, and hence a different function or “meaning” within a larger context, than carefully putting an object into a confined space; showing differs from giving; setting a baby gym in motion with specific whole-body movements has a different contour than lifting the bottom in nappy change; pulling the blanket over one’s head and back down to cover and uncover the face in peekaboo has a different contour than pulling the blanket over one’s head to go to sleep; a command or threat uttered in a playful voice has a different dynamic and meaning than one shouted in the fearful tone of emergency, etc. (compare with, and note the difference to, Stern, 2010).

6) The intuitive form of action arcs allows specific actions to be enacted together, repeatedly, in a consistent and understandable way, and hence to become established as a common shared routine. Once established as a routine, the actions become reified as a shared familiar structure providing the material to jointly relate to, in other words reflectively engage with, either commenting on and communicating about, or modifying the action. The arc shape suggests and helps to localize and co-ordinate around specific points for joint engagement, typically located either at the (often delayed and hence marked) climax, or between the component arcs of an activity.

   Thus, establishing (a sequence of) action arcs as a shared routine facilitates the development of joint reflective engagement, crucial for establishing a joint focus of concern, and hence triadic engagement, as well as the resulting joint negotiation of activities.

7) There have been many efforts, in particular in the 20\textsuperscript{th} century with its major interest in formal structures, to characterize the organizational structure of activities, which form the larger context of - and hence give meaning to - particular actions or utterances. How do participants come to know (often put
in terms of “representation”) and master such complex structures, that is, participate in, co-create, transform, and further develop them, e.g., by newly re-combining or transferring components between contexts. These efforts mostly emerged from struggles with the (logic-based) analysis of (primarily linguistic) meaning, where meaning could not be accounted for in terms of (the meanings of) its constituents, and with the directly related problem of how (language) “meaning” is learned (by infants or by computers) and taught.

Reviewing these accounts (from Wittgenstein to Bateson, Goffman, Minsky…) goes unfortunately well beyond the scope of the present work. For our purposes, we note that these accounts have in common is that they are relatively abstract, be it an arbitrary, albeit structured, list of events of actions in the case of Schank & Abelson’s scripts (Schank & Abelson, 1977), be it a hierarchically branching structure-tree, as Bruner’s famous depiction of peek-a-boo adapting Chomsky’s syntax tree (Bruner and Watson, 1983). While capturing the analytic insight of the researcher, such framing of activities does not give us any clue to how infants may learn such complex multi-step actions, whose goals may not even be apparent to them.

We suggest that jointly moving through action arcs may play a crucial role to bridge the gap between apparently “simple” dyadic interactions, and “complex” “cognitive” triadic interactions: we propose that infants can learn new complex object- and/or conventional actions, including sequences of sub-actions and distributed roles, even without knowing the goal and instrumental steps towards it, if those complex actions are mapped onto familiar shared action arcs, anchoring and locating object-related and role-specific sub-actions in their affect-imbued, dynamically predictable arc structure, thus creating a complex shared and embodied – and hence meaningful – action(land)scape (see figure 6.3).

Thus the concept of jointly moving through action arcs provides a potential (set of) learning mechanism(s) at a meso-level, a key missing bridge between local learning mechanisms based on association (as e.g. described by Rączaszek-Leonardi et al. 2013; and captured by the concept of adjacency pairs by Heller & Rohlfing, 2017), and how infants navigate, parse, and learn the complex structures of larger activities at multiple levels in a meaningful way – and
through this may even come to know goals of activities and intentions of other people.

**Figure 6.3**: Understanding complex actions through mapping them on affect-imbued shared action arcs

6.8 **Shifting the perspective from triadic interaction to jointly structuring shared multi-modal sense-and-action(land)scapes.**

Linking these results and interpretations back to the development of triadic interactions, we thus suggest that the third pole in triadic interaction, which we assume emerges, crystallizes, and differentiates gradually over time, is not simply a single object – indeed, successful object labelling marks only the easily visible tip of an much larger iceberg. Rather, the third pole is constituted by the jointly established interaction structures (including action arcs) themselves, forming a shared orientation framework, a sense-and-action-(land)scape, (parts of) which in the course of being enacted may be manifested in and hence tied to material objects, body parts, etc. However, objects (self, other, as well as specific activity sub-parts and the specific roles they may be assigned to) all get their jointly established meanings (including at some point the symbolic and conventional aspects thereof) through their place in such a dynamic, constantly jointly refined shared orientation framework. Jointly enacted action arcs are the first basic version of these jointly created orientation frameworks, affect-imbued, shared, predictable and hence meaningful – and over time differentiate into rich jointly created and systemically structured sense-and-action(land)scapes co-weaving the substrate of culture.
7 General Discussion

The results from conceptual and empirical analysis accumulating over the course of the thesis culminated in a change of framework:

Typically participation, joint action, co-ordination, and triadic interaction are all conceived of in a cognitive and rational way. Put in a slightly oversimplified nutshell: co-ordination is considered trivial with respect to motor control (e.g. gaze-alternation), but requires a lot of knowledge: 1) to be motivated to perform action co-ordination at all (one needs to know that other people have goals, intentions, beliefs in order to be motivated to engage in joint attention and perform gaze alternations to see where they are looking), and 2) to select the right “building block” in a serial sequence of actions (one needs to know the complex hierarchical structure and role distribution to appropriately contribute to the action sequence) (e.g. Tomasello et al. 2005, Knoblich et al. 2011, etc.). What – taking a closer look – all these activities seem to have in common is that they require the ability to “take into account the relations between “things”, a characteristic neatly and minimally captured in the image of the “triad”: one node taking into account the relation between the other two nodes. But if triadic interaction is the door to participation and is supposed to rest on joint attention, itself considered to depend on complex, relational knowledge (e.g. understanding people as intentional agents with distinct goals), then it becomes hard to account for the development of these abilities (except claiming their sudden emergence). To briefly summarize cognitive approaches: the infant has to know in order to participate.

Here, however, following theoretical and empirical analysis, we shift the framework to a more embodied and situated perspective (in line with, building on and further developing approaches such as: Alač, 2005; Bühler 1927; De Barbaro et al., 2013; De Jaegher & Di Paolo, 2007; Goodwin, 2013; Fantasia et al., 2014; Rączaszek-Leonardi et al., 2013; Reddy, 2008; Rogoff, 1990; Rohlfing et al., 2016; Yu & Smith, 2013; Zukow-Goldring, 2012): From this perspective the challenges the infant has to master to fully participate as a competent member in complex cultural activities are 1) action control and co-ordination, on many levels, intra and interpersonally, as well as 2) being oriented in
complex shared action structures and sense-and-action-scapes. However from this perspective, we can also suggest a pathway, how the infant might be able to get there: generally, via participation. The infant participates in order to know.

From this perspective and based on our video data analysis we can tentatively outline participation as: appropriately contributing a specific action to a shared, larger, structured, meaningful whole, as a socially legitimate participant (also compare Rogoff, 1990; Lave & Wenger, 1991). This working definition can serve as a common framework capturing both: 1) the complex interactions emerging towards the end of the first year, typically set apart from earlier ones as “triadic”: e.g. giving and taking, following instructions, pointing to and labelling objects etc. However, instead of accounting for them in terms of “a third pole entering”, or “co-ordinating engagement with and taking into account the relation between two things”, we rather put these forms of interaction into a larger context and suggest to conceive of them as “being orientated within a complex activity structure” and “actively contributing in a co-ordinated way” (compare section 6.8.). And this also allows us to take into account 2) earlier forms of participation at different levels of complexity: differentiating how both the infants’ contributions as well as the “structured, meaningful whole” start out simple and grow increasingly complex over time allows us to better account for development. Moreover 3) the framework points out the meaningful activity structure as a scaffold or container supporting this increasing differentiation. As we have seen in chapter 4, caregivers – from the beginning of the study – invited and drew infants into participation in nearly all ecological activity contexts (except delicate ones such as bathing or when putting the baby to bed). They performed verbal utterances and gestures, which we as competent members of the culture recognize as markers organizing particular cultural scripts and interaction formats, but which to young infants do not mean anything yet. We suggest, however, that orientation within a “meaningful whole” is provided and enabled by affect imbued action arcs, which caregivers shape their actions into, helping infants to orient and drawing them into participation from the beginning. We further suggest that by mapping of novel complex actions (with their conventional markers) onto the familiar affect
imbued action arcs caregivers make these complex actions understandable to the infants. Over time they create shared sense-and-action-scapes, which are embodied and at the same time systemically structured, allowing the participants to orient each other towards distinct locations within these shared networks of meaning, which may serve as a basis for joint action, cooperation, and symbol use.

This is only a rough sketch, which invites further exploration, thorough testing, and which we need to understand in more detail. Also our empirical study was limited in many respects: while naturalistic longitudinal studies bring us close to real life, helping us to come up with a better estimation of the range of activities infants participate in, and while qualitative micro-analysis allows us to capture the co-ordination of interaction as is unfolds over time in intricate detail, their results are hard to generalize. Engaging in a time and labour intensive research method we can work with only a small sample size of participants who also may not be representative for the wider population: the 16 families of our study (6 pilot, 10 main study) were not only all coming from UK and Austria reflecting Western and Middle European culture, they also were all middle class families, which e.g. might make them overly biased towards nudging their infants into specific ways of co-ordination considered beneficial to their education, promising social advance (Bourdieu, 2013). Thus – in order to get a better understanding of the varieties of infant participation in complex cultural activities and their co-ordination – the study will have to be considered in a larger context together with studies including different groups, who might engage in different ecological activity contexts using different forms of co-ordination (as indicated e.g. by Keller, Völker, & Yovsi, 2005, indicating that Nso mothers’ interactions did not get more object oriented between three and six months). And even within our small sample, with a sampling rate of half a day visits once a month, while providing many hours of rich video data, we only get a few glimpses into highly complex non-linear processes of development (Thelen & Smith, 1994). Not imposing control on the interactions, the study provides a more realistic but also highly variable sample of infant-caregiver-object interactions, making it hard to (in particular quantitatively) compare
interactions between different families, or even within the same family at different points in time.

However, observing co-ordination in detail, analysing, and describing which actions and forms of co-ordination occur (or do not occur), comparing them across different contexts and tracing their change over multiple time points – this method allows us to come up with new questions, a tentative trajectory for the development of joint object actions, and concrete hypotheses which lend themselves to testing and further investigation:

- How are the respective modalities utilized and playing together during action-co-ordination depending on the demands of the respective contexts and the characteristics of the respective modality?
- Can we show the infant’s gaze becoming anticipatory as routines get established?
- How do eye-hand co-ordination, co-ordination between multiple action strands and social co-ordination between participants play together?
- Is it possible to make the bodily action aspects of action arcs measurable with the help of e.g. a motion capture system?
- Can we test and manipulate the functional role and effect of action-arcs for infants’ understanding, participation and learning?
- ...

To address these questions, among the next steps is taking selected activities from some of the ecological activity contexts infants are typically involved in into the laboratory to investigate them in more detail under more controlled conditions. We are piloting data collection technologies such as lightweight mobile gaze and motion tracking to allow 1) staging semi-naturalistic interactions in more lab-setting and 2) conversely, taking these tools out of the lab, to enrich the possibilities of flexible home visits and to combine engaging with rich real life situations with measurement.
8 Bibliography


del Rio, & A. Alvarez (Eds.), *Sociocultural Studies of Mind*. Cambridge, MA: Cambridge University Press.


321
### FORM UPR16

**Research Ethics Review Checklist**

Please include this completed form as an appendix to your thesis (see the Postgraduate Research Student Handbook for more information).

<table>
<thead>
<tr>
<th>Postgraduate Research Student (PGRS) Information</th>
<th>Student ID:</th>
<th>64090401</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGRS Name:</td>
<td>Nicole Rossmanith</td>
<td></td>
</tr>
<tr>
<td>Department:</td>
<td>Psychology</td>
<td></td>
</tr>
<tr>
<td>First Supervisor:</td>
<td>Vasu Reddy</td>
<td></td>
</tr>
<tr>
<td>Start Date: (or progression date for Prof Doc students)</td>
<td>October 2011</td>
<td></td>
</tr>
<tr>
<td>Study Mode and Route:</td>
<td>Part-time</td>
<td>MPhil</td>
</tr>
<tr>
<td></td>
<td>Full-time</td>
<td>PhD</td>
</tr>
<tr>
<td>Title of Thesis:</td>
<td>Culture in the Making: Jointly structuring shared spaces of meaning and action in infant-caregiver-object interactions over the first year of life</td>
<td></td>
</tr>
<tr>
<td>Thesis Word Count: (excluding ancillary data)</td>
<td>86846</td>
<td></td>
</tr>
</tbody>
</table>

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University’s Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study. Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

**UKRIO Finished Research Checklist:**

(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee rep or see the online version of the full checklist at: [http://www.ukrio.org/what-we-do/code-of-practice-for-research/](http://www.ukrio.org/what-we-do/code-of-practice-for-research/))

<table>
<thead>
<tr>
<th>Question</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame?</td>
<td>☒</td>
<td></td>
</tr>
<tr>
<td>b) Have all contributions to knowledge been acknowledged?</td>
<td>☒</td>
<td></td>
</tr>
<tr>
<td>c) Have you complied with all agreements relating to intellectual property, publication and authorship?</td>
<td>☒</td>
<td></td>
</tr>
<tr>
<td>d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration?</td>
<td>☒</td>
<td></td>
</tr>
<tr>
<td>e) Does your research comply with all legal, ethical, and contractual requirements?</td>
<td>☒</td>
<td></td>
</tr>
</tbody>
</table>

**Candidate Statement:**

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s).

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):

If you have not submitted your work for ethical review, and/or you have answered ‘No’ to one or more of questions a) to e), please explain below why this is so:

Signed (PGRS): [Signature]  
Date: 28.09.2017