Intermodal Match Differentiation in the First Year of Life

Sarah Klein-Radukic
Intermodal Match Differentiation in the First Year of Life: Methodological Considerations and its Role in Interaction Difficulties and Later Self-Recognition

Inaugural – Dissertation
zur Erlangung des Grades eines Doktors der Naturwissenschaften
in der
Fakultät für Psychologie
der
RUHR-UNIVERSITÄT BOCHUM

Vorgelegt von:
Dipl.-Psych. Sarah Klein-Radukic
Gedruckt mit Genehmigung der Fakultät für Psychologie der

RUHR-UNIVERSITÄT BOCHUM

Referent: Prof. Dr. Norbert Zmyj
Koreferent: Prof. Dr. Axel Schölmerich
März 2017

Datum der mündlichen Prüfung: 27. September 2017
# Table of Contents

Abstract .................................................................................................................................................. 7

1 Intermodal Match Differentiation in the First Year of Life ................................................................. 11
   1.1 Experiencing Intermodal Matches .............................................................................................. 11
   1.2 Sources and Nature of Perfect Matches, Less-than-perfect Matches and No Matches in Intermodal Match Differentiation ......................................................................................... 12
      1.2.1 Vote for a Sensible Use of the Terms “imperfect contingent”, “less-than-perfect contingent”, and “noncontingent” ........................................................................................................ 12
      1.2.2 Sources of Less-than-perfect Matches: Contingency, Congruency, Timing .......................... 13
         1.2.2.1 Contingency .................................................................................................................. 14
         1.2.2.2 Congruency .................................................................................................................. 14
         1.2.2.3 Timing ......................................................................................................................... 14
      1.2.3 Three Categories: Perfect Match, Less-than-perfect Match, No Match ......................... 16
   1.3 Types of Intermodal Match Differentiation Tasks ......................................................................... 19
      1.3.1 Visual-Proprioceptive Matches ............................................................................................ 20
      1.3.2 Visual-Tactile Matches ........................................................................................................ 20
      1.3.3 Visual-Auditory Matches ..................................................................................................... 21
   1.4 Intermodal Match Differentiation vs. Intermodal Match Preference ............................................ 22
   1.5 Relationship of Intermodal Match Differentiation and Self-Recognition .................................... 23

2 Relationship between Intermodal Match Preference and Social Interaction .................................... 71
   2.1 Intermodal Match Preference and Social Interaction .................................................................... 71
   2.2 Study 2 – Zmyj, N., & Klein-Radukic, S. (2015). Six-month-old infants’ interaction difficulties are mirrored in their preference for perfect contingencies. Infant Mental Health Journal ................................................. 73

3 Dropout in Infancy Research ............................................................................................................... 91
   3.1 Dropout in Looking Time Studies ................................................................................................ 91
   3.2 Dropout in Longitudinal Studies ................................................................................................ 92

4 General Discussion ............................................................................................................................ 123
   4.1 Relevance .................................................................................................................................. 124
   4.2 Limitations ................................................................................................................................ 127
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3 Future Studies</td>
<td>129</td>
</tr>
<tr>
<td>5 Conclusion</td>
<td>133</td>
</tr>
<tr>
<td>References</td>
<td>135</td>
</tr>
<tr>
<td>List of Figures</td>
<td>143</td>
</tr>
<tr>
<td>List of Tables</td>
<td>144</td>
</tr>
<tr>
<td>List of Publications</td>
<td>147</td>
</tr>
<tr>
<td>Curriculum Vitae</td>
<td>149</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>151</td>
</tr>
<tr>
<td>Erklärung über den Umfang des eigenen Beitrags zu jeder anzurechnenden Publikation</td>
<td>153</td>
</tr>
<tr>
<td>Eigenständigkeitserklärung</td>
<td>155</td>
</tr>
</tbody>
</table>
Abstract

The processing of information arising from different perceptual systems is assumed to have important implications for the development of infants’ ability to differentiate between the self and others (Bahrick & Watson, 1985; Rochat & Striano, 2000). The current dissertation aims at disentangling different definitions and paradigms used in previous theoretical frameworks and studies on infants’ differentiation between different degrees of intermodal matches. Three main characteristics (contingency, congruency, and timing) are identified in order to define the nature of perfect, less-than-perfect and no intermodal matches. The differentiation between these different degrees of intermodal matches is assumed to be related to the ability to recognize the self in reflecting surfaces (Rochat, 2003), which is an important milestone in infants’ development. The central question of the first publication of the current dissertation was whether infants’ intermodal match differentiation at 6 and 9 months is a precursor of their ability to recognize themselves in the mirror at 18 months and on video at 26 months, which was investigated in a longitudinal study. Findings of study 1 indicated that infants’ intermodal match differentiation at 6 and 9 months was not predictively related to their mirror self-recognition status at 18 and 26 months or to their video self-recognition status at 26 months. The absence of predictive relationships between intermodal match differentiation and later self-recognition supports the assumption that the representation of the self in the imagination develops independently of infants’ ability to differentiate between different degrees of intermodal matches. Thus, an alternative explanation for the development of self-recognition is likely: Self-recognition might emerge as the result of a sudden cognitive change (Bischof-Köhler, 1991, 2012). In study 1, infants’ intermodal match differentiation was measured in a task that presents infants with different degrees of self-generated intermodal matches. A previous study showed that infants’ intermodal match differentiation in social interactions at 9 months predicted mirror self-recognition at 24 months (Kristen-Antonow, Sodian, Perst, & Licata, 2015). This finding offers an alternative
explanation, as it highlights the importance of intermodal match experiences in communication and interaction between parents and their infants. Intermodal match experiences regarding nonsocial stimuli like self-generated behavior, as used in study 1, do not seem to contribute to the development of a representation of the self with a comparable impact. Accordingly, intermodal match differentiation in the first year of life could be closely related to the differentiation between self-generated and other-generated behavior in social interactions (Bigelow, 2001). Infants’ strong preference for self-generated behavior (perfect intermodal matches) might result in interaction and communication difficulties with others, especially their parents. Thus, the second publication aimed to test whether infants’ intermodal match preference is related to parent-reported problems in interaction and communication with their infants. Infants’ preference for perfect intermodal matches as opposed to a preference for no matches was associated with parent-reported difficulties in the interaction between parent and infant.

In both studies, infants’ intermodal match differentiation and preference was assessed in a looking time task. Additionally, the findings of study 1 and 2 were gathered using data from a longitudinal study. Looking time studies are especially prone to dropout due to infants’ fussiness. Moreover, it might be that stable infant characteristics play a role in infants’ dropout due to fussiness. If so, it is likely that infants’ dropout in one test phase could be associated with infants’ dropout in another test phase, which is especially problematic in longitudinal studies. The third publication therefore took the form of a superordinate study assessing whether infants’ dropout during looking time tasks (such as the tasks used in study 1 and 2) affects the representativeness of the final sample in terms of infants’ temperament and cognitive developmental status, and whether dropout in one test phase is associated with dropout in another test phase 3 months later. Findings of all three studies presented in the current dissertation will be summarized and their relevance to the fields of infancy research,
clinics, and robot technology, as well as limitations and suggestions for future studies will be discussed in a General Discussion section.
1 Intermodal Match Differentiation in the First Year of Life

Self-recognition in the mirror is one of many important milestones in infants’ development (Asendorpf & Baudonniere, 1993; Lewis, Brooks-Gunn, & Jaskir, 1985; Rochat, 2003). In order to help infants to reach this milestone, websites on pregnancy and parenting advise parents to imitate their infants or play games like peek-a-boo (Heuristic, 2016). Parents’ imitation of the infant involves high degrees of intermodal matches in terms of parents’ highly contingent and congruent behavior compared to the infant’s behavior. So far, research has provided no evidence for a relationship between the suggested parent-infant activities and self-recognition. The central question of the current dissertation project is whether infants’ ability to differentiate between different degrees of intermodal matches is related to their ability to recognize themselves in the mirror.

1.1 Experiencing Intermodal Matches

The processing and coordination of information arising from multi-perceptual systems has important implications for infants’ development of the self in the first years of life. Over the past decades, researchers around the world have assessed how infants react to different degrees of temporal and spatial matches of sensory stimuli.

Perfect intermodal matches can be experienced when, for example, infants observe the movement of their own hand and compare the efferent signal (e.g., the motor signal of the hand movement) and the afferent signal (e.g., the proprioceptive signal of the hand movement). In the case of observation of one’s own hand movements, there is a perfect match – a perfect correlation – between the two signals. Perfect matches are thus specified by actions of the self (Bahrick & Watson, 1985; Rochat & Striano, 2000). “Other-than-perfect” matches can occur when infants experience a variation in the correlation between efferent and afferent signals. This is the case when, for example, parents imitate their infant; parents’ behavior is not perfectly matched to infants’ behavior, but is related and very similar. Thus, “other-than-perfect” matches are specified by actions of others (Bahrick & Watson, 1985;
Intermodal Match Differentiation in the First Year of Life

Rochat & Striano, 2000). Additionally, parents’ behavior can be unrelated to infants’ behavior.

Infants’ ability to differentiate between different degrees of intermodal matches is assumed to be related to the differentiation between the self and others (Bahrick & Watson, 1985; Rochat & Striano, 2000). Many infancy researchers have investigated the differentiation between different degrees of intermodal matches, using a variety of definitions of the different degrees of matches and with various different paradigms. The aim of this first chapter is to disentangle the definitions and paradigms used in these studies. In a next step, I will provide an overview of several paradigms used to assess infants’ ability to differentiate between different degrees of intermodal matches in the first year of life.

Intermodal match differentiation is most often measured by using preferential looking tasks, in which infants are presented with different degrees of intermodal matches successively or simultaneously. Infants’ preferential looking time to one over the other presentation is interpreted as a detection of the difference in the degrees of intermodal matches between presentations, and as a preference for one over the other presentation. The term contingency detection was coined to describe this differentiation (Bahrick & Watson, 1985; Gergely & Watson, 1996, 1999). However, as further elucidated in the next sections, I will refrain from using “contingency detection” in order to avoid confusion regarding the term contingency.

1.2 Sources and Nature of Perfect Matches, Less-than-perfect Matches and No Matches in Intermodal Match Differentiation

1.2.1 Vote for a Sensible Use of the Terms “imperfect contingent”, “less-than-perfect contingent”, and “noncontingent”

Researchers in the field of intermodal match differentiation agree about the nature of perfect intermodal matches. Perfect matches provide perfect temporal and spatial correlations between executed and resulting actions regardless of whether infants are presented with a
real-time video image of their own leg movements (Abravanel, Levan-Goldschmidt, & Stevenson, 1976; Bahrick & Watson, 1985; Hiraki, 2006; Killen & Užgiris, 1981; Nielsen, Simcock, & Jenkins, 2008; Osofsky, 1987; Zmyj, Hauf, & Striano, 2009), their own hand movements (Anisfeld, 1996), a mirror image of the self (Watson, 1967), or a mobile attached to the leg with a ribbon (Dixon et al., 2012; Field, Cohen, Garcia, & Greenberg, 1984). In the various different methods and stimuli, the presentation of perfect spatial and temporal matches is most commonly and consistently called “contingent”.

However, there is less agreement about the nature of other-than-perfect matches. The terms “imperfect contingent” (e.g., Gergely & Watson, 1999), “semi-contingent” (Hiraki, 2006), “less-than-perfect contingent” (e.g., Gergely & Watson, 1999; Osofsky, 1987), and “noncontingent/non-contingent” (e.g., Bahrick & Watson, 1985) have all been used for presentations of less-than-perfect as well as no intermodal matches. In order to avoid further confusion, some researchers decided to use alternative terms, for example “reversed view” (Abravanel et al., 1976), “observer’s view” (Abravanel et al., 1976), or “delayed view” (Hiraki, 2006; Killen & Užgiris, 1981; Osofsky, 1987) to specify the nature of the presentations of less-than-perfect and no intermodal matches.

I suggest avoiding the use of the terms “contingent”, “imperfect contingent”, “less-than-perfect contingent”, and “noncontingent” in future studies. In order to disentangle the different natures of the different degrees of intermodal matches, I will present three main characteristics, which lead to a schema. This helps to assort and term the different degrees of matches in a new manner.

1.2.2 Sources of Less-than-perfect Matches: Contingency, Congruency, Timing

Imagine an infant who is observing his or her own hand movements. The stimulus (i.e., the motor signal of that hand movement) perfectly predicts a response. There is a perfect match between stimulus and response: contingency, congruency, and timing are perfectly related. Less-than-perfect matches originate from a variation in at least one of these sources,
and no match emerges if the response matches none of the three sources.

1.2.2.1 Contingency

In a mother-infant interaction, for example, the infant waves his or her arm and the mother might wave back in a mirror-like fashion. If the mother manages to wave back every time the infant waves, there is a perfect contingency between stimulus (i.e., the infant’s proprioception) and response (i.e., visual feedback of the parent waving). But there might be a situation in which the mother sometimes—but not always—waves back. In this case, the stimulus does not perfectly predict the response; contingency is not given.

1.2.2.2 Congruency

When contingency is given, there still might be a variation in the congruency between stimulus and response. For example, even if the mother manages to wave back every time the infant waves (contingency given), it might still be the case that her waving differs in terms of her posture, the movement amplitude, and velocity (congruency not given). Another example of an incongruent but contingent presentation is a left-right reversed real-time video presentation of the infant’s own hand movements. Every time the infant moves his or her own hand, the movement can be seen on the real-time video (contingency given). However, due to the left-right reversed presentation of the video, the direction of the performed movement and the direction of the movement observed on the video presentation differ from each other.

In the case of the mother-infant interaction, it might also be that no contingency in the mother’s waving is given. In this case too, congruency between stimulus and response cannot be achieved. Contingency is therefore a prerequisite for congruency.

1.2.2.3 Timing

Finally, it might be that the mother waves back in a mirror-like fashion every time the infant waves, and additionally that her imitation is very accurate regarding the congruency of her movements. But still, the temporal match between stimulus and response might be not
perfect. The mother’s response is temporally delayed compared to the infant’s proprioception, e.g., due to her processing and reaction time.

The magnitude of the latency between stimulus and response influences the infant’s ability to reference the response to the stimulus. This raises the question of how large the greatest possible latency between stimulus and response can be while still being conceived as temporally matched. The limit for detecting an association between stimulus and response probably lies at around 3-7 seconds (Millar & Watson, 1979; Watson, 1967).

Figure 1. Schema on the nature of perfect matches and less-than-perfect-matches in intermodal match differentiation tasks. Note that for reasons of simplicity, the schema does not depict the no-match category, in which contingency and congruency are given but the latency between stimulus and response exceeds 7 s.

Figure 1 presents the three characteristics contingency, congruency, and timing in a schema that helps to identify different degrees of matches. Three different categories of matches resulting from variations in contingency, congruency, and timing can be identified: perfect match, less-than-perfect match, and no match.
1.2.3 Three Categories: Perfect Match, Less-than-perfect Match, No Match

Perfect matches between stimulus and response can only be perceived if contingency and congruency as well as perfect timing are given. This perfect match is specified by actions of the self and can be perceived in many presentation types (e.g., a mirror view of one’s own face Abravanel et al., 1976; or a real-time view of one’s own leg movements; Abravanel et al., 1976; Bahrick & Watson, 1985; Hiraki, 2006; Killen & Užgiris, 1981; Osofsky, 1987; Zmyj et al., 2009; see Table 1 for more).

Less-than-perfect matches between stimulus and response can be perceived if at least one of the sources (contingency, congruency, and timing) is given within the feedback. For example, it might be that contingency and congruency are given, whereas timing is not. This is the case, for instance, in the presentation of an egocentric (as if the infant was looking down at his or her own legs) but delayed view of one’s own leg movements (Hiraki, 2006; Zmyj et al., 2009). Theoretically, it might also be the case that only one of the three sources allows the reference between stimulus and response. This can occur, for instance, when contingency is given in the feedback, whereas congruency and timing are not. However, to the best of my knowledge, no study to date has used the presentation of less-than-perfect feedback originating from only one of the three sources that allows the reference between stimulus and response.

No match between stimulus and response can be perceived if none of the three sources (contingency, congruency, and timing) is given within the feedback. This occurs, for instance, in the presentation of a previously recorded video of a peer’s face (Bahrick, Moss, & Fadil, 1996) or a peer’s leg movements (Bahrick & Watson, 1985).

Additionally, while contingency and congruency might be given within the feedback, the magnitude of the latency between stimulus and response might be so large that it exceeds the limit for detecting the association between stimulus and response. This limit probably lies at around 3-7 seconds (Millar & Watson, 1979; Watson, 1967). Therefore, the presentation of
an egocentric (as if the infant was looking down at his or her own legs) but delayed view of one’s own leg movements might be perceived as no match if the delay exceeds the limit of 7 seconds (Bahrick & Watson, 1985; Killen & Užgiris, 1981; Osofsky, 1987). Please note that this category is not depicted in figure 1.

Table 1

<table>
<thead>
<tr>
<th>Sources of different degrees of intermodal matches</th>
<th>Intermodal match differentiation tasks: Forms of visual-proprioceptive presentations and citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓ Contingency</td>
<td>- Mirror view of the self (Watson, 1967)</td>
</tr>
<tr>
<td>✓ Congruency</td>
<td>- Real-time, egocentric view of own face or of own leg movements (video, Abravanel et al., 1976)</td>
</tr>
<tr>
<td>✓ Timing</td>
<td>- Real-time, egocentric view of own leg movements (video, Abravanel et al., 1976; Bahrick &amp; Watson, 1985; Hiraki, 2006; Killen &amp; Užgiris, 1981; Osofsky, 1987; Zmyj et al., 2009)</td>
</tr>
<tr>
<td></td>
<td>- Real-time, egocentric view of own arm and hand movements (video, Anisfeld, 1996)</td>
</tr>
<tr>
<td></td>
<td>- Real-time, left-right reversed observer’s view of own leg movements (video, Abravanel et al., 1976)</td>
</tr>
<tr>
<td></td>
<td>- Point light display of own real-time leg movements (video, Heimann, Nelson, &amp; Schaller, 1989)</td>
</tr>
</tbody>
</table>
Table 1 (cont.)

<table>
<thead>
<tr>
<th>Sources of different degrees of intermodal matches</th>
<th>Intermodal match differentiation tasks: Forms of visual-proprioceptive presentations and citations</th>
</tr>
</thead>
</table>

**Less-than-perfect match**

- Delayed, egocentric view of own leg movements (video, delay 2 s, Hiraki, 2006; delay 3 s, Zmyj et al., 2009)
- Real-time view of mobile attached to own legs (video, Dixon et al., 2012)
- Direct view of mobile which turns when activated via foot kicks (Field et al., 1984) or via pillow pressure (Watson & Ramey, 1972)
- Real-time, left-right reversed view of own leg movements (video, Abravanel et al., 1976; Nielsen et al., 2008)
- Real-time, observer’s view of own leg movements (video, Abravanel et al., 1976)
- Real-time, left-right reversed observer’s view of point light display of own real-time leg movements (video, Heimann et al., 1989)

None

None
Table 1 (cont.)

<table>
<thead>
<tr>
<th>Sources of different degrees of intermodal matches</th>
<th>Intermodal match differentiation tasks: Forms of visual-proprioceptive presentations and citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>× Contingency</td>
<td>- Previously recorded video of own face (moving and still, delay 10 min, Bahrick et al., 1996)</td>
</tr>
<tr>
<td>× Congruency</td>
<td>- Peer face-to-face interaction (Watson, 1967)</td>
</tr>
<tr>
<td>× Timing</td>
<td>- Previously recorded video of peer’s leg movements (Bahrick &amp; Watson, 1985)</td>
</tr>
<tr>
<td></td>
<td>- Previously recorded video of point light display of peer’s leg movements (Heimann et al., 1989)</td>
</tr>
<tr>
<td></td>
<td>- Direct view of mobile, which turns when activated by experimenter (Field et al., 1984)</td>
</tr>
<tr>
<td></td>
<td>- Direct view of mobile, which turns every 3-4 s unrelated to pillow pressure (Watson &amp; Ramey, 1972)</td>
</tr>
<tr>
<td></td>
<td>- Direct view of mobile, which does not turn (Watson &amp; Ramey, 1972)</td>
</tr>
<tr>
<td></td>
<td>- Previously recorded video of observer view of point light display of peer’s leg movements (Heimann et al., 1989)</td>
</tr>
<tr>
<td></td>
<td>- Previously recorded video of mobile attached to peer’s legs (Dixon et al., 2012)</td>
</tr>
<tr>
<td></td>
<td>- Previously recorded video of peer’s hand movements (Anisfeld, 1996)</td>
</tr>
<tr>
<td></td>
<td>- Previously recorded, left-right reversed video of peer’s hand movements (Anisfeld, 1996)</td>
</tr>
<tr>
<td></td>
<td>- Previously recorded video of peer’s hand movements filmed from underneath (Anisfeld, 1996)</td>
</tr>
<tr>
<td></td>
<td>- Previously recorded video of peer’s face (moving and still, Bahrick et al., 1996)</td>
</tr>
<tr>
<td>✓ Contingency</td>
<td>- Delayed, egocentric view of own leg movements (video, delay 10 min., Bahrick &amp; Watson, 1985; delay 7.5 s, Killen &amp; Užgiris, 1981; Osofsky, 1987)</td>
</tr>
<tr>
<td>✓ Congruency</td>
<td></td>
</tr>
<tr>
<td>× Timing (more than 7 s delay)</td>
<td></td>
</tr>
</tbody>
</table>

1.3 Types of Intermodal Match Differentiation Tasks

The integration of information arising from multi-perceptual systems has been assessed in different types of intermodal match differentiation tasks. The intermodal match differentiation tasks presented in Table 1 assessed infants’ differentiation of different degrees of visual-proprioceptive matches. However, the assessment of the integration of information from sources other than vision and proprioception is possible as well. In the following
sections, I will briefly describe which other multi-perceptual systems have been assessed. As an exhaustive presentation of all pairings of modalities and task types is beyond the scope of the current dissertation, I will describe visual-proprioceptive match differentiation tasks. Additionally, I will describe visual-tactile and visual-auditory match differentiation tasks. Compared to visual-proprioceptive match differentiation tasks, visual-tactile and visual-auditory match differentiation tasks do not use infants’ self-generated behavior in the presentations. Thus, the pairings of these modalities can be seen as special cases of intermodal match differentiation tasks.

1.3.1 Visual-Proprioceptive Matches

Visual-proprioceptive match differentiation tasks are easy to implement as they require the manipulation of only one of the sources (contingency, congruency, timing; see section 1.2.2) to achieve the presentation of a less-than-perfect intermodal match. To date, no study has manipulated more than one of the three sources in order to achieve the presentation of a less-than-perfect intermodal match in visual-proprioceptive match differentiation tasks. Most visual-proprioceptive match differentiation tasks use and manipulate self-generated behavior of the infants to present different degrees of intermodal matches. This is not the case in visual-tactile and visual-auditory match differentiation tasks, as I will elucidate in the next sections.

1.3.2 Visual-Tactile Matches

In a visual-tactile match differentiation task, infants are presented with two different videos simultaneously (e.g., two video presentations of the stroking of one of two life-like puppet’s legs in different patterns; Zmyj, Jank, Schütz-Bosbach, & Daum, 2011). At the same time, the corresponding leg of the infant is stroked synchronously with the stroking pattern on one of the video presentations and asynchronously with the stroking pattern on the other video presentation (Zmyj et al., 2011). Infants’ looking time to the different presentations is measured in order to assess their differentiation between different degrees of visual-tactile matches. Although other task types have also been used to assess infants’ visual-tactile match
differentiation (e.g., Filippetti, Johnson, Lloyd-Fox, Dragovic, & Farroni, 2013; Filippetti, Orioli, Johnson, & Farroni, 2015), it is beyond the scope of the current dissertation to elucidate them all.

Visual-proprioceptive match differentiation tasks most commonly present self-generated behavior; thus, infants have control over the actions in at least one view in most of the presentation types (except for the presentations in which previously recorded behaviors of peers are used). Infants can actively explore the difference between, for example, the real-time and the delayed or the peer view (e.g., Bahrick & Watson, 1985; Hiraki, 2006; Osofsky, 1987; Zmyj et al., 2009). In contrast, in visual-tactile match differentiation tasks, such control over the actions is not given in any of the views.

1.3.3 Visual-Auditory Matches

In a visual-auditory match differentiation task, infants are presented with two different videos simultaneously (e.g., a video of a woman playing peek-a-boo and a video of a hand playing a tambourine; Spelke, 1976). At the same time, a soundtrack plays. The soundtrack corresponds to one video presentation but does not correspond to the other (e.g., the rhythm of the tambourine corresponds to the video of the hand playing the tambourine but not to the video of the woman playing peek-a-boo; Spelke, 1976). Infants’ looking time to the different presentations is measured in order to assess their differentiation between different degrees of visual-auditory matches. Again, although other tasks have also been used to assess infants’ visual-auditory match differentiation (e.g., Bahrick, 1988; Slater, Quinn, Brown, & Hayes, 1999), it is beyond the scope of the current dissertation to elucidate them all here.

Like in visual-tactile match differentiation tasks, and in contrast to visual-proprioceptive match differentiation tasks, infants have no control over the presentations in visual-auditory match differentiation tasks. The lack of infants’ control over the presentations in other modality pairings as opposed to the pairing of vision and proprioception makes a comparison difficult, which is why I chose not to include such tasks in the schema presented in sections
1.4 Intermodal Match Differentiation vs. Intermodal Match Preference

Intermodal match differentiation is most often measured by using preferential looking tasks. The studies included in section 1.2.3 presented infants with different levels of intermodal matches successively or simultaneously. Infants’ preferential looking time to one over the other presentation was interpreted as a differentiation between the different levels of intermodal matches in terms of a differentiation between presentations, and as a preference for one over the other presentation. Thus, what infants actually perceive is only construed by assessing and analyzing the looking times to the presentations. Comparisons of total or of mean looking times to each of the views constitute rather inadequate measures for this task, especially when the views are presented successively instead of simultaneously. Looking time measures in such tasks have to take into account the total looking time to both views, because the looking time to one view is not independent of the looking time to the other. In particular, total and mean looking time to the second presentation in successive presentations can be influenced by decreasing attention and by learning experiences from the first presentation.

Most researchers stick to the use of proportional looking measures. In order to calculate the proportional looking time to view 1, the looking time to view 1 is divided by the total looking time to both views. The possible values of the proportion of looking time range from 0 to 1. A proportional looking time value of .5 represents a looking behavior of 50-50% to both views. A proportional looking time value of 0 represents no looking (0%) to view 1 and 100% looking to view 2. The proportion of looking time as well as the difference score are used to indicate infants’ preference for one over the other view (intermodal match preference).

Another conception of a looking time measure is to calculate a difference score from the proportional looking times (Anisfeld, 1996). Schmuckler (1996) subtracted the proportion of looking time to view 1 from the proportion of looking time to view 2 and obtained values that
could range from -1 to 1. A difference score of 0 represents a looking behavior of 50-50% to both views. A difference score of -1 represents no looking to view 1 and 100% looking to view 2. The calculation of the difference score therefore adds no additional information to the proportion of looking time measure, and both the proportion of looking time and the difference score treat the two views as antagonistic entities.

I suggest adding another score to the existing conceptions: the differentiation score. To calculate the differentiation score, absolute values of the proportional looking time are subtracted from .5. The possible (absolute) values of the differentiation score range from 0 to .5. A differentiation score of 0 represents a looking behavior of 50-50% to both views. A differentiation score of .5 represents no looking to view 1 and 100% looking to view 2, or no looking to view 2 and 100% looking to view 1. A differentiation score value above 0 therefore represents the amount of interest in one of the two views, regardless of which one. This score might be used to measure the degree to which infants differentiate between the two views, regardless of which one the infants prefer (intermodal match differentiation). In contrast, the proportion of looking time as well as the difference score can be used to measure and indicate the preference for either one or the other view (intermodal match preference).

1.5 Relationship of Intermodal Match Differentiation and Self-Recognition

The ability to recognize the self in reflecting surfaces is an important milestone in infants’ development. When tested with the rouge test, around half of children were able to recognize themselves in a mirror at 18 months (Amsterdam, 1972; Nielsen & Dissanayake, 2004; Schulman & Kaplowitz, 1977). During this test, the child’s face is surreptitiously marked with a spot of rouge and the child is subsequently presented with a mirror. The child is able to recognize him/herself if he or she notices the spot in the mirror and reaches towards the corresponding spot on his or her own face.

Several studies assessed infants’ recognition of different body parts using different reflecting surfaces (Amsterdam, 1972; Nielsen, Dissanayake, & Kashima, 2003; Nielsen,

Within the first theory, it has been suggested that infants first recognize that the mirror image of themselves provides “something unique” (Rochat, 2003, p. 720), which is not provided, for example, by images of others. This uniqueness is caused by the experience of the perfect match between vision and proprioception of self-performed movements in front of a reflecting surface. Accordingly, the perfect intermodal match when observing self-performed movements in front of a mirror has been identified as one key component of the mirror image (termed as “contingency of the mirror image on the infant’s own activity”; Papoušek & Papoušek, 1974, p. 150) together with other components such as eye-to-eye contact. In a next step, infants become able to systematically test the unique relation between visual and proprioceptive feedback in front of the mirror, before they become able to
recognize themselves in the mirror in a further step (Rochat, 2003). Hence, infants’ experience with different degrees of intermodal matches is assumed to pave the way for later self-recognition. This notion is supported by the finding that infants’ movement-testing behavior in front of a mirror (systematical test of the relation between visual and proprioceptive feedback) increased just before the infants showed mirror self-recognition for the first time (Bigelow, 1981). In line with this notion, a seminal study provided evidence that at 3 months, infants are able to differentiate their own face from that of a peer (Bahrick et al., 1996). It has been argued that this ability is “developed through prior visual experience with the mirror. […] The mirror allows the infant to become familiar with […] the visual-proprioceptive contingency that specifies self.” (Bahrick et al., 1996, p. 204).

Within the second theory, it has been suggested that infants’ ability to recognize themselves in the mirror emerges along with a rather sudden cognitive change in the second year of life. This cognitive change allows infants to build representations of objects and of themselves in their imagination that refer to real objects and their real self, respectively (Bischof-Köhler, 1991, 2012). It has been argued that this cognitive change constitutes a developmental burst which arises independently of the experience of intermodal matches.

So far, no longitudinal study has assessed whether infants’ intermodal match differentiation is associated with their later self-recognition and thus provided evidence for either the first or second theory. Study 1 therefore aims at assessing the relationship between infants’ differentiation of different degrees of intermodal matches and their later self-recognition using a longitudinal study design. To this aim, infants’ ability to differentiate between different levels of intermodal matches at 6 and 9 months and their ability to recognize themselves in a mirror at 18 months and on a video presentation at 26 months are measured.

Please note that the terms intermodal match preference, intermodal match differentiation, and intermodal match differentiation task are not used throughout the
manuscript of study 1; instead, the terms contingency preference, contingency detection, and contingency task are used. This inconsistency arose from the fact that the schema on the nature of perfect, less-than-perfect and no matches in intermodal match differentiation tasks (see section 1.2.3) was developed later than the manuscript.
No evidence for infants’ contingency detection as a precursor of toddlers’ self-recognition

Sarah Klein-Radukic and Norbert Zmyj

TU Dortmund University, Dortmund, Germany

Submitted for publication in Developmental Psychology

The authors would like to thank Marie-Christin Frerich, Irene Gettmann, Sina Hamester, Tessa Heinrich, Hannah Hermschulte, Nadja Herten, Sina Hulten, Stefanie Kleinschmidt, Nathalie Marcinkowski, Milena Meyers, Erich Molz, Susanne Röttgers, and Paula Siegmann for their assistance with the data collection and data analysis as well as the infants and parents who participated in this study.

Correspondence concerning this article should be addressed to Sarah Klein-Radukic, Institute of Psychology, TU Dortmund University, Emil-Figge-Str. 50, 44227 Dortmund, Germany, Phone: +49 231 755 5243, E-mail: sarah.klein-radukic@tu-dortmund.de

This research was supported by the Deutsche Forschungsgemeinschaft (ZM54/2-1).

Abstract

The aim of the present study was to assess the relationship between infants’ contingency preference and detection in the first year of life and toddlers’ mirror and video self-
recognition in the second and third year of life in a longitudinal study \((N = 113)\). Infants’ preference for a noncontingent over a perfect contingent view (contingency preference) and their differentiation between the two views – indicated by longer looking times to either view – (contingency detection) were assessed in two contingency tasks at 6 and 9 months. A mirror-face-recognition task and a mirror-leg-recognition task were conducted at 18 months. A video-face-recognition task and a mirror-leg-recognition task were conducted at 26 months. Results revealed no predictive relationships between infants’ contingency preference and contingency detection in the first year of life and their ability to recognize themselves in a mirror or on a video monitor in the second and third year of life. This finding supports the notion that self-recognition emerges independently from the experience of contingencies (Bischof-Köhler, 1991, 2012; Suddendorf & Whiten, 2001). Thus, a representation of the self seems to rely on more than a specific developmental pathway leading from contingency preference and detection to self-recognition.

**Keywords**: contingency, self-recognition, infancy, cognitive developmental status

Self-awareness is crucial to human existence. Mirror self-recognition in the second year of life has been identified as a hallmark of the development to adult-like self-awareness (Rochat, 2003). A developmental trajectory in the development of self-awareness has been suggested, in which infants first recognize that there is “something unique about the experience” (Rochat, 2003, p. 720) of self-performed movements when they look in the mirror: They perceive the perfect match between visual and proprioceptive feedback. Next, infants become “capable of systematically exploring the intermodal link” (Rochat, 2003, p. 720) between visual and proprioceptive feedback when they look in the mirror. Subsequently, toddlers show mirror self-recognition when they identify their mirror reflection with their own body (Rochat, 2003), suggestive of a linear increase in the self-awareness capacities.
Other theories on the development of self-awareness suggest that mirror self-recognition emerges due to a cognitive change in the second year of life. Infants become able to form representations that refer to real objects and these representations can be manipulated in the imagination (Bischof-Köhler, 1991, 2012). Thus, this cognitive change enables infants to represent not only other objects but also the self. The ability to represent the self in the imagination constitutes a developmental burst, for which experiences of intermodal matches merely provide the material. Conversely, the experience of intermodal matches is not a precursor of the ability to represent the self in the imagination (Bischof-Köhler, 1991). Accordingly, mirror self-recognition is assumed to emerge independently from the experience of intermodal matches. In line with this account, it has been argued that mirror self-recognition “reflects the capacity to generate and compare multiple models of the same thing” (Suddendorf & Butler, 2013, p. 122). At around 2 years of age, infants become able to represent, for example, a banana as a banana and as a telephone at the same time (Suddendorf & Whiten, 2001), that is to generate secondary representations of situations and objects in mind. The “ability to consider a mental model of a situation not currently perceived” (Suddendorf & Whiten, 2001, p. 630) enables the infant to pretend that the banana is a telephone. In the case of mirror self-recognition, infants are able to compare the primary representation of the direct perception of the own body to the secondary representation of the self in the mirror (Suddendorf & Whiten, 2001).

The question of how the developmental milestones – the experience of intermodal links, for example between visual and proprioceptive feedback (labeled as contingency detection) and self-recognition – are related to each other, is uncharted territory. The aim of the current study is to address this question in a longitudinal study.

**Contingency**

Infants experience perfect contingency when they move their own body parts and compare the corresponding efferent signal with the afferent feedback. When infants move
their leg, for example, the efferent motor signal of the leg movement perfectly predicts the afferent proprioceptive input of the leg movement. Less-than-perfect contingency is characterized by a variation in the correlation between afferent and efferent information, for instance when infants observe their parents imitating them. In contrast, noncontingent behavior is characterized by no correlation between afferent and efferent information, for instance when parents perform everyday actions without any relation to the infants’ behavior. Hence, perfect contingent feedback arises from self-generated actions, whereas less-than-perfect contingent and noncontingent feedback arises from other-generated actions (Bahrick & Watson, 1985).

Infants’ contingency detection is commonly tested by means of preferential looking. In a seminal study (Bahrick & Watson, 1985), infants were presented with two monitors simultaneously. One monitor provided perfect contingent feedback (i.e., a real-time video image of infants’ own leg movements) and the other monitor provided noncontingent feedback (e.g., a previously recorded video image of a peer’s leg movements; Bahrick, Moss, & Fadil, 1996; Bahrick & Watson, 1985; Schmuckler, 1996; Schmuckler & Fairhall, 2001). Hence, infants controlled one display and did not control the other display. Five-month-olds, but not 3-month-olds, looked longer at the noncontingent view than at the perfect contingent view, indicating an ability to differentiate between the two views. By using a similar paradigm but implementing less-than-perfect contingent feedback (a delayed video image of infants’ own leg movements) instead of the noncontingent feedback of a previously recorded peer, researchers are able to control the video feedback for the congruency of the movements (i.e., variation in the amplitude and velocity of the movement). Seven-month-olds – but not 5-month-olds or younger infants – looked longer at the less-than-perfect contingent feedback (up to 3-s delayed video image) than at the perfect contingent feedback (real-time video image; Hiraki, 2006; Rochat & Striano, 2000; Zmyj, Hauf, & Striano, 2009). Infants’ preference for one of the two views in the preferential looking task is interpreted as the ability
to differentiate between the views. Accordingly, at around 6 months of age, infants might be in a transitional stage in terms of differentiating between contingent feedback and less-than-perfect contingent or noncontingent feedback in a visual-proprioceptive contingency task.

Indeed, in two previous studies (Klein-Radukic & Zmyj, 2015; Zmyj & Klein-Radukic, 2015), 6-month-olds showed no preference for perfect contingent over noncontingent feedback of their own leg movements or vice versa (a delay of 7.5 s was used, which is probably not perceived as contingent; Watson, 1967).

Infants’ contingency detection can also be tested when they have no control over the displays in a preferential-looking task. Similar to the visual-proprioceptive contingency task, infants were presented with two video monitors simultaneously (Zmyj, Jank, Schütz-Bosbach, & Daum, 2011). Both video monitors showed videos of lifelike baby doll legs, being stroked in different patterns. At the same time and out of the infant’s view, the infant’s own corresponding leg was stroked synchronously with one of the video displays. This provided less-than-perfect contingent feedback, as infants’ leg movements were not displayed on the video feedback. The infant’s leg was stroked asynchronously with the other display, providing noncontingent feedback. Ten-month-olds, but not 7-month-olds, looked longer at the less-than-perfect contingent feedback (Zmyj et al., 2011) as opposed to a looking preference towards noncontingent feedback in visual-proprioceptive contingency tasks. Another study presented newborns with two successive videos of the stroking of another newborn’s cheek. At the same time and out of the newborn’s view, the newborn’s own corresponding cheek was stroked. The cheek was stroked either synchronously with the video, as if the newborn could see him/herself in a mirror, thus providing less-than-perfect feedback, or asynchronously with the video, providing noncontingent feedback. Newborns looked longer at the less-than-perfect contingent feedback as opposed to a looking preference towards a noncontingent feedback in this visual-tactile contingency task (Filippetti, Johnson, Lloyd-Fox, Dragovic, & Farroni, 2013). Thus, for visual-proprioceptive contingency tasks, infants develop a looking preference...
towards less-than-perfect contingent feedback rather than perfect contingent feedback, whereas in visual-tactile contingency tasks, infants prefer to look at less-than-perfect contingent feedback rather than noncontingent feedback. Less-than-perfect contingent feedback resembles feedback of a responsive other, as is the case, for example, when parents imitate their infant (Bahrick & Watson, 1985). Consequently, the development of a preference for less-than-perfect contingent feedback over noncontingent and perfect contingent feedback is assumed to orient infants towards responsive others in social interactions (Gergely & Watson, 1999).

**Self-Recognition**

When toddlers start to imagine the visual appearance of the self, they become able to identify themselves in reflecting surfaces (Rochat, 2003). This self-recognition can be assessed with the rouge test (Amsterdam, 1972). In this task, a dot of rouge or a sticker is surreptitiously applied to the child’s face. Subsequently, a mirror or a real-time video image of the child is presented. The child passes the self-recognition task if he/she reaches towards the mark on his/her face when presented with the mirror or real-time video image (Amsterdam, 1972; Nielsen, Suddendorf, & Slaughter, 2006).

**Mirror self-recognition.** Around half of infants were able to recognize their own face (marked with a spot on the cheek) in the mirror at 18 months of age (Amsterdam, 1972; Nielsen & Dissanayake, 2004; Schulman & Kaplowitz, 1977). In another study, infants were presented with a view of only their lower body parts, with one of their legs marked with a sticker. At 18 months of age, infants were also able to recognize their own legs in the mirror (Nielsen et al., 2006). These findings support the assumption that infants start to build a representation of their own body—not only of their face—at around 18 months of age.

**Video self-recognition.** When the mirror was replaced with a real-time video image, at 24 months of age, 35% of children were able to recognize their own face marked with a spot on the cheek in a real-time video image presented on a projector screen (Suddendorf,
Simcock, & Nielsen, 2007). Other researchers used a television monitor instead of a projector and showed that only some 15-16-month-olds recognized their own face in a real-time video image (Vyt, 2001). Some studies showed that infants recognize their own face in a real-time video image at around 21 to 24 months of age (Johnson, 1983; Miyazaki & Hiraki, 2006; Vyt, 2001). At around 24 months of age, children were also able to recognize their own legs from a real-time video image presented on a projector screen (Suddendorf et al., 2007).

Self-recognition from a real-time video image is therefore more difficult, as children pass the rouge test at a later age when a real-time video image is presented than when a mirror is used. This asynchrony may be caused by two main characteristics of the real-time video image compared to the mirror. First, while the mirror provides direct eye-to-eye contact, the real-time video image does not. Second, the resolution of the monitor and the rendering of the colors differ between video and mirror image. Accordingly, self-recognition emerges in the second year of life and subsequently develops further.

**Relationship between Contingency Detection and Self-Recognition**

To our knowledge, only one study has investigated the relationship between contingency detection at 9 months, contingency preference at 12 months, and self-recognition at 24 months (Kristen-Antonow, Sodian, Perst, & Licata, 2015). Contingency detection was measured via a variation of the Still-Face Task (Tronick, Als, Adamson, Wise, & Brazelton, 1978). At the beginning of the task, the experimenter interacted naturally with the infant. The experimenter then interrupted the interaction by adopting a neutral face in the face-to-face phase or by adopting a neutral face and turning away from the infant in the ignore phase. Infants’ reactions during the interruption phases (i.e., amount of smiling, gazing and reengagement behaviors towards the experimenter) were measured. Contingency preference was measured via a social mirroring task (Meltzoff, 1990). Two experimenters interacted with the infant and both experimenters reacted contingently to the infant with regard to timing. One experimenter imitated the infant’s behaviors while the other performed control actions on the
same activity level as the infant. Infants’ reactions to each of the experimenters (i.e., amount of smiling, gazing and testing behavior) were measured. Self-recognition was measured via a mirror self-recognition task (Asendorpf & Baudonnière, 1993). A dot of lipstick color was surreptitiously applied to the child’s face. Following this, the child was presented with a mirror. The child passed the mirror self-recognition task if he/she reached towards the mark on the face – and not the corresponding part in the mirror. This study revealed that contingency detection – but not contingency preference – predicted mirror self-recognition. In this seminal study by Kristen-Antonow et al. (2015), the infants’ awareness of contingency was tested in social interactions. This gives rise to the question whether infants’ awareness of contingency when viewing the self also predicts later self-recognition.

The Present Study

The aim of the present longitudinal study was to assess the relationship between contingency detection in the first year of life and mirror and video self-recognition in the second and third year of life. Therefore, infants were invited at 6, 9, 18, and 26 months of age. At 6 and 9 months, infants’ contingency preference and detection was assessed. At 6 months, a visual-proprioceptive contingency task (6-VPC) was conducted. We tested 6-month-olds because this age marks a transitional period in the development of contingency preference for differentiating between contingent feedback and less-than-perfect contingent or noncontingent feedback in the visual-proprioceptive contingency tasks used in this study (Hiraki, 2006; Rochat & Striano, 2000; Zmyj et al., 2009). Infants were presented with a real-time and a delayed video image. A delay of 7.5 s was used, which exceeds the maximum threshold of short-term memory in infants (Watson, 1967) and is therefore expected to be perceived as noncontingent. At 9 months, a visual-tactile contingency task (9-VTC) was conducted. We tested 9-month-olds because this age marks a transitional period in the development of contingency preference for differentiating between less-than-perfect contingent and noncontingent feedback in visual-tactile contingency tasks used in this study (Zmyj et al.,
At 18 and 26 months of age, self-recognition was assessed. At 18 months, a mirror-face-recognition task (18-MFR) and a mirror-leg-recognition task (18-MLR) were conducted. At 26 months, a video-face-recognition task (26-VFR) and the mirror-leg-recognition task (26-MLR) were conducted. The 26-MLR was conducted due to the 18-month-olds’ poor performance in the corresponding 18-MLR. We tested 18- and 26-month-olds, because this age marks a period in which some but not all infants are able to recognize themselves in the mirror and on a video monitor (for mirror self-recognition; Amsterdam, 1972; Nielsen & Dissanayake, 2004; Schulman & Kaplowitz, 1977, for video self-recognition; Johnson, 1983; Miyazaki & Hiraki, 2006; Suddendorf et al., 2007; Vyt, 2001). The Cognitive Scale of the Bayley Scales of Infant and Toddler Development was applied at each test phase to assess children’s cognitive developmental status with the aim of examining whether relationships between the performance at different test phases was independent of the children’s cognitive developmental status (BSID-III, Bayley, 2006).

Our first hypothesis was that infants who prefer the noncontingent feedback in the 6-VPC and 9-VTC are more likely to recognize themselves in the mirror in the 18-MFR, 18-MLR, 26-MLR and in a real-time video in the 26-VFR. Our second hypothesis was an alternative hypothesis to our first hypothesis. We hypothesized that infants who differentiate between different degrees of contingency in the 6-VPC and 9-VTC – indicated by longer looking times to either view – are more likely to recognize themselves in the mirror in the 18-MFR, 18-MLR, 26-MLR and in a real-time video in the 26-VFR. We introduced this hypothesis because we assumed that some infants might prefer the perfect contingent view, which also suggests that they differentiated between the two views. Our third hypothesis was that there is a relationship between the performance in the 18-MFR and the 18-MLR. Our fourth hypothesis was that infants who recognize themselves in the mirror in the 18-MFR and 18-MLR are more likely to recognize themselves on the monitor in the 26-VFR. The
relationships in the four hypotheses should remain after controlling for cognitive developmental status.

**Method**

**Participants**

Infants and their parents were recruited from a database of parents who had previously agreed to participate in child development studies. All families lived in one of three medium-sized cities in the Ruhr area. Hundred-thirteen infants participated in the study, although the sample size in later test phases was smaller due to dropout.¹ Ten additional infants took part in the tests at 6 months of age, but had to be excluded because it was not possible to invite them again to the subsequent tests at 9, 18, and 26 months of age.

| Table 1 | Number of Infants, Infants’ Sex, Infants’ Mean Age (Months;Days) with Standard Deviation in Parentheses (Days) and Age Range at 6, 9, 18, and 26 Months of Age Test Phases |
|---|---|---|---|---|
| 6 months of age | 9 months of age | 18 months of age | 26 months of age |
| N | 113 | 111 | 105 | 93 |
| Sex | 53 female | 53 female | 47 female | 41 female |
| Infants’ age M (SD) | 6:05 (0:07) | 9:10 (0:10) | 18:01 (0:05) | 26:12 (0:09) |

Parents were asked to make appointments at a time of the day when their children were likely to be awake and alert. All infants were Caucasian. The main native language of mothers and fathers was German (76.1% for mothers, 71.7% for fathers) and all parents spoke German

¹The sample size in the different tasks at 6, 9, 18, and 26 months varied due to dropout over the test phases. We do not provide exhaustive information on the sample sizes, but we provide the sample size for each analysis in the Results section. Detailed information on the sample sizes and the reasons for dropout in the different tasks in the test phases can be provided by the corresponding author.
fluently. Mothers were between 21 and 42 years old ($M = 32$ years, $SD = 4$) and fathers were between 23 and 49 years old ($M = 35$ years, $SD = 5$). Most of the parents had either a university degree (mothers 44.2%, fathers 47.8%) or university entrance-level qualifications (Abitur; mothers 20.4%, fathers 17.7%). Parents were asked not to intervene in the test phase at any time. All infants received a small gift and all parents received 5 Euros as an acknowledgement for their participation at the end of each test phase. Families who participated in all four test phases additionally received 20 Euros.

**Design**

The study had a longitudinal design and tested infants at 6, 9, 18, and 26 months of age.

**Apparatuses and Procedure**

All tests and tasks were conducted in the same laboratory at the University of Bochum with the same experimenter. The testing room was partitioned into two and surrounded by beige curtains. A fan produced a 30 dB white noise in order to produce a slight background noise during testing. We used the same cameras for all tasks at all test phases (Canon Legria HV40, HDV 1080).

**6 months of age.** The test at 6 months of age consisted of three tasks: a visual-propiroceptve contingency task (6-VPC), a parent-infant interaction phase, and the BSID-III (Bayley, 2006). The interaction phase was not part of the current research question and will therefore not be analyzed. The order of the tasks (6-VPC, parent-infant interaction, and BSID-III) was not randomized in order to keep effects of attrition constant across all infants.

**6-VPC.**

*Apparatus.* In the 6-VPC, the infant sat in a dark blue baby seat (Maxi-Cosi, Dorel-Industries Inc., Netherlands) in front of camera 1, which filmed the infant’s face (distance 130 cm). Camera 2 recorded the infant’s leg movements from above the baby seat (distance 130 cm). Two 20-inch displays (Dell, UltraSharp 2007FPb) were positioned on the left and right side of camera 1 and showed the video stimuli. The distance between the centers of the
displays was 80 cm.

Stimuli. In the 6-VPC, the video stimuli consisted of the image of the infant’s leg movements recorded by camera 2. The image was fed into a device (Delay Line, Ovation Systems, Milton Common, Great Britain) which allows short-term storage and adjustable delayed output of real-time video recording. One monitor showed a real-time video image of the infant’s legs, while the other monitor showed a delayed (7.5 s) video image of the infant’s legs. The video stimuli resembled an egocentric view as if the infant was looking down at his/her legs (see Figure 1). The side of the delayed display was randomized across all infants. The duration of the task was 4 min.

![Image](image-url)

Figure 1. Experimental setup in the visual-proprioceptive contingency task at 6 months of age.

Procedure. The parent was given white trousers and was asked to dress their infant in these trousers. The trousers had two black points painted on the upper side of each foot. The parent was asked to put their infant into the baby seat in front of the monitors. The parent sat
behind the infant and was allowed to stroke the infant’s head in order to soothe him or her during the task. The experimenter then left the testing area. A bell was rung behind the curtain at the height of camera 1 in order to draw the infant’s attention to the monitors. The monitors were then switched on for the duration of the task.

**Bayley Scales of Infant and Toddler Development (BSID-III).** Infants’ cognitive developmental status was assessed via the Cognitive Scale of the BSID-III (Bayley, 2006) following the guidelines in the administration and technical manual. During testing, the infant was seated on his or her parent’s lap at a table. The BSID-III provides standardized scores for infants’ cognitive developmental status resembling IQ scores ($M = 100, SD = 15$).

**9 months of age.** The test at 9 months of age consisted of three different tasks: a visual-tactile contingency task (9-VTC), a habituation and novelty-preference task, and the BSID-III (Bayley, 2006). The habituation and novelty-preference task not part of the current research question and will therefore not be analyzed. The order of the tasks (9-VTC, habituation and novelty-preference task, and BSID-III) was not randomized in order to keep effects of attrition constant across all infants.

**9-VTC.**

**Apparatus.** In the 9-VTC, the infant sat on a table (90 (length) x 90 (width) x 70 (height) cm) in a blue baby seat (Bumbo Floor Seat, Bumbo International Trust, Gauteng, South Africa) in front of camera 1, which filmed the infant’s face (distance 115 cm). Camera 2 recorded the infant’s leg movements from underneath the table through a cut-out opening. Two 20-inch displays (Dell, UltraSharp 2007FPb) were placed on boxes with a height of 85 cm on the left and right side of camera 1 and showed the video stimuli. The distance between the centers of the displays was 80 cm. The experimenter sat under the table and stroked the infant’s left leg through the cut-out opening in the table. Beige linen covered the sides of the table and the infant’s legs and prevented the infant from looking at his or her legs or at the experimenter. The duration of the task was 4.5 min.
Stimuli. In the 9-VTC, the video stimuli consisted of two previously recorded videos of lifelike baby doll legs as used in Zmyj et al.’s study (2011). The doll’s legs were filmed from above, and in both videos, the doll’s left leg was stroked. One monitor showed a synchronous video image of the stroking of the doll’s legs, while the other monitor showed an asynchronous video image of the stroking of the doll’s legs. The video stimuli resembled an egocentric view, as if the infant was looking down at his or her own legs (see Figure 2). In the synchronous video display, the displayed stroking matched the stroking of the infant’s leg delivered by the experimenter. In the asynchronous video display, the displayed stroking did not match the stroking of the infant’s leg. Intervals of 1 min of stroking were separated by 9 s of attention getters (i.e., a flashing sun) in order to direct the infant’s attention to the monitors. The side of the asynchronous display was randomized across all infants.

Procedure. The parent was provided with red socks and was asked to dress the infant in these socks. The socks matched the socks presented in the video stimuli. Afterwards, the parent was asked to put the infant into the baby seat in front of the monitors. The parent sat behind the infant and was asked to secure the position of the infant on the table during the
task. Additionally, the parent was allowed to stroke the infant’s head in order to soothe him or her during the task. A bell was rung behind the curtain at the height of camera 1 in order to draw the infant’s attention to the monitors. As soon as the infant looked in the direction of the monitors, the experimenter crouched under the table, from where she stroked the infant’s left leg as soon as the first minute of stroking on the video started.

**BSID-III.** The assessment with the BSID-III at 9 months of age was identical to that at 6 months of age.

**18 months of age.** The test at 18 months of age consisted of three different tasks: a mirror-face-recognition task (18-MFR), a mirror-leg-recognition task (18-MLR), and the BSID-III (Bayley, 2006). The order of the tasks was semi-randomized. The 18-MFR and the 18-MLR were conducted at the beginning of the testing and were conducted in a randomized order. The BSID-III assessment was always conducted at the end of the testing in order to keep effects of attrition constant across all infants.

**18-MFR.**

*Stimuli.* The mark in the 18-MFR consisted of purple lipstick (pure color lipstick 042, P2 Kosmetik GmbH, Vienna, Austria). The purple color was chosen in order to avoid toddlers’ associations with blood.

*Apparatus.* A white multipurpose barrier (Baby Vivo MA Trading GmbH & Co. KG, Nürnberg, Germany) separated the testing area from the rest of the room (see Figure 3). The barrier was used in order to keep the child in the testing area and near the mirror. The testing area was 2.30 m². A mirror (138 (height) x 68 (width) cm) was mounted on the wall on the left side in the testing area. A fabric roller shutter was mounted on the top of the mirror. The experimenter could time exactly when the toddler saw his or her mirror image by operating the roller shutter. Two video cameras filmed the child’s behavior in front of the mirror. The cameras were mounted outside of the testing area on the left and right side of the mirror. The distance between the left camera and the mirror was 150 cm (height 170 cm). The distance
between the right camera and the mirror was 80 cm (height 107 cm). The experimenter sat outside the testing area next to the right camera. The experimenter adjusted, when necessary, the camera angle in order to capture the child’s movements within the testing area. The parent was asked to stay with the child in the testing area during the task.

Figure 3. Experimental setup in the mirror-leg-recognition task at 18 and 26 months of age. For the mirror-face-recognition task at 18 months of age the mirror was completely uncovered and the chair was removed. The infants were allowed to explore the mirror freely.

Procedure. During a warm-up play phase in a play area, the experimenter asked the child “Uhh, what’s that?” while pointing at the child’s face. The experimenter then asked “Do you want me to clean that off?” The experimenter then took a previously prepared soft tissue and applied the mark to the child’s cheek. The experimenter put the soft tissue into the garbage can and continued to play with the child for at least one minute. During this time, the experimenter observed whether the child touched his or her cheek and noticed the mark. If the child did not touch the mark during this phase, the parent and the child were led to the testing area. If the child touched the mark before the mirror was uncovered, he or she was excluded.
from the analyses. If the child did not allow the experimenter to clean his or her face with the soft tissue, the parent was asked to apply the mark to their child’s cheek. Then, the experimenter, the child, and the parent went into the testing area together. When they entered the testing area, the fabric roller shutter covered the mirror. The experimenter presented the child with a puppet (Chou Chou 95138, Zapf Creation AG, Rödental, Germany) which had a mark on its cheek. The mark was the same color as used for the mark on the child’s cheek. The experimenter asked the child: “What’s the matter with the puppet’s face? That doesn’t belong there! Do you want to clean her face?” and handed a tissue to the child. The child was allowed to wipe the mark off the puppet’s cheek. When the child was done, the experimenter additionally wiped off the remainder of the mark and said: “Well done. Now the puppet is clean.” If the child did not clean the puppet’s face, the experimenter asked the child: “Do you want me to clean the puppet’s face?” and wiped the mark off puppet’s cheek. The experimenter placed the puppet in front of the mirror, and then left the testing area. The experimenter drew the child’s attention to the roller shutter and uncovered the mirror. The experimenter sat next to the camera on the right side of the mirror. The child could explore the mirror and the testing area for 2 min. If necessary, the experimenter drew the child’s attention to the mirror by pointing at the mirror, snapping her fingers in front of the mirror, and asking the child to look at the mirror ($M = 1$ time pointing at mirror ($SD = 1$), $M = 1$ time snapping fingers ($SD = 2$), $M = 2$ vocalizations of experimenter ($SD = 2$)).

18-MLR.

Stimuli. The mark in the 18-MLR was a 5-cm-diameter sticker made out of a holographic, self-adhesive film.

Apparatus. The 18-MLR was conducted in the same testing area as used in the 18-MFR. A wooden high chair (Stokke® Tripp Trapp®, Stokke AS, www.stokke.com) was placed in the testing area. The distance between the high chair and the mirror was 65 cm. A tray (Play Tray, 4mykid aps, Jyllinge, Denmark) which was mounted onto the high chair prevented the
child from looking at his or her legs when sitting in the high chair. An LED light (LEDstixx, Osram GmbH, Munich, Germany) was secured under the tray and illuminated the child’s legs during the task. One camera filmed the child in the high chair, and the other camera filmed the child’s legs.

*Procedure.* The experimenter asked the parent to take off the child’s shoes. The experimenter then led the child and the parent into the testing area. When they entered the testing area, the fabric roller shutter covered the mirror and the light under the tray was then switched on. The experimenter asked the parent to place their child in the high chair and to sit next to the high chair on the left side. The experimenter sat on the right side of the child and presented the child with the same puppet as used in the 18-MFR. The only difference was that instead of a mark on the cheek, the puppet had a small sticker (2 cm in diameter) on its right leg. The mark was made out of the same holographic self-adhesive film as used for the mark on the child’s knee. The experimenter asked the child: “What’s the matter with the puppet’s leg? That doesn’t belong there! Can you take it off, please?” The child could take the sticker off the puppet’s leg and hand it either to the experimenter or the parent. The experimenter then said: “Well done. Now the puppet looks nice.” If the child did not take the sticker off the puppet’s leg, the experimenter asked the child: “Do you want me to take it off?” and took the sticker off the puppet’s leg. The experimenter pointed to the camera mounted on the left side of the child and asked the child: “Have you seen the camera up there?” or played with the puppet and the child. When the child turned his/her head in order to look at the camera or the puppet, the experimenter placed the self-adhesive sticker onto the child’s knee. The experimenter interacted with the child for a short time in order to ensure that the child did not notice the placement of the mark. If the child noticed the mark before the mirror was uncovered, the child was excluded from the analyses. The puppet was placed out of the child’s sight, and afterwards, the lower part of the mirror (55 cm) was uncovered. The experimenter then sat down on the seat next to the camera on the right of the mirror. Sitting in
an upright position, the child was able to see the tray and his or her legs underneath the tray in the mirror. The child could explore the mirror image for 1 min. If necessary, the experimenter drew the child’s attention to the mirror by pointing at the mirror, snapping her fingers in front of the mirror, and asking the child to look at the mirror ($M = 3$ times pointing at mirror ($SD = 1$), $M = 2$ times snapping fingers ($SD = 2$), $M = 3$ vocalizations of experimenter ($SD = 2$)).

**BSID-III.** The assessment with the BSID-III at 18 months of age was identical to the assessment with the BSID-III at 6 and 9 months of age.

**26 months of age.** The test at 26 months of age consisted of three different tasks: a video-face-recognition task (26-VFR), the mirror-leg-recognition task (26-MLR), and the BSID-III (Bayley, 2006). The leg-recognition task was again conducted in front of a mirror instead of using a video monitor due to the low percentage of self-recognizers (6%) in the 18-MLR. The order of the tasks was semi-randomized. The 26-VFR and the 26-MLR were conducted first, in a randomized order. The BSID-III assessment was always conducted at the end of the testing in order to keep effects of attrition constant across all infants.

**26-VFR.**

*Stimuli.* The mark consisted of the same colored lipstick as used in the 18-MFR. The application of the mark was conducted in the warm-up play phase in the same way as in the test at 18 months.

*Apparatus.* The 26-VFR was conducted at a white table (75x75 cm, 74 cm high). A white plastic high chair (Antilop, IKEA Deutschland GmbH & Co. KG, Hofheim-Wallau, Germany) was placed on one side of the table (see Figure 4). A 20-inch monitor (Dell, UltraSharp 2007FPb) was placed on the opposite side of the table and presented a real-time close-up video image of the child’s face and upper body, which was filmed by a camera (height 130 cm) above the monitor. The distance between the monitor and the child’s face was 74 cm. The distance between the camera and the child’s face was 94 cm. The monitor was covered with a blue blanket.
Procedure. The mark was applied to the child’s cheek during a warm-up play phase in a play area in the laboratory. The experimenter asked the parent to take off the child’s shoes. The experimenter then led the child and the parent to the table, where the blanket was covering the monitor. The experimenter asked the parent to place the child in the high chair. The parent sat on a chair behind and to the left of the child. The experimenter sat on the right side of the child and presented the same puppet as used in the 18-MFR and 18-MLR to the child. The puppet had a mark on its cheek. The cleaning of the puppet’s cheek followed the same procedure as used in the 18-MFR. Afterwards, the puppet was placed out of the child’s sight and the experimenter said: “Let’s watch TV. Are you ready? Let’s go!” Then, the monitor was uncovered. The child was free to explore the video image for 1 min. The experimenter did not interact with the child during this phase. If necessary, the experimenter
drew the child’s attention to the monitor by pointing at the monitor, snapping her fingers in front of the monitor.

26-MLR.

Stimuli, Apparatus, Procedure. The stimuli, apparatuses and procedure were identical to those in the 18-MLR.

BSID-III. The assessment with the BSID-III at 26 months of age was identical to that at 6, 9 and 18 months of age, except that the parent was asked to choose whether the child sat in an infant seat or on the parent’s lap.

Coding

6-VPC. Based on the video recordings, an observer coded how long the infant looked at each monitor using the software Interact® (version 9.3.5, Mangold Software & Consulting GmbH, Arnstorf, Germany) for a total of 4 min. The looking time was analyzed for the number of consecutive complete minutes in which the infant looked at the monitors. For example, if testing was terminated during the third minute, the first two minutes were analyzed. For analyses of the contingency tasks, two indices using proportional looking time were calculated to analyze infants’ contingency detection. First, the contingency preference index measured infants’ preference for the noncontingent over the perfect contingent view (see formula 1). High scores on the contingency preference index indicate a preference for the noncontingent view.

\[
\text{Contingency preference index} = \frac{\text{looking time at noncontingent view}}{\text{looking time at noncontingent view} + \text{looking time at contingent view}} \tag{1}
\]

Second, the contingency detection index was calculated (see formula 2).

\[
\text{Contingency detection index} = 0.5 - \text{contingency preference index} \tag{2}
\]
High scores on the contingency detection index indicate a high differentiation in terms of a high looking time to either one of the two views. An independent second observer additionally coded the videos of 40 randomly chosen infants. Interrater reliability for the contingency preference index was high ($r = .99$, Intraclass Correlation Coefficient, ICC). Both observers were naïve regarding the side of the noncontingent view.

**9-VTC.** Based on the video recordings, an observer coded how long the infant looked at each monitor using the software Interact® (version 9.3.5, Mangold Software & Consulting GmbH, Arnstorf, Germany) for a total of 4 min; only the four 1-min intervals of stroking were coded for looking time. The inclusion criteria and the calculation of the contingency preference index and the contingency detection index were analogous to the 6-VPC. An independent second observer additionally coded the videos of 38 randomly chosen infants. Interrater reliability for the contingency preference index was high ($r = .99$, ICC). Both coders were naïve regarding the side of the asynchronous view.

**18-MFR, 18-MLR, 26-VFR, and 26-MLR.** Based on the video recordings, an observer coded different behaviors using the software Interact® (version 9.3.5, Mangold Software & Consulting GmbH, Arnstorf, Germany). The behaviors coded are depicted in Table 2.
Table 2

Minutes of Coding, Looking Behavior and Definition of Mark-Directed Behavior in the Mirror-Face-Recognition Task at 18 Months of Age (18-MFR), the Video-Face-Recognition Task at 26 Months of Age (26-VFR), and the Mirror-Leg-Recognition Task at 18 and 26 Months of Age (18-MLR and 26-MLR)

<table>
<thead>
<tr>
<th></th>
<th>18-MFR</th>
<th>26-VFR</th>
<th>18-MLR and 26-MLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total minutes of coding</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Looking behavior</td>
<td>total looking time at the mirror</td>
<td>total looking time at the monitor</td>
<td>total looking time at the mirror</td>
</tr>
<tr>
<td>Definition of mark-directed behavior</td>
<td>The infant touches the own cheek where the mark was applied/on the opposite side of the mark or points with the index fingers at the marked side/opposite side while the infant directs his/her attention to the mirror, the experimenter or the parent.</td>
<td>The infant touches the sticker either by - guiding the own hand along the body (between body and tray) down to the leg, - or by stretching out the leg in a way that it pokes out of the tray and then reaching around the tray with the hand, - or by standing up in the chair or drawing up the marked knee.</td>
<td></td>
</tr>
</tbody>
</table>

**BSID-III at 6, 9, 18, and 26 months of age.** The coding of the BSID-III at 6, 9, 18, and 26 months followed the guidelines in the administration and technical manual. Four independent second observers coded randomly chosen videotaped assessments of the BSID-III at 6, 9, 18, and 26 months. Interrater reliability for the BSID-III scores was high ($r = .864; n = 40$ at 6 months, $r = .933; n = 41$ at 9 months, $r = .947; n = 44$ at 18 months, and $r = .924; n = 30$ at 26 months, ICC).
Results

Looking Behavior, Contingency Preference, and Contingency Detection

Children’s average looking behavior in the 18-MFR, 18-MLR, 26-VFR, and 26-MLR is depicted in Table 3.

Table 3
Children’s Average Looking Time and Range of Looking Time to Mirror and Monitor in the 18-MFR, 18-MLR, 26-VFR, and 26-MLR

<table>
<thead>
<tr>
<th></th>
<th>18-MFR</th>
<th>18-MLR</th>
<th>26-VFR</th>
<th>26-MLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average looking time to mirror / monitor in s</td>
<td>20</td>
<td>18</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>Range</td>
<td>4 - 39</td>
<td>7 - 29</td>
<td>9 - 60</td>
<td>4 - 53</td>
</tr>
</tbody>
</table>

**6-VPC - Contingency preference.** Infants’ looking time at the noncontingent and at the perfect contingent view did not differ \( (M = 82.58 \text{ s}; SD = 29.38, M = 86.86 \text{ s}; SD = 28, \) respectively; \( t(108) = 1.062, p = .291, d = -0.15 \). Infants’ looking time at the left and at the right view differed \( (M = 92.09 \text{ s}; SD = 28.95, M = 77.36 \text{ s}; SD = 26.64, \) respectively; \( t(108) = -3.883, p < .001, d = 0.53 \). The mean contingency preference index was \( M = .48 (SD = 0.12) \) and ranged from .19 to .81. A single-sample \( t \)-test against the chance level of .50 revealed no differences between the mean contingency preference index and chance level, \( t(108) = -1.348, p = .18, d = -0.26. \)

**6-VPC - Contingency detection.** The mean contingency detection index in the 6-VPC was \( M = .09 (SD = 0.07). \)

**9-VTC - Contingency preference.** Infants’ looking time at the synchronous and at the asynchronous view did not differ \( (M = 60.37 \text{ s}; SD = 27.78, M = 61.02 \text{ s}; SD = 26.75, \)
respectively; \(t(101) = -1.192, p = .849, d = -0.02\). Infants’ looking time at the left and at the right view differed \((M = 65.44 \text{ s}; SD = 29.09, M = 55.93 \text{ s}; SD = 24.41, \text{ respectively}; t(101) = -2.914, p = .004, d = 0.35\). The mean contingency preference index was \(M = .51 (SD = 0.14)\) and ranged from .20 to .87. A single-sample \(t\)-test against the chance level of .50 revealed no differences between the mean contingency preference index and chance level, \(t(101) = -0.656, p = .513, d = -0.13\).

### 9-VTC - Contingency detection

The mean contingency detection index in the 9-VTC was \(M = .11 (SD = 0.08)\).

### Self-recognition

The number of children assessed and the percentages of children passing the self-recognition tasks at 18 and 26 months of age are depicted in Table 4.

### Table 4

**Number of Children Who Completed and Percentage of Children Who Recognized Themselves in the Self-Recognition Task at 18 and 26 Months of Age**

<table>
<thead>
<tr>
<th></th>
<th>18-MFR</th>
<th>18-MLR</th>
<th>26-VFR</th>
<th>26-MLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of completers of the task</td>
<td>98</td>
<td>84</td>
<td>80</td>
<td>75</td>
</tr>
<tr>
<td>Percentage of self-recognizers</td>
<td>31%</td>
<td>6%</td>
<td>40%</td>
<td>73%</td>
</tr>
</tbody>
</table>

A contingency table of the numbers of children who did or did not recognize themselves in the 18-MFR in relation to whether they did or did not recognize themselves in the 18-MLR is depicted in Table 5. There was no association between children’s self-recognition status in the 18-MFR and their self-recognition status in the 18-MLR, \(p = .149\), Fisher’s exact test. Due
to the low percentage of self-recognizers in the 18-MLR, the task was excluded from further analyses.

Table 5

*Contingency Table of the Numbers of Children (with Percentage in Parentheses) Who Did or Did Not Recognize Themselves in the Mirror-Face-Recognition Task in Relation to whether the Children Did or Did Not Recognize Themselves in the Mirror-Leg-Recognition Task at 18 Months of Age*

<table>
<thead>
<tr>
<th>18-MLR</th>
<th>Self-recognizer</th>
<th>Non self-recognizer</th>
</tr>
</thead>
<tbody>
<tr>
<td>18-MFR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-recognizer</td>
<td>3 (4%)</td>
<td>20 (26%)</td>
</tr>
<tr>
<td>Non self-recognizer</td>
<td>2 (2%)</td>
<td>53 (68%)</td>
</tr>
</tbody>
</table>

*Cognitive Developmental Status*

**BSID-III at 6, 9, 18, and 26 months of age.** The number of infants who completed the BSID-III, their mean cognitive developmental status scores with standard deviations and minimum and maximum scores in the Cognitive Scale of the BSID-III (Bayley, 2006) are given in Table 6. For further analyses, infants’ average scores on the Cognitive Scale of the BSID-III (Bayley, 2006) at 6, 9, 18, and 26 months are used.
Table 6

Number of Infants Who Completed the BSID-III, Mean Score, Standard Deviations, Minimum Score, and Maximum Score for the Cognitive Scale of the BSID-III at 6, 9, 18, 26 Months of Age and Average Score

<table>
<thead>
<tr>
<th></th>
<th>6 months</th>
<th>9 months</th>
<th>18 months</th>
<th>26 months</th>
<th>Average score</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>109</td>
<td>110</td>
<td>97</td>
<td>89</td>
<td>113</td>
</tr>
<tr>
<td>M (SD)</td>
<td>106.83 (8.60)</td>
<td>100.45 (9.71)</td>
<td>101.44 (11.09)</td>
<td>106.29 (11.14)</td>
<td>103.77</td>
</tr>
<tr>
<td>Min</td>
<td>75</td>
<td>75</td>
<td>85</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>Max</td>
<td>125</td>
<td>125</td>
<td>140</td>
<td>135</td>
<td>123.75</td>
</tr>
</tbody>
</table>

**BSID-III and Contingency.**

**6-VPC.** A correlation analysis (Pearson correlation) revealed no relation between the average BSID-III score and the contingency preference index in the 6-VPC, $r = .023, p = .814, n = 109$, or between infants’ average BSID-III score and the contingency detection index in the 6-VPC, $r = .175, p = .068, n = 109$.

**9-VTC.** A correlation analysis (Pearson correlation) revealed no relation between the average BSID-III score and the contingency preference index in the 9-VTC, $r = .065, p = .515, n = 102$, or between infants’ average BSID-III score and the contingency detection index in the 9-VTC, $r = -.042, p = .678, n = 102$.

**BSID-III and Self-Recognition.**

**18-MFR.** A correlation analysis (point-biserial correlation) revealed no relation between infants’ average BSID-III score and the self-recognition in the 18-MFR, $r = .102, p = .319, n = 98$.

**26-VFR and 26-MLR.** A correlation analysis (point-biserial correlation) revealed no relation between infants’ average BSID-III score and the self-recognition in the 26-VFR, $r =
.001, \( p = .996, n = 80 \), or between the average BSID-III score and the self-recognition in the 26-MLR, \( r = .137, p = .242, n = 75 \).

The Relation between Contingency Preference, Contingency Detection and Self-Recognition

The correlation matrix in Table 7 depicts the relations between the contingency preference indices and the contingency detection indices in the 6-VPC and 9-VTC, and the self-recognition in the 18-MFR, 18-MLR, 26-VFR, and 26-MLR.
Table 7

**Correlation Matrix Displaying the Relations Between Contingency Preference Indices, Contingency Detection Indices in 6-VPC and 9-VTC, and Self-Recognition in the 18-MFR, 26-VFR, and 26-MLR. The Lower Left Part Represents Partial Correlations After Controlling for the Average Score of the Cognitive Scale of the BSID-III at 6, 9, 18, 26 Months of Age (Bayley, 2006)**

<table>
<thead>
<tr>
<th></th>
<th>Contingency Preference Index</th>
<th>Contingency Detection Index</th>
<th>Self-Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6-VPC</td>
<td>9-VTC</td>
<td>6-VPC</td>
</tr>
<tr>
<td>Contingency preference index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-VPC</td>
<td>-</td>
<td>.047^a</td>
<td>-.039</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>106</td>
<td>97</td>
</tr>
<tr>
<td>9-VTC</td>
<td>.046</td>
<td>-</td>
<td>.080</td>
</tr>
<tr>
<td></td>
<td>97</td>
<td>97</td>
<td>99</td>
</tr>
<tr>
<td>Contingency detection index</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-VPC</td>
<td>-.035^a</td>
<td>.090^a</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>109</td>
<td>100</td>
<td>-</td>
</tr>
<tr>
<td>9-VTC</td>
<td>.096^a</td>
<td>.115^a</td>
<td>.122</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>102</td>
<td>97</td>
</tr>
<tr>
<td>Self-Recognition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-MFR</td>
<td>-.010</td>
<td>.050</td>
<td>.203*</td>
</tr>
<tr>
<td></td>
<td>92</td>
<td>85</td>
<td>92</td>
</tr>
<tr>
<td>26-VFR</td>
<td>-.041</td>
<td>-.045</td>
<td>.053</td>
</tr>
<tr>
<td></td>
<td>76</td>
<td>70</td>
<td>76</td>
</tr>
<tr>
<td>26-MLR</td>
<td>.144</td>
<td>-.162</td>
<td>.054</td>
</tr>
<tr>
<td></td>
<td>71</td>
<td>67</td>
<td>71</td>
</tr>
</tbody>
</table>

*p < .05; ^a Pearson correlation; ^b point-biserial correlation

Binary logistic regression analysis (inclusion method) was applied with self-recognition in the 18-MFR as dependent variable and contingency preference indices and contingency detection indices in the 6-VPC and 9-VTC as predictors. The average score on the BSID-III at 6, 9, 18, and 26 months of age was included as control variable. As shown in Table 8, there were no significant predictors of self-recognition in the 18-MFR.
Table 8

Results of Binary Logistic Regression Analysis of Mirror-Face-Recognition at 18 Months of Age

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>95% CI</th>
<th>Wald</th>
<th>SE</th>
<th>e^B</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average score for the Cognitive Scale of the BSID-III at 6, 9, 18, 26 months of age</td>
<td>0.03</td>
<td>[-0.06; 0.11]</td>
<td>0.463</td>
<td>0.04</td>
<td>1.03</td>
<td>.52</td>
</tr>
<tr>
<td>Contingency preference index 6-VPC</td>
<td>0.88</td>
<td>[-3.69; 6.89]</td>
<td>0.214</td>
<td>2.52</td>
<td>2.41</td>
<td>.66</td>
</tr>
<tr>
<td>Contingency detection index 6-VPC</td>
<td>5.41</td>
<td>[-3.24; 15.60]</td>
<td>2.735</td>
<td>4.31</td>
<td>222.67</td>
<td>.12</td>
</tr>
<tr>
<td>Contingency preference index 9-VTC</td>
<td>0.04</td>
<td>[-4.34; 4.51]</td>
<td>0</td>
<td>2.10</td>
<td>1.04</td>
<td>.99</td>
</tr>
<tr>
<td>Contingency detection index 9-VTC</td>
<td>2.97</td>
<td>[-4.28; 9.81]</td>
<td>0.913</td>
<td>3.85</td>
<td>19.44</td>
<td>.38</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.87</td>
<td>[-13.94; 3.55]</td>
<td>1.514</td>
<td>4.72</td>
<td>0.08</td>
<td>.24</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval. Pseudo R² = .057 (Cox & Snell), .082 (Nagelkerke), Model χ² (5) = 5.131, p = .40. n = 87. 95% bias corrected and accelerated standard errors, confidence intervals and p-values based on 1000 bootstrap samples.

Binary logistic regression analysis (inclusion method) was applied with self-recognition in the 26-VFR as dependent variable, and contingency preference indices and contingency detection indices in the 6-VPC and 9-VTC, and self-recognition in the 18-MFR as predictors. The average score on the BSID-III at 6, 9, 18, and 26 months of age was included as control variable. As shown in Table 9, there were no significant predictors of self-recognition in the 26-VFR.
Table 9

Results of Binary Logistic Regression Analysis of Video-Face-Recognition at 26 Months of Age

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$B$</th>
<th>95% CI</th>
<th>Wald</th>
<th>SE</th>
<th>$e^B$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average score for the Cognitive Scale of the BSID-III at 6, 9, 18, 26 months of age</td>
<td>-0.02</td>
<td>[-0.11; 0.04]</td>
<td>0.371</td>
<td>0.05</td>
<td>0.98</td>
<td>.55</td>
</tr>
<tr>
<td>Contingency preference index 6-VPC</td>
<td>-0.16</td>
<td>[-5.52; 6.79]</td>
<td>0.005</td>
<td>3.11</td>
<td>0.85</td>
<td>.95</td>
</tr>
<tr>
<td>Contingency detection index 6-VPC</td>
<td>0.90</td>
<td>[-9.32; 15.23]</td>
<td>0.058</td>
<td>5.48</td>
<td>2.47</td>
<td>.81</td>
</tr>
<tr>
<td>Contingency preference index 9-VTC</td>
<td>0.57</td>
<td>[-5.78; 7.01]</td>
<td>0.070</td>
<td>2.87</td>
<td>1.78</td>
<td>.79</td>
</tr>
<tr>
<td>Contingency detection index 9-VTC</td>
<td>-4.80</td>
<td>[-14.56; 2.92]</td>
<td>1.522</td>
<td>4.93</td>
<td>0.01</td>
<td>.22</td>
</tr>
<tr>
<td>Self-recognition in 18-MFR</td>
<td>-0.75</td>
<td>[-2.48; 0.65]</td>
<td>1.518</td>
<td>0.75</td>
<td>0.48</td>
<td>.23</td>
</tr>
<tr>
<td>Constant</td>
<td>2.65</td>
<td>[-9.88; 16.47]</td>
<td>0.392</td>
<td>5.11</td>
<td>14.08</td>
<td>.54</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval. Pseudo $R^2 = .044$ (Cox & Snell), .061 (Nagelkerke), Model $\chi^2 (6) = 3.004, p = .81$. $n = 66$. 95% bias corrected and accelerated standard errors, confidence intervals and $p$-values based on 1000 bootstrap samples.

Binary logistic regression analysis (inclusion method) with self-recognition in the 26-MLR as dependent variable, and contingency preference indices and contingency detection indices in the 6-VPC and 9-VTC, and self-recognition in the 18-MFR as predictors could not be applied. This was due to unreliable estimates for the predictors contingency preference index 6-VPC and contingency detection index 6-VPC, as indicated by wide confidence intervals and unreasonably high standard errors (Irala, Fernandez-Crehuet Navajas, & Serrano del Castillo, 1997).
The Relation between Self-Recognition at 18 and at 26 Months

Binary logistic regression analysis (inclusion method) was applied with self-recognition in the 26-VFR as dependent variable, and self-recognition in the 18-MFR as predictor. The average score on the BSID-III at 6, 9, 18, and 26 months of age was included as control variable. As shown in Table 10, there were no significant predictors of self-recognition in the 26-VFR.

Table 10

Results of Binary Logistic Regression Analysis of Video-Face-Recognition at 26 Months of Age

<table>
<thead>
<tr>
<th>Predictor</th>
<th>B</th>
<th>95% CI</th>
<th>Wald</th>
<th>SE</th>
<th>e^B</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average score for the Cognitive Scale of the BSID-III at 6, 9, 18, 26 months of age</td>
<td>-0.01</td>
<td>[-0.08; 0.08]</td>
<td>0.018</td>
<td>0.04</td>
<td>1.00</td>
<td>.90</td>
</tr>
<tr>
<td>Self-recognition in the 18-MFR</td>
<td>-0.72</td>
<td>[-1.76; 0.34]</td>
<td>2.038</td>
<td>0.52</td>
<td>0.49</td>
<td>.14</td>
</tr>
<tr>
<td>Constant</td>
<td>0.56</td>
<td>[-7.39; 8.08]</td>
<td>0.024</td>
<td>3.87</td>
<td>1.75</td>
<td>.88</td>
</tr>
</tbody>
</table>

Note. CI = confidence interval. Pseudo $R^2$ = .027 (Cox & Snell), .037 (Nagelkerke), Model $\chi^2$ (2) = 2.060, $p = .36$. $n = 74$. 95% bias corrected and accelerated standard errors, confidence intervals and $p$-values based on 1000 bootstrap samples.

Discussion

The aim of the current study was to examine the relationship between contingency preference and contingency detection at 6 and 9 months and mirror self-recognition at 18 months and video self-recognition at 26 months.

Contingency Preference and Contingency Detection do not Predict Self-Recognition

Our first hypothesis was that infants who prefer less-than-perfect contingency at 6 and 9
months are more likely to recognize themselves in the mirror at 18 and 26 months and in a real-time video at 26 months. The findings of the current study did not support this hypothesis.

Our second hypothesis was that infants who differentiate between different degrees of contingency in the contingency tasks at 6 and 9 months are more likely to recognize themselves in the mirror in the face- and leg-recognition tasks at 18 and 26 months and in a real-time video image at 26 months. The findings of the current study did not support this hypothesis. Overall, we could not find evidence that infants’ contingency detection early in life predicts their ability to reference the mirror image of their own face to the mental representation of their own face (Rochat, 2003).

The lack of relationships in the present study is in contrast to the finding that contingency detection at 9 months predicts self-recognition at 24 months (Kristen-Antonow et al., 2015). A difference between the two studies was that in Kristen-Antonow et al.’s study, contingency detection was measured via a variation of the Still-Face Task (Tronick et al., 1978). In their study, infants’ gazing during the still face phase – but not smiling and reengagement behavior during the still face phase – was a predictor of mirror self-recognition. This difference in the measurement of contingency detection between Kristen-Antonow et al.’s and our study suggests that infants’ expectations of social contingency are related to the development of a representation of the self (Bigelow, 2001), but infants’ expectations of contingency when viewing the self are not. The latter interpretation is limited by the positive relationship found between contingency detection at 6 months and mirror self-recognition at 18 months after controlling for cognitive developmental status in the present study. However, this correlation did not remain statistically significant in the regression analysis, which does not support the assumption of a relationship between contingency detection in a non-social situation and later self-recognition. Future studies should test the relative importance of
contingency detection in social and nonsocial situations in the first year of life for later self-recognition.

Our interpretation rests on the idea that the contingency tasks used in the current and in earlier studies (Klein-Radukic & Zmyj, 2015; Zmyj et al., 2009; Zmyj et al., 2011; Zmyj & Klein-Radukic, 2015) measure infants’ ability to perceive intermodal contingencies in real life. However, the visual-proprioceptive contingency task at 6 months allows infants to experience two different degrees of self-generated visual-proprioceptive feedback at the same time, which is unique to this task. In their everyday life, infants experience either perfect contingent visual-proprioceptive feedback arising from self-generated movements or less-than-perfect or noncontingent feedback arising from other-generated movements. The visual-proprioceptive contingency task used in this study provides infants with perfect contingent and noncontingent feedback, both arising from self-generated movements, which is novel for the infant. The task allows researchers to measure the current contingency preference and current contingency detection, but both can be biased by the novelty of the task.

**Performance in Mirror Face Recognition is not Associated with Performance in Mirror Leg Recognition at 18 Months**

Contrary to our third hypothesis, we found no association between performance in the face- and in the leg-recognition task in the mirror at 18 months. This finding indicates that recognition of the own legs was more difficult than recognition of the own face at 18 months of age and does not replicate a similar study by Nielsen et al. (2006). Although our mirror-face-recognition task was not more difficult than that used in Nielsen et al.’s study, our mirror-leg-recognition task was more difficult. We identified at least three differences in terms of procedure and coding of the mirror-leg-recognition task between Nielsen et al.’s study and the current study. First, the high chair used by Nielsen et al. (2006) in the mirror-leg-recognition task differed from the high chair we used. Thus, the space between the infants’ upper body and the plastic tray mounted on the high chairs were unequal, which
might have made the sticker more easily accessible for infants in Nielsen et al.’s study. However, this appears to be an unlikely explanation, as the few children who reached the sticker on the leg in the current study did so easily. Second, the experimenter in the study by Nielsen et al. “said ‘Look at that. Can you get it for me?’ while pointing toward the mirror” (Nielsen et al., 2006, p. 178) if the child was “uncertain as to whether they are allowed […] to reach for it” (Nielsen et al., 2006, p. 178). Children were classified as self-recognizers if they touched the sticker spontaneously without showing signs of uncertainty and consequently without the prompt of the experimenter. They were also classified as self-recognizers if they showed signs of uncertainty, were consequently asked to get the sticker, and reached for the sticker the experimenter’s prompt. In contrast, the current study only classified children as self-recognizers when they spontaneously reached towards the sticker without prompts from the experimenter. We applied this procedure in order to keep the experimenter’s instructions constant across all self-recognition tasks. Third, in the current study, the children had 60 s to reach for the sticker, whereas the children in Nielsen et al.’s task only had 30 s. However, this difference increased the task difficulty in Nielsen’s study. To sum up, differences between Nielsen et al.’s study and the current study in terms of the procedure and coding in the mirror-leg-recognition task might have contributed to a lower rate of passers in the current study compared to previous research (Nielsen et al., 2006).

**Self-Recognition at 18 Months does not Predict Self-Recognition at 26 Months**

Our fourth hypothesis expected a positive relationship between the self-recognition tasks at 18 months and video face recognition at 26 months. Mirror face recognition at 18 months did not predict video face recognition at 26 months, and mirror leg recognition at 18 months could not be included in this analysis due to the low percentage of passers. The lack of relationship can be explained by the different task characteristics. A mirror was used in the task at 18 months of age, whereas a real-time video image was used at 26 months of age. As children pass the rouge test conducted in front of a real-time video image later compared to
the rouge test conducted in front of a mirror, self-recognition from a real-time video image is more difficult. The greater difficulty of video self-recognition tasks compared to mirror self-recognition tasks may be caused by two main characteristics. First, while a mirror can provide direct eye-to-eye contact, a real-time video image cannot. However, 24-month-olds are less likely to pass a video-leg-recognition task than a mirror-leg-recognition task, both of which do not involve eye-to-eye contact (Suddendorf et al., 2007). Thus, this explanation is unlikely. Second, the image resolution and the rendering of the colors differ between mirror and video image. An undistorted and unstained mirror reflects objects and surroundings more realistically than a video image. Additionally, the observer’s movement in front of a mirror results in a change in the details of the reflected objects and surroundings visible to the observer. In contrast, the observer’s movement in front of a real-time video image does not result in a change in details visible to the observer, as the camera provides a steady image of the rendered objects and surroundings (Loveland, 1986). Thus, successful video self-recognition requires higher representational flexibility of infants’ representation of the self compared to successful mirror self-recognition (Barr, 2010). So far, no study has resolved the issue of differing resolutions, color renderings and reflected details between video and mirror images, although it has been suggested that video training can decrease the difficulties infants experience when presented with a video self-recognition task (Barr, 2010; Suddendorf et al., 2007). For reasons that remain unclear, infants have more difficulties to collate the representation of the self on the video with the representation of the self in their mind, than to collate the representation of the self in the mirror with the representation of the self in their mind. The lack of relationship between performance on video-face-recognition tasks and mirror-face-recognition tasks might also be attributed to differences in the processing of different modalities. Passing the mirror-face-recognition task therefore must not necessarily lead to passing the video-face-recognition task and vice versa.
Limitations

A limitation of the current study is the lack of information about the validity and reliability of the contingency and self-recognition tasks used, which is a common shortcoming in infancy research. To the best of our knowledge, no study has assessed the validity and reliability of the contingency tasks. Some studies investigated the retest reliability of the mirror self-recognition task, and there is evidence that some non-passers of the mirror self-recognition task are probably false negatives, who were therefore capable of self-recognition but did not show it during the task (Asendorpf & Baudonnière, 1993; Courage, Edison, & Howe, 2004; Nielsen & Dissanayake, 2004). Furthermore, there is evidence that children’s self-recognition status was the best predictor of children’s self-recognition status two weeks later (Kärtner, Keller, Chaudhary, & Yovsi, 2012). Thus, although evidence from previous studies points to acceptable retest reliability, it is nevertheless possible that the findings of the current study underestimate the percentage of self-recognizers due to false negatives. To date, no study has provided findings on the internal reliability of self-recognition tasks or the retest reliability of contingency tasks.

Six- and 9-month-olds looked longer at the left video image than at the right video image in both contingency tasks. This finding indicates a side preference at both ages, although we endeavored to reduce the possibility of the emergence of side biases, for instance by including a bell which was rung behind the curtain in the middle of the two video monitors in order to draw the infant’s attention to the space between the two monitors at the beginning of both tasks. Additionally, infants sat at identical distances to both monitors during both tasks. One explanation for the side bias might be that infants were placed into the infant seat from the left side, and the experimenter left the testing room through a curtain on the left side. However, we do not think that this side bias influenced the main results of this study, because the presentation of the contingent and the noncontingent video image, as well as the
presentation of the synchronous and the asynchronous video image, was randomized across all infants in both tasks.

**Future directions**

Both contingency tasks in the present study implemented a rather artificial situation in which infants observed video images that were contingent to own proprioception. Based on Kristen-Antonow et al.’s (2015) emphasis on the relevance of social situations for later self-recognition, future research should focus on more lifelike tests of contingency detection (Kristen-Antonow et al., 2015). A promising approach is to assess the relationship between parental contingent behavior during infancy and toddlers’ self-recognition. It has been proposed that infants discover themselves through interaction with contingent others (Bigelow, 2001). Empirical evidence provided first support for this idea: Children in a culture with relatively low parental contingent behavior recognize themselves in the mirror later than children in a culture with relatively high parental contingent behavior (Keller, Kärtner, Borke, Yovsi, & Kleis, 2005). The next step would be to test whether this relationship can also be found within one culture, by training one group of parents to interact more contingently with their infants. Toddlers of parents in the training group should recognize themselves earlier than toddlers in a control group (Riksen-Walraven, 1978).

**Conclusion**

The present longitudinal study provides no support for the notion that infants’ contingency preference and contingency detection in the first year of life predicts their ability to recognize themselves in a mirror or on a video monitor in the second and third year of life. Our findings support the notion that self-recognition emerges independently from the experience of intermodal matches (Bischof-Köhler, 1991, 2012; Suddendorf & Whiten, 2001). The ability to represent the self in the imagination emerges as a cognitive change in the second year of life, providing the prerequisite for the ability to compare this representation with the mirror image of the moving self (Bischof-Köhler, 2012). Thus, a representation of
the self seems to rely on more than a specific developmental pathway leading from contingency preference and contingency detection to self-recognition.
Intermodal Match Differentiation in the First Year of Life

References


Cameroonian Nso and German families. *International Journal of Behavioral Development*, 29, 496–504. doi:10.1080/01650250500147485


doi:10.1016/j.cognition.2011.03.001


doi:10.1016/j.cognition.2011.03.001

2 Relationship between Intermodal Match Preference and Social Interaction

Intermodal match preference and differentiation in the first year of life have been suggested to be related to self-recognition in the second and third year of life (Rochat, 2003). Study 1 provided no evidence for this relationship. The finding of study 1 rather supports the assumption that the emergence of a representation of the self occurs independently of the experience of intermodal matches, as a sudden cognitive change in the second year of life (Bischof-Köhler, 1991, 2012; Suddendorf & Whiten, 2001). Nevertheless, it has been suggested that the experience of intermodal matches in the first year of life provides material which forms the basis for this cognitive change in the second year of life (Bischof-Köhler, 1991).

2.1 Intermodal Match Preference and Social Interaction

The cognitive change, allowing representations of the self in the imagination in the second year of life, might be based on material provided in the first year of life (Bischof-Köhler, 1991). Such material might stem from infants’ experience interaction with their parents. It has been suggested that infants experience themselves through interaction with others, who provide them with high degrees of intermodal matches (Bigelow, 2001). In a longitudinal study, Kristen-Antonow et al. (2015) used a variation of the Still-Face Task (Tronick, Als, Adamson, Wise, & Brazelton, 1978) in which the experimenter first interacted naturally with the infant, and then interrupted the interaction. Infants’ reactions to this interruption were measured as intermodal match differentiation in social interactions. Indeed, infants’ intermodal match differentiation in social interactions at 9 months predicted mirror self-recognition at 24 months. Note that in contrast, study 1 measured infants’ intermodal match differentiation in a task that presented infants with different degrees of self-generated intermodal matches. The finding of Kristen-Antonow et al. (2015) thus highlights the importance of intermodal match experiences in communication and interaction between parents and their infants.
Accordingly, intermodal match differentiation in the first year of life could be closely related to the differentiation between self-generated behavior and other-generated behavior in social interactions in the first year of life (Bigelow, 2001). Infants’ high interest with respect to a preference for less-than-perfect matches (which are specified by other-generated actions; Bahrick & Watson, 1985; Rochat & Striano, 2000) could facilitate interaction and communication with others. Conversely, infants’ high preference for self-generated behavior could result in interaction and communication difficulties with their parents. The aim of study 2 is to examine the relationship between infants’ intermodal match preference and the parent-reported interaction and communication difficulties at 6 months. Data from study 2 stem from the longitudinal study presented in study 1.

Please note that the terms intermodal match preference and intermodal match preference task are not used throughout the manuscript of study 2; instead, the terms contingency preference and contingency-preference task are used. This inconsistency arose from the fact that the schema on the nature of perfect, less-than-perfect and no matches in intermodal match differentiation tasks (see section 1.2.3) was developed later than the manuscript.
2.2 Study 2 – Zmyj, N., & Klein-Radukic, S. (2015). Six-month-old infants' interaction difficulties are mirrored in their preference for perfect contingencies. *Infant Mental Health Journal*

Six-month-old infants' interaction difficulties are mirrored in their preference for perfect contingencies

Norbert Zmyj and Sarah Klein-Radukic

Ruhr-Universität Bochum, Bochum, Germany

Published in *Infant Mental Health Journal*, 2015, 36, 1-7. doi: 10.1002/imhj.21503

We thank Hannah Hermschulte, Sina Hulten, Nathalie Marcinkowski, Milena Meyers, Susanne Röttgers, and Paula Siegmann for their assistance in collecting and analyzing the data, and all the parents and their infants for participating in this study.

Direct correspondence to Norbert Zmyj, Ruhr-Universität Bochum, Faculty of Psychology, Universitätsstraße 150, 44801 Bochum, Germany; e-mail: norbert.zmyj@rub.de

Abstract

During the first months of life, infants switch from a preference for perfect contingent feedback to a preference for less-than-perfect contingent feedback of their own movements. This is an indicator of increasing social interest since others provide less-than-perfect contingent feedback whereas the self provides perfect contingent feedback. We presented 117 6-month-olds with real-time and delayed video feedback of self-performed leg movements
and asked parents about difficulties in various socioemotional domains. It was hypothesized that the more infants prefer real-time and therefore perfect contingent feedback over 7 s delayed and therefore perceived non-contingent feedback, the more difficulties parents will report in interaction and communication with their child. Preference for real-time feedback was related to difficulties in interaction, but not to difficulties in communication. Implications of this finding for infants’ socioemotional development and health are discussed.

It is of vital importance for infants to interact with and learn from their caregivers. Infants’ attachment to their caregivers has a phylogenetic root beyond the mere provision of food and a safe environment: Human culture depends on transmitting knowledge from one generation to the next (Csibra & Gergely, 2011). Imitation – which starts in infancy – plays an important role in this process (Tomasello, 1999). Since humans do not necessarily understand the evolutionary benefit of the transmission of knowledge, infants and parents should be interested in interacting with each other by default. Interacting and communicating with others is not perfectly predictable. It follows that infants should be interested in nonpredictable actions compared to perfectly predictable actions provided by themselves. In turn, an increased interest in perfectly predictable actions provided by the self might be accompanied by problems in parent-infant interactions and communication. This study was designed to test the relationship between infants’ interest in feedback provided by own actions and socioemotional difficulties in general and difficulties in interaction and communication with their parents in particular.

To interact and communicate with others, infants must be able to differentiate between the self and others. One way to capture this problem is by differentiating between various levels of matches between intermodal feedback (Watson, 1972); that is, when the sensory information of one modality is redundant with respect to the sensory information of another modality. One’s own behavior provides perfect matches in intermodal feedback whereas the
behavior of others does not. For example, when infants kick their legs, they receive visual and proprioceptive feedback that matches perfectly. Intermodal mismatches or imperfect matches between proprioceptive and visual feedback occur when infants observe others’ behavior (Watson, 1972).

Less-than-perfect intermodal matches can originate from different sources (Gergely & Watson, 1999). First, when infants wave their arms, for example, their parents might wave back – but not always. The contingency between stimulus (i.e., the infant’s proprioception) and response (i.e., visual feedback of the parent waving) is imperfect when a stimulus does not predict a response perfectly. Second, even if the parent waves back every time the infant waves, the match between proprioception and visual feedback can be imperfect. For example, the parent’s movement entails some variation in the amplitude and velocity of the movements, leading to imperfect congruency between proprioception and visual feedback. Third, even if the parent waves back in a mirror-like fashion every time the infant waves and thus provides contingent and congruent feedback, the match between proprioceptive and visual feedback can still be imperfect. That is, the parent’s response is temporally delayed compared to the infant’s proprioception, leading to imperfect timing. This short latency of parental behaviors adapts to infants’ ability to detect delayed responses as contingent to their own behavior.

There is mixed evidence regarding the maximum delay between stimulus and response which can be conceived as contingent. In delayed-reinforcement studies, a threshold of 3 s was reported (Millar & Watson, 1979), although short-term memory in infants might extend to 5 s (Watson, 1967). Contingent responses with a delay of more than 5 s are thus expected to be perceived as noncontingent. Timing of parental reactions is an important characteristic in interactions between parents and their infants because parents’ reactions usually occur within 1 s after the infants’ behavior (Keller, Lohaus, Völker, Cappenberg, & Chasiotis, 1999). Infants are sensitive to delayed feedback of their parents’ reactions in comparison to real-time feedback of their parents’ reactions by showing increased negative emotions (Henning &
It was shown that infants develop a preference for mismatches or imperfect matches in visual-proprioceptive feedback over perfect matches in visual-proprioceptive feedback. In a seminal study, infants were presented with two monitors, each showing leg movements: One monitor displayed a real-time video image of infants’ own leg movements, providing a perfect match between proprioception and visual feedback. The other monitor displayed a video image of leg movements of a peer, or delayed feedback of infants’ own leg movements recorded prior to the experiment, providing mismatches between proprioception and visual feedback (Bahrick & Watson, 1985). In this study, 5-month-olds, but not 3-month-olds, were found to look longer at the peer video or delayed video image of their own movements than at the real-time video image. A recent study, however, demonstrated that only 9-month-olds, and not 3- and 6-month-olds, showed this pattern of preferential looking (Geangu, Benga, Stahl, & Striano, 2011). In this paradigm, the two views differed in terms of contingency, congruency, and timing thus; it was difficult to disentangle the relative contribution of each of these variables to the infants’ preference. Subsequent studies have revealed that infants consider congruency and timing of video feedback of self-performed movements. As for congruency, 5-month-olds looked longer at incongruent real-time video images of self-performed movements (i.e., if infants moved their right leg, then the leg on the left side of the monitor moved) than they did at congruent real-time video images of self-performed movements (i.e., if infants moved their right leg, then the leg on the right side of the monitor moved; Morgan & Rochat, 1997). As for timing, there is convergent evidence that infants younger than 5 months do not and infants older than 7 months do discriminate between a real-time video image and an up to 3-s delayed video image of self-performed movements (Hiraki, 2006; Rochat & Striano, 2000; Zmyj, Hauf, & Striano, 2009). Accordingly, 6 months might represent a transitional age for differentiating between congruent and contingent feedback with different timing.
It is commonly assumed that the preference for imperfect or no contingencies reflects a preference for social stimulation, which does not provide perfect contingency by default (e.g., Gergely, 2001). However, infants do not have this preference from the beginning of their lives. Some researchers have believed that infants prefer perfect contingencies in the first few months after birth (Bahrick & Watson, 1985). Nevertheless, solid empirical evidence to support this proposition is still lacking. According to the contingency-switch hypothesis, infants start to prefer less-than-perfect contingencies over perfect contingencies in the first months of life (Bahrick & Watson, 1985; Geangu et al., 2011; Hiraki, 2006; Zmyj et al., 2009). It has been suggested that this transitional age is around 3 months of age (Gergely, 2001); however, studies using preferential looking paradigms with video image of leg movements showed that infants between 3 and 6 months do not always prefer less-than-perfect feedback over contingent feedback (Geangu et al., 2011; Zmyj et al., 2009). What has been described as a switch is not an all-or-nothing principle, but rather a gradual change that the majority of infants go through. A minority of infants continue to prefer perfect contingencies.

When infants maintain a preference for perfect contingencies, this might indicate a risk of autism-spectrum disorder (ASD), which is characterized, among other things, by communication and social interaction problems (American Psychiatric Association, 2013). The first empirical support for this hypothesis was reported by Gergely (2001), who presented children with and without autism with two monitors, each displaying a cursor. The cursor of one monitor was operated by the child him-/herself (perfect contingency) while the other cursor was operated by an experimenter who was not visible to the child and who matched the movements of the child’s cursor as closely as possible (imperfect contingency). Children with autism preferentially observed the cursor that they themselves operated over the cursor that was operated by the experimenter whereas children without autism showed the reversed preference (Gergely, 2001). Since autism is characterized by communication and social
interaction problems (American Psychiatric Association, 2013), one could ask whether problems in interaction are generally related to a preference for perfect contingencies. Therefore, the aim of the current study was to examine the relation between infants’ contingency preference and parents’ perceived problems in interaction with their infant.

In the present study, we tested infants’ preference for real-time and delayed views of their leg movements (as in Zmyj et al., 2009). We chose a delay of 7 1/2 s to present infants with perfect contingent and noncontingent feedback. (A delay of more than 5 s is probably not perceived as contingent; Watson, 1967.) The video image of delayed own movement has an important advantage over a video image of a peer’s leg movements in that it provides the same amount of movement and patterns of movements as does the video image of one’s own movements in real time. Both the amount of movements and the patterns of movements might influence infants’ visual preference, but are beyond the dimension of contingency. We tested 6-month-old infants because this age marks a transitional period toward a preference for the delayed view with delays shorter than 3 s (Hiraki, 2006; Rochat & Striano, 2000; Zmyj et al., 2009). Another reason was the mixed evidence regarding 5- to 6-month-olds’ preferential looking behavior between a video image showing congruent real-time feedback of their own leg movements and a video image of leg movements of a peer, or delayed feedback of infants’ own leg movements recorded prior to the experiment (Bahrick & Watson, 1985; Geangu et al., 2011). Thus, whether infants perceived the contingency of the 7 1/2-s delayed feedback, we expected that 6-month-olds would show no clear-cut preference for either view.

Parents were asked to rate their infants’ socioemotional difficulties on the Ages & Stages Questionnaire: Social-Emotional (ASQ:SE; Squires, Diane, & Elizabeth, 2003). This questionnaire provides a total score and six subscales, including subscales for difficulties in social interaction and communication. We hypothesized that the more infants prefer the real-time view, the more problems parents will report in the subscales Social Interaction and Communication. To control for general cognitive development, we applied the Cognitive

Method

Participants

Infants were tested at a time of day when they were likely to be awake and in an alert state. Families were recruited from a list of parents who had previously agreed to participate in child development studies. Participants were 117 six-month-old, healthy, full-term children (53 female; $M$ age = 6 months 5 days; $SD = 7$ days). Six additional children participated, but had to be excluded due to crying ($n = 2$), equipment failure ($n = 2$), or because the parents did not fill out the ASQ:SE ($n = 2$). All participants received a small gift, 5 euros, and a certificate for their participation as an acknowledgement at the end of their visit.

Design

Three tasks were administered: a contingency-preference task, the assessment with the BSID-III, and the ASQ:SE. For the contingency-preference task, infants sat in a baby seat and for the BSID-III, infants sat on their parent’s lap at a table. The contingency-preference task was conducted first, followed by the BSID-III. We chose this fixed order because infants tend to fuss in the contingency-preference task if they have just participated in the BSID-III. The ASQ:SE was completed by the accompanying parent before the experiment started.

Apparatus, materials and procedure

Testing took place in a room partitioned into two sections and surrounded by beige curtains. A fan produced a 30-dB white noise.

Contingency-preference task. In the contingency-preference task, Camera 1 (Canon Legria HV40, HDV 1080) was positioned in front of the infant, who sat in a baby seat (Maxi-Cosi, Dorel-Industries Inc., The Netherlands). The distance from the infant’s face was approximately 130 cm. The infant’s leg movements were filmed from above the baby seat with Camera 2 (Canon Legria HV40, HDV 1080) from a distance of approximately 130 cm.
This image was fed into a device (Delay Line, Ovation Systems) which allows short-term storage and adjustable delayed output of real-time video recording. In front of the infant (on the left and right side of Camera 1), two displays (Dell, UltraSharp 2007FPb) showed the video image recorded by Camera 2. One monitor showed a real-time video image of the infant’s legs while the other monitor showed a delayed video image (delay 7 1/2 s) of the infant’s legs. The side at which the delayed display was positioned was randomized across infants. All infants were dressed in white tights, which contrasted with the dark blue baby seat. The tights had a black dot on the middle of the upper side of the feet. The duration of the contingency-preference task was 4 min.

**BSID-III.** The current study used the Cognitive Scale of the BSID-III to assess infants’ cognitive development. The BSID-III provides standard IQ scores \( M = 100, SD = 15 \) for cognitive development. The application of the BSID-III followed the administration and technical manual. Forty (34%) randomly chosen videotaped assessments of the BSID-III were additionally coded by an independent second observer. Intraclass correlation for interrater reliability for the cognitive development score was high, \( r = .86 \).

**The ASQ:SE.** To examine infants’ socioemotional development, the German version of the ASQ:SE for 6-month-olds was used. The questionnaire consists of 22 items. Internal consistency ranged from .67 to .91 (Cronbach’s \( \alpha \)) and test-retest reliability was 94%. Concurrent validity ranged from 81 to 95% (overall agreement 93%). The ASQ:SE targets six different dimensions of socioemotional development: self-regulation, interaction, affect, communication, adaptive functioning, and general concerns. Table 1 provides a summary of the definitions of the subscales as well as item examples and the numbers of items in the subscales. The items are answered on a scale indicating whether the behavior occurs 0 (*most of the time*), 1 (*sometimes*), or 2 (*rarely or never*). In addition, each question can be marked by the parent if this particular behavior is a matter of concern to them. (“Check if this is a concern.”) Four items are open-ended and relate to eating, sleeping and toileting. An example
of an open-ended item is: “Is there anything that worries you about your baby? If so, please explain.”

<table>
<thead>
<tr>
<th>ASQ:SE Subscale</th>
<th>Associated Definition</th>
<th>Example</th>
<th>No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Regulation</td>
<td>Items address the child’s ability or willingness to calm or settle down or adjust to physiological or environmental conditions or stimulation.</td>
<td>“Is your baby able to calm himself down (for example, by sucking on his hand or a pacifier)?”</td>
<td>5</td>
</tr>
<tr>
<td>Communication</td>
<td>Items address the child’s ability or willingness to respond to or initiate verbal or nonverbal signals to indicate feelings, affective, or internal states.</td>
<td>“Does your baby let you know when she is hungry or sick?”</td>
<td>2</td>
</tr>
<tr>
<td>Adaptive Functioning</td>
<td>Items address the child’s success or ability to cope with physiological needs (e.g., sleeping, eating, elimination, safety).</td>
<td>“Does your baby sleep at least 10 hours in a 24-hour period?”</td>
<td>6</td>
</tr>
<tr>
<td>Affect</td>
<td>Items address the child’s ability or willingness to demonstrate his or her own feelings and empathy for others.</td>
<td>“Does your baby like to be picked up and held?”</td>
<td>2</td>
</tr>
<tr>
<td>Interaction With People</td>
<td>Items address the child’s ability or willingness to respond to or initiate social responses to parents, other adults, and peers.</td>
<td>“When awake, does your baby seem to enjoy watching or listening to people?”</td>
<td>3</td>
</tr>
</tbody>
</table>

**Coding**

**Contingency-preference task.** The software Interact 9.3.5 (Mangold Software & Consulting GmbH, Arnstorf, Germany) was used to determine how long infants looked at each monitor based on the video recordings. The proportion of looking time at the delayed video image (looking at delayed video/looking at both videos) was calculated. Values higher than .50 indicated a preference for the delayed video image. The range of the proportion of looking time at the delayed video image was .19 to .81. In addition, randomly chosen videos of 40 participants were coded by an independent second observer. Interrater reliability for the proportion of looking time at the delayed video image using intraclass correlation was high ($r = .99$). Both coders were naïve regarding the side at which the delayed video image was positioned.

**ASQ:SE.** Parents’ responses in the ASQ:SE were scored with 0 (*rarely or never*), 5 (*sometimes*), or 10 (*most of the time*), indicating whether difficulties were present. An additional 5 points were given if parents were concerned about a specific behavior. If missing values occurred, the final score was computed in a two-step procedure: First, the average score (child’s total score of answered items/number of answered items) was computed. Second, the final score [child’s total score of answered items + (average score x number of
unanswered items) was computed. Questionnaires with more than two missing values were excluded. A total score, and scores for self-regulation, interaction, affect, communication, adaptive functioning, and general concerns were determined. The higher the score, the more difficulties parents perceive in this area of socioemotional development. Open-ended questions were not included in the data analysis.

**Results**

**Contingency preference**

There was no difference between infants’ looking time at the delayed video image and at the real-time video image ($M = 82.04$ s; $SD = 30.27$, and $M = 86.16$ s, $SD = 28.76$, respectively), $t(116) = 1.04$, $p = .299$. Mean proportion of looking time at the delayed video image (looking at delayed video/looking at both videos) was $M = .48$ ($SD = 0.12$). A single sample $t$ test against the chance level of .50 revealed no differences between the mean proportion of looking time and chance level, $t(116) = -1.40$, $p = .165$.

In addition, infants were dichotomized into a group who looked longer at the delayed display and a group who looked longer at the real-time display. Fifty-one of 117 infants (43.6%) spent more time looking at the delayed display than they did at the real-time display, $p = .195$, two-tailed (binomial exact test).

**The ASQ: SE**

Mean total scores were as follows: $M = 18.06$ ($SD = 12.60$) for the ASQ:SE, $M = 9.67$ ($SD = 7.70$) for the Self-Regulation scale, $M = 0.88$ ($SD = 2.14$), for the Interaction scale, $M = 0.51$ ($SD = 1.66$) for the Communication scale, $M = 2.49$ ($SD = 3.32$) for the Affect scale, $M = 4.20$ ($SD = 6.50$) for the Adaptive Functioning scale, and $M = 0.30$ ($SD = 1.06$) for the General Concerns scale.

**The BSID-III**

Assessment with the Cognitive Scale of BSID-III revealed a mean cognitive development score of the participating infants of $M = 106$ ($SD = 8.46$, range = 75-125).
Correlation between Contingency Preference and the ASQ:SE

Table 2 shows the correlation matrix displaying the relations between proportion of looking time at the delayed video image in the contingency-preference task, the scores of the ASQ:SE scales, and the total score, respectively. A correlation analysis revealed a significant relation between the proportion of looking time at the delayed video image and the score in the Interaction scale, $r = -.273$, $p = .003$, Spearman rank correlation. No further significant correlations between proportion of looking time at the delayed video image and scales of the ASQ:SE were found. A partial correlation analysis controlling for cognitive development revealed a significant relation between the proportion of looking time at the delayed video image and the score of the Interaction scale, $r = -.275$, $p = .003$, Spearman rank correlation.

<table>
<thead>
<tr>
<th>Proportion of Looking Time at Delayed Video</th>
<th>Self-Regulation</th>
<th>Interaction</th>
<th>Communication</th>
<th>Affect</th>
<th>Adaptive Functioning</th>
<th>General Concerns</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman Rank Correlation</td>
<td>$-.048$</td>
<td>$-.273^*$</td>
<td>$-.138$</td>
<td>$-.147$</td>
<td>$-.090$</td>
<td>$.079$</td>
<td>$-.113$</td>
</tr>
<tr>
<td>Partial Spearman Rank Correlation</td>
<td>$-.076$</td>
<td>$-.275^*$</td>
<td>$-.125$</td>
<td>$-.146$</td>
<td>$-.108$</td>
<td>$.061$</td>
<td>$-.141$</td>
</tr>
<tr>
<td>Controlling for Cognitive Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $p < .01$.

Discussion

The present study showed that the more infants look at the real-time and therefore perfect contingent feedback to their own behavior, the more likely their parents are to report difficulties in the interaction with their infant. This finding was predicted by the contingency-switch hypothesis (Gergely, 2001). According to this account, infants with a preference for perfect contingencies pay more attention to their own behavior than they do to others’ behavior, which results in interaction difficulties. In other words, infants who attend less to their parents have more difficulties in regulating their behavior in interactions with them. The results of the present study support the idea that the emerging interest in less-than-perfect contingent feedback is accompanied by an increased interest in social interaction (Gergely,
There are different alternative explanations for the present finding. First, one could argue that infants with more difficulties in interactions and a preference for perfect contingencies might simply be more delayed in their general cognitive development. When controlling for general cognitive development, however, the relation was still present. Second, the causal direction might be in the opposite direction. That is, parents themselves might have difficulties in interacting with their infants and project these problems onto their children. In turn, infants would then focus more on their own behavior than they would on their parents’ behavior to avoid problematic interactions. Although we cannot rule out this possibility, we do not believe it to be a very likely explanation due to the low mean total score of the ASQ:SE, indicating that parents perceived problems in only a few situations. If parents had difficulties in interacting with their child, which is a very basic parenting skill, then we would expect them to report difficulties in other more vulnerable domains such as the Affect scale. Third, the relationship between interest in perfectly contingent feedback and presence of interaction problems might be influenced by the infants’ temperament. For example, activity level, duration of orienting responses, and distress reactions to novel stimulation – typical dimensions of temperament – were found to predict the amount of infants’ crying when the contingent feedback changed unexpectedly (Fagen & Ohr, 1985). Although this study cannot explain the present findings, it does point to a potential relationship between temperament and contingency preference. The switch in contingency preference (from perfect to less-than-perfect) might develop at different time points depending on the individual temperament. For example, extrovert infants might be interested in less-than-perfect contingent feedback earlier than might their peers. Extrovert infants also may have fewer problems in interactions with their parents. The differential perspective on contingency preferences in this study could be further developed by studying infants’ temperament in future studies.

Contrary to our hypothesis, we were unable to find a correlation between a preference
for perfect contingent feedback and difficulties in communication. A possible explanation is that parents perceived almost no difficulties in communication displayed by their children, making a floor effect likely. Moreover, in the Communication scale, parents were asked, for example, whether children let the parent know if they are hungry, sick, or tired. However, at 6 months of age, parents do not usually expect words or gestures when their child is in need of something, but rather try to infer the need based on their child’s overall behavior.

In the present study we tested infants’ preference for perfectly contingent feedback over noncontingent feedback and examined the relationship of this preference with interaction problems with the parents. Parents’ behavior is not only noncontingent but also imperfectly contingent. Parents respond to their infants’ behavior within a small latency window of approximately 1 s (Keller et al., 1999). Since parents do not provide contingent responses to every one of their infants’ behaviors, the parents’ behavior can be considered as imperfectly contingent. Thus, a future research avenue would be to investigate infants’ interest in imperfectly contingent feedback and the absence of interaction problems. The reason why this research has not yet been conducted is that so far, there is no manipulation of the infants’ movements available that results in an imperfectly contingent feedback with the visual characteristics of the infants’ body.

The idea of a switch from a preference for perfect contingent feedback to less-than-perfect contingent feedback has been articulated on several occasions (Bahrick & Watson, 1985; Gergely, 2001; Gergely & Watson, 1999). Unfortunately, there is no direct evidence – including in the present study – for this switch. However, the present study was able to support a central assumption of this idea. Thus, future research should not only investigate the developmental trajectories of contingency preference but also account for difficulties in interactions during this development. Furthermore, there has been a recent surge of interest in identifying ASDs as early as possible (Robins, Fein, Barton, & Green, 2001). Key items in the Modified Checklist for Autism in Toddlers (Robins et al., 2001) in terms of predicting ASD
ask about social interaction and communication. In this study, we reported a relationship between interaction problems identified in the ASQ:SE and interest in perfectly contingent feedback. This relationship could be an indicator that ASD-associated behavior might be detected via questionnaire and behavioral observation in the first year of life (for similar lines of thought, see Yirmiya et al., 2006). It therefore might be a promising idea to investigate children at risk of an ASD by the age of 6 months to ascertain whether a preference for perfect contingent feedback predicts later traits related to ASD.

This is the first study to show that infants’ difficulties in interactions are related to their interest in perfect contingent feedback. Developmental psychologists, psychiatrists and pediatricians are aiming to achieve a better and earlier diagnosis of problems in parent-child relations in general and developmental disorders such as ASD in particular. This study suggests that future research also should focus on infants’ interest in perfect contingencies, which are related to interaction difficulties with their parents – a key issue of infant mental health.
References


3 Dropout in Infancy Research

High rates of dropout from specific tasks due to fussiness can sometimes arise during test phases in infancy research (Slaughter & Suddendorf, 2007; Wachs & Smitherman, 1985). This raises the question whether the final sample is still representative as compared to the original sample. The issue of representativeness is especially problematic in looking time studies and in longitudinal studies.

3.1 Dropout in Looking Time Studies

The findings of study 1 and 2 rest on the interpretation of infants’ behavior in looking time tasks. Looking time tasks are commonly used among infancy researchers to assess infants’ processing of and preference for objects and stimuli. One could argue that infants’ ability to reach towards objects or other stimuli could be a more direct measure especially of infants’ preferences. However, the ability to perform precision grasping, that is to reach accurately towards objects, is distinctly developed in the early second year of life (Butterworth, Verweij, & Hopkins, 1997). Looking behavior, along with more complex methods such as EEG measurements, thus constitute more straightforward ways to measure infants’ processing of visual stimuli at younger ages. A great advantage of looking time tasks is that they are easy to implement at a low cost in comparison to EEG measurements. However, during looking time tasks, infants are required to sit still for several minutes, which is why such tasks are prone to dropout. Additionally, infants are repeatedly presented with recurring stimuli during looking time tasks, which might be perceived as predictable and rather boring for the infants. Therefore, study 3 aimed at examining whether dropout in looking time tasks such as the intermodal match differentiation tasks used in study 1 and 2 could have affected the representativeness of the final sample in terms of infants’ cognitive development. In addition, it was of interest which infant temperament dimensions could be related to infants’ dropout from such looking time tasks.
3.2 Dropout in Longitudinal Studies

In study 1, data were gathered from a longitudinal study. The general dropout rate, in terms of complete dropout from test phases due to moves or other reasons, was low. The dropout rates from specific tasks during test phases due to fussiness at the different test phases were not reported due to space restrictions. However, dropout due to fussiness occurred in the different test phases, and sample sizes for each analysis were provided, ranging from 63 to 113. If stable infant characteristics play a role in infants’ dropout due to fussiness, it is likely that infants’ dropout in one test phase could be associated with their dropout in another. Thus, study 3 aims at examining whether infants’ dropout in one test phase is associated with their dropout in a similar test phase 3 months later. Data from study 3 stem from the longitudinal study presented in study 1.

Please note that the term intermodal match differentiation task is not used throughout the manuscript of study 3; instead, the term looking time task is used. This inconsistency arises from the fact that the purpose of study 3 is rather superordinate in nature, concerning methodological considerations of longitudinal studies and the use of looking time tasks.

Dropout in looking time studies: The role of infants’ temperament and cognitive developmental status

Sarah Klein-Radukic and Norbert Zmyj

Technical University of Dortmund, Dortmund, Germany

Published in Infant Behavior & Development, 2015, 41, 142-153.

doi:10.1016/j.infbeh.2015.10.001

Corresponding author at: Institute of Psychology, Technical University of Dortmund, Emil-Figge-Str. 50, 44227 Dortmund, Germany. E-mail address: sarah.klein-radukic@tu-dortmund.de (S. Klein-Radukic).

Abstract

Dropdown of infants in looking time studies sometimes occurs at high rates, raising concerns that the representativeness of the final sample might be reduced in comparison to the originally obtained sample. The current study investigated which infant characteristics play a role in dropout. Infants were presented with a preferential looking task at 6 and 9 months of age. At 9 months of age, an additional habituation task and a subsequent novelty preference task were conducted. In addition, temperament was assessed via the Infant Behavior
Questionnaire – Revised (IBQ-R, Gartstein & Rothbart, 2003), and cognitive developmental status was assessed via the Cognitive Scale of the Bayley Scale of Infant and Toddler Development (BSID-III, Bayley, 2006). Dropout was positively related to the IBQ-R temperament scales Distress to Limitations and Approach, and negatively related to the scales Falling Reactivity and Cuddliness. The representativeness of the final sample regarding situation-specific temperament dimensions is affected by dropout. Dropout was not related to cognitive developmental status as measured via the BSID-III, habituation speed and novelty preference. Dropout at 6 months of age was associated with dropout at 9 months of age. We concluded that in looking time studies, the representativeness of the final sample regarding performance-relevant temperament dimensions or cognitive developmental status is not affected by dropout.

Keywords: Subject loss, Infant research, Preferential looking, Mental development, Personality, Character

1. Introduction

Dropout in infant studies is a problem which all researchers have to face when testing infants and toddlers at home or in a laboratory. Possible reasons for the termination of a test session are crying, lack of responding, restlessness or sleepiness. This change in some infants’ affective and arousal state can be caused either by random factors such as the infants’ state on the day or by predictable factors such as the infants’ temperament. If the latter is the case, the obtained sample is not representative of the original sample. Especially in research with pre-linguistic infants, the looking time task is a widely used method. Looking time tasks enable the assessment of attention-related processes and infants’ discrimination abilities (Saayman, Ames, & Moffett, 1964). Looking time tasks are especially prone to dropout because infants are required to sit still for several minutes. Average dropout rates of up to 26% are normal (Wachs & Smitherman, 1985), and studies examining which factors account for dropout in
looking time tasks have provided mixed results so far. In order to investigate this phenomenon in more detail, we tested whether infants’ temperament and cognitive developmental status predict dropout in a looking time study at 9 months of age. Additionally, dropout in a looking time study at 6 months of age was assessed in order to predict dropout in a looking time study at 9 months of age.

1.1. Dropout and temperament

Temperament is an infant characteristic which might influence dropout in looking time studies. A standard method to measure temperament in infancy is the assessment via the Infant Behavior Questionnaire – Revised (IBQ-R, Gartstein & Rothbart, 2003). The questionnaire is based on a model of temperament which “refers to individual differences in reactivity and self-regulation” that are assumed to be a “relatively enduring biological makeup of the organism” (Rothbart & Derryberry, 1981, p. 40).

Temperament has been defined as “relatively enduring” (Rothbart & Derryberry, 1981, p. 40), which suggests a high stability of temperament over time in the first years of life. There is mixed evidence for the idea of a high stability of temperament in empirical studies. On the one hand, some longitudinal studies on infant temperament showed high stability for different aspects of temperament (e.g., Gartstein, Putnick, Kwak, Hahn, & Bornstein, 2015; Mink, Henning, & Aschersleben, 2013, for general infant temperament; Buss, Block, & Block, 1980, for activity level; Putnam, Gartstein, & Rothbart, 2006; Rothbart, Derryberry, & Hershey, 2000, for positive emotionality). On the other hand, other longitudinal studies on infant temperament showed low stability for temperament (e.g., Bridgett et al., 2009, for regulatory capacity; Carranza Carnicero, Pérez-López, González Salinas, & Martínez-Fuentes, 2000, for emotional tone, social orientation, activity level, and vocalization). There are several explanations for these inconsistent findings. First, low stability of infant temperament measures might be related to problems in the concept and measurement of temperament (Hubert, Wachs, Peters-Martin, & Gandour, 1982). For example, it is difficult to
control for other factors in the longitudinal measurement of temperament. That is, disentangling temperament development and environmental and parental influences can be problematic; it has been shown that parenting practices are related to temperament development (Bridgett et al., 2009). Second, the expression of temperament in young infants might change over time as the infant develops. The low stability of temperament might therefore result from developmental changes in the expression of temperament rather than from changes in the underlying structure of temperament (Riese, 1987). The authors of the IBQ acknowledge the idea of developmental changes by recommending taking the “possibility of developmental change and individual differences in developmental timetables [...] into account” (Rothbart, 1986, p. 356).

Several studies have addressed the relationship between temperament and dropout, with questionnaires being the most widely used method to assess infant temperament (Fagen, Ohr, Singer, & Fleckenstein, 1987, Fagen, Singer, Ohr, & Fleckenstein, 1987; Miceli, Whitman, Borkowski, Braungart-Rieker, & Mitchell, 1998; Mink, Henning, & Aschersleben, 2013; Treiber, 1984; Vonderlin, Pahnke, & Pauen, 2008; Wachs & Smitherman, 1985). However, the findings in this regard have been mixed: On the one hand, researchers reported differences in temperament between non-completers (NC) and completers (C) of looking time tasks, with NCs being rated as more fussy and inadaptable (for female Cs and NCs, see Wachs & Smitherman, 1985; for both boys and girls, see Treiber, 1984), as sadder and less able to maintain orientation for longer periods of time (Fagen, Ohr et al., 1987; Mink et al., 2013), and as more active (Miceli et al., 1998) than Cs. NCs were found to smile and laugh more often than Cs (Miceli et al., 1998), which is inconsistent with the finding that NCs were sadder than Cs (Mink et al., 2013). On the other hand, some studies reported no differences between NCs and Cs in temperament (for boys, see Wachs & Smitherman, 1985; for boys and girls, see Vonderlin et al., 2008) and in their general amount of movements (Lewis & Johnson, 1971). However, the comparability of the studies is limited due to different task
types, different sample sizes, different methods to assess infant temperament, and different infant ages.

All things considered, studies assessing whether Cs and NCs of test sessions differ in their temperament have produced mixed findings. NCs were reported to be less able to maintain orientation, to be more active, fussier and less adaptable than Cs. In contrast, some studies reported no differences between NCs and Cs. Therefore, there is a clear necessity to shed further light on this question.

1.2. Dropout and cognitive developmental status

Cognitive developmental status is another infant characteristic which might influence dropout in looking time studies. Standard methods to measure cognitive developmental status in infancy are assessments via the Cognitive Scale of the Bayley Scale of Infant and Toddler Development (BSID-III, Bayley, 2006), habituation speed, and novelty preference tasks (McCall & Carriger, 1993). To the best of our knowledge, only a small number of studies have analyzed the influence of infant cognitive developmental status on dropout. One study has examined habituation speed in Cs and NCs of a habituation task (Richardson & McCluskey, 1983). Although no differences in habituation speed between Cs and NCs were found, it was not reported how the number of trials to habituation in NCs was analyzed.

There might be two possible relationships between dropout and cognitive developmental status: First, infants with high cognitive developmental status might be more likely to drop out of looking time studies because these tasks might be too undemanding, leading the infants not to complete them. This idea received support from a study in which NCs of an attention-getting procedure were more likely to drop out when presented with simple compared to complex stimuli (DeLoache, Rissman, & Cohen, 1978). The authors suggested that developmentally advanced infants were more likely to drop out of the simple stimuli conditions than less advanced infants. Second, the ability to complete a task might be a cognitive demand in itself. High cognitive developmental status might therefore enable
infants to complete tasks, whereas infants with lower cognitive developmental status might be
less likely to complete tasks due to the high cognitive demands of the task. This suggestion
was supported by a longitudinal study in which NCs were found to have lower developmental
scores on the Bayley Mental Scales than Cs (Bathurst & Gottfried, 1987). No studies to date
have used habituation speed to assess the influence of cognitive developmental status on
dropout rates of other tasks. The ambiguity of findings in this respect is unsatisfactory,
because there might be a fundamental relationship between these domains.

1.3. Dropout as predictor of subsequent dropout

If potentially stable characteristics such as temperament and cognitive developmental
status have an impact on dropout in looking time studies, then NCs at time 1 should also be
NCs at time 2. This notion has received only limited empirical support. A longitudinal study
showed no association between dropout at time 1 and time 2: Infants were tested at 30, 36 and
42 months of age, and dropout at 30 months of age was not associated with dropout at later
test sessions (Bathurst & Gottfried, 1987). However, the statistics used in order to assess
stability were not reported. Another study reported that NCs of a looking time task were less
likely than Cs to complete a looking time task one week later and a problem-solving task nine
months later (Bell & Slater, 2002). However, the authors used a chi-square test, despite the
fact that Cohen’s Kappa might have provided a more stringent test of the existence of stability
of dropout. Another study reported no association between the dropout rates of looking time
tasks conducted at 6 and 12 months of age (Mink et al., 2013). Given these inconsistent
findings, the present study was conducted in order to shed further light on the question
whether infants’ dropout at one time point of assessment is associated with infants’ dropout at
a later time point of assessment.

1.4. The present study

Findings on the relationship between infant temperament and dropout are contradictory,
and only a small number of studies have investigated whether differences in cognitive
developmental status are related to dropout in infant looking time studies. In order to obtain further findings in this regard, the current study, which was part of a longitudinal study, assessed several infant characteristics and their relationship to dropout. We investigated the following questions: First, is there a relationship between infant temperament and dropout from preferential looking tasks? In order to answer this question, infant temperament was measured at 9 months of age via the Infant Behavior Questionnaire – Revised (IBQ-R, Gartstein & Rothbart, 2003, Vonderlin, Ropeter, & Pauen, 2012). Dropout was measured in three different tasks: a habituation task, a novelty preference task and a preferential looking task. Second, we asked: Is there a relationship between infants’ cognitive developmental status and dropout from preferential looking tasks? In order to answer this question, we assessed infants’ cognitive developmental status using the Cognitive Scale of the BSID-III (Bayley, 2006), the habituation speed, and the novelty preference. The BSID-III score, habituation speed and novelty preference of Cs and NCs of the aforementioned tasks were compared. It should be noted that the habituation task provided data on dropout and also served as a measurement of cognitive developmental status. Third, we asked: Is there a relationship between the dropout of infants at one time point of assessment and the dropout at a later time point of assessment? In order to answer this question, the relationship between the dropout of 6-month-olds and the dropout of 9-month-olds was assessed. Thus, a preferential looking task was conducted at 6 and 9 months of age.

2. Methods

2.1. Participants

One hundred and six healthy, full-term infants participated in this study. At their first visit to the laboratory at the Ruhr-Universität Bochum, infants were 6 months old (M = 6 months, 4 days; SD = 6 days; range: 5 months, 19 days to 6 months, 23 days; 55 boys and 51 girls). At the second visit, the same infants were 9 months old (M = 9 months, 9 days; SD = 9 days; range: 8 months, 18 days to 9 months, 30 days). An additional 17 infants had to be
excluded from the study for the following reasons: Infants did not participate at the age of 9 months ($N = 12$), parents did not fill out the IBQ-R at 9 months of age ($N = 3$), infants were substantially older than the other infants at the time of the first and the second visit, respectively ($N = 2$). Infants were recruited from a database of families who lived in the Ruhr area and had previously agreed to participate in studies on child development. Parents were asked to make appointments for testing at a time of the day when their infant is most likely to be awake and alert. The IBQ-R was sent to the parents one week prior to the second visit. Parents were asked not to interfere at any time of testing. At the end of their visit, all infants received a small gift and the parents received 5 Euro.

The dropout rate of 8% due to fussiness in the preferential looking task at 6 months of age was equal to other similar studies (e.g., 9%, Bahrick & Watson, 1985; 16%, Zmyj et al., 2009). The dropout rate of 29% due to fussiness in the preferential looking task at 9 months of age was higher than in another similar study (17%, Zmyj, Jank, Schütz-Bosbach, & Daum, 2011). The dropout rate of 8% due to fussiness in the habituation and novelty preference task at 9 months of age was equal to other similar studies (9%, Kelly et al., 2009, 9%, Kirkham, Slemmer, & Johnson, 2002).

2.2. Design

Testing consisted of four different tasks, one of which was conducted at 6 months of age and the other three tasks at 9 months of age. Six- and 9-month-olds participated in a preferential looking task. Nine-month-olds also participated in tasks that measured their cognitive developmental status: Infants participated in a task testing infants’ habituation speed and novelty preference as well as in the Cognitive Scale of the BSID-III. The order of the tasks was the same for all infants in order to keep effects of attrition constant for all infants.

2.3. Apparatus and procedure

Infants were tested in a room that was partitioned into two adjacent areas. Each area was surrounded by beige curtains. A fan produced a 30 dB white noise that served as a comforting
background noise during testing.

2.3.1. Preferential looking task at 6 and 9 months of age

At 6 months of age, a preferential looking task was conducted. Infants sat in a baby seat in front of two monitors. Both monitors provided video feedbacks of the infants’ legs movements, which were videotaped by a camera that was mounted above the baby seat. One monitor displayed real-time feedback of the leg movements whereas the other monitor displayed delayed feedback of the leg movements. For a more precise description of the preferential looking task at 6 months of age, see Klein-Radukic & Zmyj (2015). At 9 months of age, an analogous preferential looking task was conducted. Again, infants sat in a baby seat in front of two monitors. Both monitors showed videos of lifelike baby doll legs being stroked in different patterns. At the same time, one of the infants’ legs was stroked. The leg was stroked synchronously with one of the video displays and asynchronously with the other display. For a more precise description of the preferential looking task at 9 months of age, see Zmyj et al. (2011). The duration of both tasks was 4 minutes. Parents were allowed to stroke their infants’ head in order to console them during the preferential looking tasks at 6 and 9 months of age.

2.3.2. Habituation and novelty preference task at 9 months of age

During the habituation task, infants sat on their parents’ lap in front of a monitor which was approximately 75 cm away. A camera was mounted above the monitor, providing a video image of the infant for the experimenter. Behind the curtain, the experimenter used this video to time-code infants’ looking behavior online during habituation trials. Infants habituated to a stimulus that contained geometrical forms (square, circle, and triangle) in different colors (red, blue, and yellow, see Figure 1 a). On repeated trials, infants saw the same picture with the geometrical stimuli. Habituation trials were separated by attention-getters. These attention-getters were pictures of toys that moved across the screen and made different sounds in order to caption the attention of the infant. If an infant did not reach the habituation
criterion within 20 trials, the habituation phase was terminated. When infants reached the habituation criterion of the habituation task, they were presented with the novelty preference task. After the last trial of the habituation task, another attention-getter was presented. When infants observed the monitor, both the habituated stimulus and a new stimulus were presented. The new stimulus contained the same geometrical forms in a new arrangement and with exchanged colors (see Figures 1b and 1c). The side of the new stimulus was randomized across all infants. The duration of the task was 30 s.

![Fig. 1](image.png)

**Fig. 1.** (a) Stimulus presented during habituation task, (b) and (c) Stimulus presented during novelty preference task, after infants reached the habituation criterion. (For interpretation of the references to color in the text, the reader is referred to the web version of this article.)

### 2.3.3. BSID-III at 9 months of age

In order to assess infants’ cognitive developmental status, the Cognitive Scale of the BSID-III (Bayley, 2006) was applied. The assessment with the BSID-III followed the
guidelines in the administration and technical manual. The BSID-III provides standardized IQ scores \((M = 100, SD = 15)\) for cognitive developmental status.

**2.3.4. IBQ-R at 9 months of age**

Infant temperament was assessed via the German version of the IBQ-R (Gartstein & Rothbart, 2003; Vonderlin et al., 2012), which was sent to the parents one week before testing. The IBQ-R consists of 191 items. Parents were asked to judge the frequency of different behaviors of their child during the last 7 days. An example of an item concerning Distress to Limitation is: “When placed in an infant seat or car seat, how often did the baby show distress at first; then quiet down?” Answers were given on a 7-point Likert scale ranging from 1 (never) to 7 (always). Additionally, parents had the possibility to choose “does not apply” if they were not able to observe their child in the depicted situation. The questionnaire contains 14 scales which are described in Table 1. Internal consistency of the German version of the IBQ-R ranges from \(r = .74\) to \(r = .92\) (Cronbach’s alpha, Vonderlin et al., 2012). The six-month stability of temperament rating (Pearson correlation) ranged from moderate \((r = .39\) for Fear) to high \((r = .62\) for Smiling and Laughter), except for the scales Cuddliness and Soothability, which had no stability for 6 months (Mink et al., 2013). A recent analysis of reliability and validity of the English version of the IBQ-R revealed good internal consistency and good interrater reliability, but low convergent validity, which was additionally moderated by parent depression (Parade & Leerkes, 2008).
Table 1
Summary of associated definitions of the IBQ-R subscales, item examples and numbers of items in the subscales.

<table>
<thead>
<tr>
<th>IBQ-R subscale</th>
<th>Associated definition (Gartstein &amp; Rothbart, 2003)</th>
<th>Example</th>
<th>Number of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity level</td>
<td>Gross motor activity, including movement of arms and legs, squirming and locomotor activity.</td>
<td>When placed in an infant seat or car seat, how often did the baby wave arms and legs?</td>
<td>15</td>
</tr>
<tr>
<td>Distress to limitations</td>
<td>Fussing, crying or showing distress while (a) in a confining place or position; (b) in caretaking activities; (c) unable to perform a desired action.</td>
<td>When the baby wanted something, how often did s/he become upset when s/he could not get what s/he wanted?</td>
<td>16</td>
</tr>
<tr>
<td>Fear</td>
<td>Startle or distress to sudden changes in stimulation, novel physical object or social stimuli; inhibited approach to novelty.</td>
<td>When introduced to an unfamiliar adult, how often did the baby never “warm up” to the unfamiliar adult?</td>
<td>16</td>
</tr>
<tr>
<td>Duration of orienting</td>
<td>Attention to and/or interaction with a single object for extended periods of time.</td>
<td>How often during the last week did the baby spend time just looking at playthings?</td>
<td>12</td>
</tr>
<tr>
<td>Smiling and laughter</td>
<td>Smiling or laughter during general caretaking and play.</td>
<td>How often during the last week did the baby laugh aloud in play?</td>
<td>10</td>
</tr>
<tr>
<td>High pleasure</td>
<td>Pleasure or enjoyment related to high stimulus intensity, rate, complexity, novelty, and incongruity.</td>
<td>How often did your baby enjoy bouncing up and down while on your lap?</td>
<td>13</td>
</tr>
<tr>
<td>Low pleasure</td>
<td>Amount of pleasure or enjoyment related to low stimulus intensity, rate, complexity, novelty, and incongruity.</td>
<td>How often during the last week did the baby enjoy being sung to?</td>
<td>13</td>
</tr>
<tr>
<td>Soothability</td>
<td>Reduction of fussing, crying, or distress when soothing techniques are used by the caregiver.</td>
<td>When rocking your baby, how often did s/he soothe immediately?</td>
<td>18</td>
</tr>
<tr>
<td>Falling reactivity/rate of recovery from distress</td>
<td>Rate of recovery from peak distress, excitement, or general arousal; ease of falling asleep.</td>
<td>When put down for a nap, how often did your baby settle down quickly?</td>
<td>13</td>
</tr>
<tr>
<td>Cuddliness</td>
<td>Expression of enjoyment and molding of the body to being held by the caregiver.</td>
<td>When being held, how often did the baby mold to your body?</td>
<td>17</td>
</tr>
<tr>
<td>Perceptual sensitivity</td>
<td>Detection of slight, low intensity stimuli from the external environment.</td>
<td>How often does the infant look up from playing when the telephone rang?</td>
<td>12</td>
</tr>
<tr>
<td>Sadness</td>
<td>Lowered mood and activity related to personal suffering, physical state, object loss, or inability to perform a desired action; general low mood.</td>
<td>At the end of an exciting day, how often did your baby become tearful?</td>
<td>14</td>
</tr>
<tr>
<td>Approach</td>
<td>Rapid approach, excitement, and positive anticipation of pleasurable activities.</td>
<td>When your baby saw a toy s/he wanted, how often did s/he get very excited about getting it?</td>
<td>12</td>
</tr>
<tr>
<td>Vocal reactivity</td>
<td>Amount of vocalization exhibited by the baby in daily activities.</td>
<td>How often did your baby make talking sounds when you talked to him/her?</td>
<td>12</td>
</tr>
</tbody>
</table>

2.4. Coding

2.4.1. Preferential looking task at 6 and 9 months of age

Infants’ total looking time at the monitors was coded for the 4-min duration of the tasks. Infants were classified as Cs if the preferential looking task could be conducted without interferences. They were classified as NCs if the preferential looking task had to be terminated for various reasons. The group of NCs was subdivided into non-completers due to fussiness (NC_fuss) and non-completers due to other reasons (NC_other). Infants were judged as fussy if they cried, were fearful, or if they were too restless to remain seated facing the monitors. If infants did not want to sit alone in the infant seat, they were also judged as fussy. In the preferential looking task at 9 months of age, an additional criterion applied for the NC_fuss group: If infants tried to turn away their leg to hinder stroking, they were also judged
as fussy. Technical problems, disturbances or interruptions by siblings or parents were counted as other reasons.

### 2.4.2. Habituation and novelty preference task

During the habituation task, the experimenter coded the looking times of the infant at the monitor in an infant-controlled habituation procedure. The presentation and coding of the habituation stimulus was conducted via Presentation® (version 15.0 build 02.08.12, Neurobehavioral Systems, Inc.). If infants looked away from the monitor for more than 2 s, the current habituation trial was terminated. The mean duration of looking time during the first two trials defined the baseline. The habituation phase was terminated when two consecutive looks of the infant reached on average less than 50% of the baseline (habituation criterion, see Bornstein, Arterberry, Mash, & Manian, 2011 for a similar procedure). The number of habituation trials the infant needed to reach the habituation criterion was used as habituation speed. If infants did not reach the habituation criterion within 20 trials they were excluded from habituation speed analysis because habituation speed could not be calculated.

Infants who reached the habituation criterion were presented with the novelty preference task. Based on the video recordings, an observer coded how long infants looked at each stimulus using the software Interact® (version 9.3.5, Mangold Software & Consulting GmbH, Arnstorf, Germany) for a total of 30 s. For further analyses, the proportion of looking time at the new stimulus was calculated (proportion of looking time at new stimulus = looking time at new stimulus/total looking time at both stimuli). Intraclass correlation for interrater reliability for the proportion of looking time at the new stimulus was $r = .77$.

Infants were classified as Cs if they reached the habituation criterion within 20 habituation trials or if they participated in 20 habituation trials without reaching the habituation criterion. Additionally infants were classified as Cs if they were able to take part in the novelty preference task for 30 s. Infants were classified as NCs if they failed to complete the habituation or the novelty preference task. The group of NCs was subdivided
into non-completers due to fussiness (NC_fuss) and non-completers due to other reasons (NC_other). Infants were judged as fussy if they cried, were fearful, or if they were too restless to remain seated facing the monitor either during the habituation task or the novelty preference task. If infants did not want to sit on their parents lap anymore, they were also judged as fussy. Technical problems, disturbances or interruptions by siblings or parents were counted as other reasons.

2.4.3. BSID-III at 9 months of age

The coding of the BSID-III followed the guidelines in the administration and technical manual (Bayley, 2006). An independent second observer coded 41 randomly chosen videotaped assessments of the BSID-III. Intraclass correlation for interrater reliability was high \( r = .93 \).

2.4.4. IBQ-R

The scoring of the IBQ-R consisted of averaging the items from each subscale to scores of that subscale. Items for which the parent marked “does not apply” as well as omitted items were counted as missing values. Questionnaires with more than 20% of missing items were excluded. Missing values from questionnaires with less than 20% of missing items were replaced with the mean of the subscale.

2.5. Analyses

In order to assess differences between Cs and NCs_fuss, a MANOVA and \( t \)-tests were used. For some analyses, the groups were too small to apply \( t \)-tests, and Mann-Whitney \( U \)-tests were therefore conducted instead. In order to assess the association between being C or NC_fuss in the preferential looking task at 6 months of age and at 9 months of age, Cohen’s kappa was used. The reasons for dropout among the group of NCs_other were assumed to be unsystematic and unrelated to temperament or cognitive developmental status. NCs_other were therefore not included in the analyses.
3. Results

3.1. Dropout rates

The number of Cs, NCs and its subgroups in the preferential looking tasks at 6 and 9 months of age, the habituation and novelty preference task are depicted in Table 2.

<table>
<thead>
<tr>
<th>Dropout status</th>
<th>Preferential looking task at 6 months of age</th>
<th>Preferential looking task at 9 months of age</th>
<th>Habituation and novelty preference task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completers (C)</td>
<td>96 (90.0%)</td>
<td>73 (68.9%)</td>
<td>84 (79.3%)</td>
</tr>
<tr>
<td>Non-completers (NC)</td>
<td>10 (9.4%)</td>
<td>33 (31.1%)</td>
<td>22 (20.7%)</td>
</tr>
<tr>
<td>NC_fuss</td>
<td>9 (90.0%)</td>
<td>31 (99.0%)</td>
<td>8 (28.6%)</td>
</tr>
<tr>
<td>NC_other</td>
<td>1 (10.0%)</td>
<td>2 (6.1%)</td>
<td>14 (63.6%)</td>
</tr>
</tbody>
</table>

3.2. Age, sex, and dropout

Cs and NCs_fuss did not differ in terms of age or sex in the preferential looking tasks at 6 and 9 months of age, the habituation and novelty preference task (all $p_s > .308$).

3.3. Temperament

The mean scores of the IBQ-R subscales and their standard deviations for the total sample are given in Table 3.

3.4. Temperament and dropout

3.4.1. Temperament and dropout in the preferential looking task at 9 months of age

The mean scores of the IBQ-R subscales and their standard deviations for Cs and NCs_fuss in the preferential looking task at 9 months of age are given in Table 3. A
MANOVA revealed a marginally significant effect of dropout on the 14 IBQ-R subscales (Pillai’s Trace = 0.2, $F(14, 89) = 1.57$, $p = .10$). Subsequent $t$-tests comparing Cs and NCs_fuss in the preferential looking task at 9 months of age revealed differences in some subscales of the IBQ-R (see Table 3). NCs_fuss had a higher mean Distress to Limitations score, and a lower mean Cuddliness score than Cs. Finally, Cs had a higher mean Falling Reactivity score than NCs_fuss. No further differences between NCs_fuss and Cs were found.\(^2\)

### 3.4.2. Temperament and dropout in the habituation and novelty preference task

NCs_fuss had a higher mean Approach score ($M = 6.25$, $SD = 0.65$) than Cs ($M = 5.67$, $SD = 0.70$, $U = 505$, $p = .019$, $r = .24$). No further differences between NCs_fuss and Cs were found.

### 3.5. Cognitive developmental status and dropout

#### 3.5.1. BSID-III and dropout

The mean score of the Cognitive Scale of the BSID-III for all infants at 9 months of age was $M = 100.48$ ($SD = 9.44$, range = 75-125). Cs and NCs_fuss did not differ in their cognitive developmental status as measured via the BSID-III at 9 months of age ($t(101) = 1.628$, $p = .107$, $d = 0.35$, for dropout in preferential looking task at 9 months of age; $t(90) = 0.513$, $p = .609$, $d = 0.08$, for dropout in habituation and novelty preference task).

#### 3.5.2. Habituation speed and dropout in the preferential looking task at 9 months of age

The mean habituation speed did not differ for Cs ($M = 10.23$, $SD = 4.62$) and NCs_fuss ($M = 8.52$, $SD = 4.51$) of the preferential looking task at 9 months of age ($t(82) = 1.596$, $p = .114$, $d = 0.37$).

---

\(^2\) There is evidence of a relationship between Negative Affectivity and attention-related processes (e.g., Kochanska, Coy, Tjebkes, & Husarek, 1998; Matheny, Riese, & Wilson, 1985). Completers and non-completers of the preferential looking task at 9 months of age did not differ in their Negative Affectivity ($t(102) = −0.868$, $p = .389$, $d = 0.18$). The factor structure of the German version of the IBQ-R (two-factor solution: Surgency/Extraversion and Negative Affectivity, Vonderlin et al., 2012; Mink et al., 2013) which we replicated in the current study differs from the factor structure obtained in the original study (three-factor solution: Surgency/Extraversion, Negative Affectivity and Orienting/Regulation, Gartstein and Rothbart, 2003).
3.5.3. Novelty preference and dropout in the preferential looking task at 9 months of age

The mean proportion of looking time at the new stimulus in the novelty preference task did not differ for Cs ($M = 0.50$, $SD = 0.19$) and NCs_fuss ($M = 0.55$, $SD = 0.19$) of the preferential looking task at 9 months of age ($t(81) = -1.149$, $p = .254$, $d = 0.26$).

3.6. Dropout in the preferential looking task at 6 months of age and dropout in the preferential looking task at 9 months of age

A contingency table of the numbers of Cs and NCs_fuss in the preferential looking task at 6 and 9 months of age is depicted in Table 4.

<table>
<thead>
<tr>
<th>Preferential looking task at 9 months of age</th>
<th>Completers (C)</th>
<th>Non-completers (NC_fuss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completers (C)</td>
<td>71 (68.3%)</td>
<td>2 (1.9%)</td>
</tr>
<tr>
<td>Non-completers (NC_fuss)</td>
<td>24 (23.1%)</td>
<td>7 (6.7%)</td>
</tr>
</tbody>
</table>

Note: Two additional infants dropped out due to other reasons: one infant from the preferential looking task at 6 months of age and one infant from the preferential looking task both at 6 and 9 months of age.

Dropout at 6 months of age was associated with dropout at 9 months of age (Cohen’s kappa = .249, $p = .001$).

4. Discussion

The current study revealed that dropout\(^3\) from infant studies is related to certain aspects of infant temperament as measured via the IBQ-R. Dropout in a preferential looking task was positively related to the IBQ-R temperament scale Distress to Limitations, and negatively related to the scales Falling Reactivity and Cuddliness. Dropout in a habituation and novelty preference task was positively related to the IBQ-R scale Approach. Dropout in any of the conducted tasks was not related to cognitive developmental status as measured via the BSID-III, habituation speed and novelty preference. Dropout at 6 months of age predicted dropout at 9 months of age.

\(^3\) For reasons of readability the following section uses the term “dropout” instead of “dropout due to fussiness” and “non-completers” instead of “non-completers due to fussiness”.
4.1. Temperament and dropout

In the present study, completers differed from non-completers in their temperament. These relations can be explained by the specific task demands of the preferential looking time task and the habituation task.

Non-completers had higher scores than completers on the Distress to Limitations scale. In other words, non-completers were more likely to fuss and cry than completers in situations in which they are restricted (e.g., in an infant seat, during caretaking). Experimental setups in preferential looking tasks and habituation tasks require a certain period of time during which infants sit still and observe objects or video screens. It might be that infants who are generally prone to fuss and cry when they are restricted were consequently more likely to drop out in the present preferential looking task.

Non-completers had lower scores than completers on the Falling Reactivity scale. In other words, non-completers were slower to recover from distress or excitement than completers. Experimental setups in preferential looking tasks often require infants to quickly familiarize themselves with a new environment, the experimenter and the test situation itself. The ability of completers to recover faster from distress or general arousal than non-completers might have made completers more likely to complete the test session than non-completers.

Non-completers had higher scores than completers on the Approach scale. In other words, non-completers were more rapid in their speed of approach and excitement toward certain stimuli than completers. Experimental setups in habituation tasks often present the same stimuli for a certain period of time repeatedly. It might be that the group of non-completers became easily excited and then quickly disappointed and tired when a rather boring task with repeated presentation of the same stimulus was presented (Mink et al., 2013).

Non-completers had lower scores than completers on the Cuddliness scale. In other words, non-completers were less likely than completers to experience enjoyment when being
held by their caregiver. Experimental setups in preferential looking tasks often require infants to sit on their parent’s lap. The characteristic of completers to be more joyful in this situation than non-completers might additionally help completers to cope with the high demands of the test session. During the preferential looking task in the current study, infants sat in an infant seat, but parents were allowed to stroke their infants’ head. Although we did not control how often each infant was stroked, it might be that completers experienced more enjoyment when their head was stroked compared to non-completers.

Differences between completers and non-completers only concerned temperament scales (e.g., Distress to Limitations), which influence infants’ likelihood of dropping out but probably do not influence their actual task-specific performance (beyond dropout) in looking time study situations. There were no differences between completers and non-completers on the scales (e.g., Duration of Orienting) that might influence this task-specific performance. Our findings are in line with other findings in which non-completers were reported as being fussier than completers (Treiber, 1984; Wachs & Smitherman, 1985). However, we could not replicate other reported differences in temperament: Completers and non-completers in our study did not differ in their sadness and their ability to maintain orientation (as reported in Fagen, Ohr et al., 1987; Mink et al., 2013), in their activity, or their likelihood of smiling and laughing (as reported in Miceli et al., 1998). Our findings are also in contrast to studies that did not find differences in temperament between completers and non-completers (Vonderlin et al., 2008; Wachs & Smitherman, 1985). Differences in findings between the current study and previous studies might be explained by the use of different paradigms and stimuli.

One could assume that infants’ negative affect influenced dropout in the current study, as reported in previous studies (e.g., Eisenberg et al., 1993; Kochanska et al., 1998; Matheny et al., 1985), and that findings on a relationship between the scale Distress to Limitation and dropout and between the scale Falling Reactivity and dropout are therefore not surprising. The current findings cannot be explained by a relationship between general negative affect and
dropout: First, the German version of the IBQ-R has been repeatedly reported with a two-factor solution (Negative Affectivity and Surgency/Extraversion Mink et al., 2013; Vonderlin et al., 2008), and we replicated this two-factor solution. We found a positive relationship between the scale Distress to Limitation and dropout and a negative relationship between Falling Reactivity and dropout. These scales are only two of the four scales (for the English version, Gartstein & Rothbart, 2003) with loadings on the Negative Affectivity factor. The other scales which load on the factor Negative Affectivity (Sadness and Fear, Gartstein & Rothbart, 2003) are not related to dropout in our study. Infants’ general negative affect is therefore not related to dropout, but the specific demands of the test situation in the looking time task (especially the ability to keep calm in a restricted situation and the ability to quickly become familiar with a new situation) were associated with dropout from this task.

One could also assume that infants’ attention-related processes influenced dropout in the current study. This explanation is not very likely, because there was no relationship between dropout and Duration of Orienting, which can also be seen as an attention-related process. Therefore, we conclude that it is not infants’ attention-related processes but rather the specific demands of the test situation in the looking time task which influence dropout from this task.

To sum up, given the findings of the current study, a sample that only consists of completers in looking time studies is not representative in terms of some parts of the temperament of the participating infants. This lack of representativeness is specific to the demands of the test situation in the looking time study. Other temperament dimensions that are potentially more important for the infants’ performance, such as Duration of Orienting and Perceptual Sensitivity, do not differ between completers and non-completers.

4.2. Cognitive developmental status and dropout

There were no differences between completers and non-completers in the cognitive developmental status as measured via the BSID-III, habituation speed and novelty preference.
Our finding is in line with a study that reported no differences in habituation speed between completers and non-completers (Richardson & McCluskey, 1983). Thus, despite a considerable proportion of dropout in looking time studies, our findings suggest that the final samples are representative in terms of cognitive developmental status of the initial sample.

4.3. Dropout at 6 months of age and dropout at 9 months of age

Our findings support the hypothesis that the dropout of an infant at one test session is associated with dropout at another test session three months later. A possible explanation for this finding is that temperament is stable during infancy (Pedlow, Sanson, Prior, & Oberklaid, 1993; Rothbart, 1986) and, in turn, certain temperament profiles are prone to dropout in looking time studies. In fact, we showed that dropout is related to certain temperament dimensions and it is likely that these temperament dimensions also affected infants’ dropout at 6 months of age. This finding is contrary to studies that reported no association between the dropout rates for an interval of 6-12 months between two test sessions (Bathurst & Gottfried, 1987; Mink et al., 2013). For longitudinal studies with multiple time points, this finding is important: If large-scaled longitudinal studies with high economic costs are conducted, some researchers might consider excluding infants who fail to complete one test session from further test sessions. This rationale could be acceptable because the remaining sample is representative in terms of cognitive developmental status and temperament scales that probably have an influence on the ability to perform in looking time study situations (e.g., Duration of Orienting). However, the relation in this study between dropout rates at 6 and 9 months of age reported here is weak. Additionally, parents might be disappointed about the exclusion if they have made a commitment to bring their infant into the laboratory for multiple tests in longitudinal studies. Looking time tasks should be adapted to the temperament dimensions that were associated with dropout: Cuddliness, Falling Reactivity, Approach, and Distress to Limitations. For example, reducing the infants’ feeling of being restricted during the looking time task could also reduce infants’ fussiness. This adaption
would not only reduce dropout but would also enhance the representativeness of the task. Additionally, researchers might consider re-scheduling a test if the infant dropped out from a test session, although this might be problematic in studies in which practice influences the outcome. Another way in which to handle infants’ dropout is through statistical methods, such as data imputation or maximum likelihood estimation (see Schafer & Graham, 2002, for a review).

4.4. Limitations and future directions

The current study used a preferential looking paradigm and a habituation paradigm to assess infants’ dropout in looking time studies. Thus, the interpretation of the current findings is relevant to research methods that investigate infants’ looking time. Since looking time studies are widely used, especially among young infants, the current study is a further step towards answering the question of why some infants drop out of testing whereas others do not. Future research should also consider dropout in studies in which infants’ behavior is analyzed, for example, in imitation studies.

There are other stable characteristics besides temperament and cognitive developmental status in infancy, such as circumstances and type of labor, birth order or maternal sensitivity, which might also influence the infants’ ability to complete looking time studies. For example, infants with mothers who had received medication during labor were reported to be less likely to complete a test session including a habituation task and a free-play session at 2 months of age (Oates, 1998). Type of labor, the point in time at which mothers start to think of their fetus as a real person during pregnancy and the infants’ range of experienced emotions as described by the mother were also reported to have an influence on the number of completed trials and the likelihood of completing the test session. It might be fruitful to additionally consider these factors in future research on infant dropout.

The MANOVA revealed only a marginally significant effect of dropout on the 14 IBQ-R subscales, and several tests of differences between completers and non-completers among
the 14 subscales using t-tests revealed statistically significant results. Due to statistical significance above the conventional .05 level, the question arises whether we found the differences between completers and non-completers in the IBQ-R scales by sheer chance. It has been argued that first conducting a MANOVA helps to protect against inflating the type 1 error rate in subsequent multiple tests, but even the use of multivariate methods such as the MANOVA does not completely control for type 1 error probability (Huberty & Morris, 1989). It has been suggested that the choice between a multivariate method should be made in accordance with the research question, as the methods address different questions. One situation in which univariate methods may “be appropriate is when the research being conducted is exploratory in nature.” (Huberty & Morris, 1989, p. 303). As previous studies assessing whether completers and non-completers of test sessions differ in their temperament have produced mixed findings, our study was rather exploratory in nature. However, due to possible type 1 errors in the analyses, the current findings should be interpreted with caution.

The current study investigated infant temperament at 9 months of age. Future studies on the relation between infant temperament and dropout could investigate the development of infant temperament and its relationship with dropout longitudinally. This might shed further light on the question of how infant temperament influences dropout.

Infant characteristics like temperament and cognitive developmental status can be assessed by standardized tests. Changes in infants’ states, however, cannot be assessed by standardized tests although these changes might be mirrored in physiological variables such as heart rate, skin conductance (Ham & Tronick, 2006) and hormone levels such as alpha amylase and cortisol (Jansen, Beijers, Riksen-Walraven, & Weerth, 2010). The investigation of these variables that are sensitive to changes in the infants’ state might further illuminate the reasons for dropout.

4.5. Conclusion

We showed that completers and non-completers of a looking time study differ in terms
of temperament but not in terms of cognitive developmental status. Moreover, there was
evidence of a stability of dropout between 6 and 9 months of age. These findings are
important in three respects: First, completers of looking time studies are better able to cope
with being held in one position and are more interested in novel events than non-completers.
Other temperament dimensions such as Duration of Orienting and Perceptual Sensitivity did
not differ between completers and non-completers. This temperament profile suggests that
non-completers did not feel comfortable in the test situation, but were not limited in their
attentional capacities. Second, dropout is stable across time. This could be taken into
consideration in high-cost, longitudinal studies, when infants are tested multiple times. Third,
dropout in looking time studies does not limit the representativeness of the final sample in
terms of cognitive developmental status. This is important information since dropout
sometimes occurs at high rates in looking time studies, raising concerns that it might
compromise the representativeness. In sum, we showed that dropout is not a major concern
for the representativeness of the final sample of looking time studies in task-relevant
characteristics of development.

Acknowledgments

The authors would like to thank Sina Hamester, Tessa Heinrich, Hannah Hermschulte,
Nadja Herten, Sina Hulten, Stefanie Kleinschmidt, Nathalie Marcinkowski, Milena Meyers,
Erich Molz, Susanne Röttgers, and Paula Siegmann for their assistance with the recruitment
of infants, data collection and data analysis. We also thank all of the infants and parents who
participated in this study.

This research was supported by the Deutsche Forschungsgemeinschaft (ZM54/2-1).
References


http://dx.doi.org/10.2307/1130552


http://dx.doi.org/10.1016/S0163-6383(02)00118-2

http://dx.doi.org/10.1016/j.infbeh.2010.10.002

http://dx.doi.org/10.1016/j.infbeh.2008.10.007

http://dx.doi.org/10.2307/1129273


http://dx.doi.org/10.2307/1130098


http://dx.doi.org/10.1016/j.cognition.2011.03.001
4 General Discussion

The current dissertation aimed at investigating the importance of infants’ early experiences with different degrees of intermodal matches as a precursor of infants’ building of a representation of the self. Their ability to differentiate between the self and others was measured via self-recognition tasks in the second and third year of life. Findings of study 1 indicated that infants’ intermodal match preference at 6 and 9 months was not predictively related to their mirror self-recognition status at 18 and 26 months or to video self-recognition status at 26 months. Moreover, infants’ intermodal match differentiation at 6 and 9 months was not predictively related to their mirror and video self-recognition status at 18 and 26 months. Overall, the absence of predictive relationships between intermodal match preference and differentiation and later self-recognition in study 1 supports the assumption that the representation of the self in the imagination develops independently of infants’ ability to differentiate between different degrees of intermodal matches. Infants’ intermodal match differentiation in study 1 was measured in a task that used self-generated behavior, which resembles a rather nonsocial stimulus. A previous study used another approach and showed that infants’ intermodal match preference, which was measured in a social interaction (a variation of the Still-Face Task; Tronick et al., 1978), at 9 months predicted mirror self-recognition. This study indicated the importance of intermodal match experiences in communication and interaction between parents and their infants, whereas intermodal match experiences regarding nonsocial stimuli like self-generated behavior do not seem to contribute to the development of a representation of the self with a comparable impact, as indicated by the findings of study 1.

The findings of study 2 indicated that infants’ preference for self-generated perfect matches as opposed to a preference for no matches is associated with parent-reported difficulties in the interaction between parent and infant. This finding supports the assumption that infants’ ability to differentiate between different degrees of intermodal matches is related
to the differentiation between the self and others (Bahrick & Watson, 1985; Rochat & Striano, 2000). Infants with a preference for perfect matches (which were self-generated in the current investigations) pay more attention to their own as opposed to others’ behavior, which might result in interaction difficulties with others. This finding highlights the importance of the experiences of intermodal matches in social interactions, and more crucially, the importance of infants’ attention to other-than-perfect intermodal matches in interaction with their parents.

Overall, the findings of study 1 do not support the assumption that the experience of self-generated intermodal matches plays an important role in the course of the development of the ability to differentiate between the self and others. By contrast, study 2 supports the assumption that the experience of intermodal matches in parent-infant interactions does play an important role in this ability (Bigelow, 2001). Parenting books and websites that advise parents to interact with their infants frequently and encourage them to imitate their infants in interaction and games in order to help infants to reach the milestone of self-recognition (Heuristic, 2016) might therefore guess right.

4.1 Relevance

The current dissertation project is the first attempt to define the nature of perfect, less-than-perfect and no matches in infants’ intermodal match perception, differentiation and preference. To this aim, the sources contingency, congruency, and timing of intermodal matches were identified. The schema developed in sections 1.2.2 and 1.2.3 offers the possibility to assort the different intermodal match preference tasks in order to make a comparison of these tasks and their findings among different age groups more comprehensible. This might be helpful for the development and classification of intermodal match preference tasks in the future.

Study 1 is the first empirical study to assess the question whether infants’ differentiation between different degrees of intermodal matches when viewing self-generated actions predicts later self-recognition. In the field of research on the development and nature of the
ability to represent the self in the imagination, the current findings contribute to theoretical frameworks on the development of self-recognition. One of these theoretical frameworks suggests a developmental trajectory which leads from the recognition of perfect intermodal matches in front of the mirror to later self-recognition, as measured via the rouge test (Rochat, 2003). This assumption cannot be supported by the findings of study 1. Another theory, by contrast, suggests that mirror self-recognition emerges rather independently of the experience of intermodal matches due to a cognitive change in the second year of life (Bischof-Köhler, 1991, 2012). This theory on the development of the self seems more likely.

The findings of study 1 stand in contrast to the only study that assessed infants’ differentiation of intermodal matches in social interactions, which showed it to be a predictor of later self-recognition (Kristen-Antonow et al., 2015). The difference in the measurement of infants’ intermodal match differentiation between Kristen-Antonow et al.’s (intermodal match differentiation of other-generated behavior in a social interaction) and our study (intermodal match differentiation of self-generated behavior), however, makes a direct comparison difficult. Nevertheless, these studies are the first empirical studies to assess the relationship between intermodal match differentiation and preference and later self-recognition in longitudinal study designs. A specific developmental pathway leading from intermodal match differentiation and preference to self-recognition seems unlikely. A representation of the self relies on more factors, such as infants’ experiences of intermodal matches in social interactions.

Moreover, the findings of study 1 might have implications for the field of robotics. Researchers in robot technology face the problem that robots have to be able to distinguish themselves from their surroundings (autonomous self-detection). Such autonomous self-detection is needed for robots that are able to change their so-called end effectors, meaning the device at the end of the robotic arm (i.e. a tool). Autonomous self-detection can enable the robot to detect which end effector is currently attached, to change the end effector according
to prospective tasks, and to detect malfunctions in end effectors in order to exchange them (self-repair, Stoytchev, 2011). Robots that used the comparison of efferent (motor commands) and afferent (visual movement) signals at pixel level were able to identify self-generated movements by direct self-observation and by observation of the movements in the mirror (Michel, Gold, & Scassellati). A recent study investigated whether a robot can learn to detect itself by analyzing the delay between efferent and afferent signals from self-observation (Stoytchev, 2011). It was shown that robots are able to estimate the efferent-afferent delay from such self-observation. Interestingly, the robot is able to estimate this “delay from data gathered while the robot performs motor babbling, i.e., random joint movements similar to the primary circular reactions described by Piaget.” (Stoytchev, 2011, p. 20). The findings of study 1, however, suggest that humans’ ability to represent the self in the imagination does not solely rely on the ability to compare efferent and afferent signals (intermodal matches). It might be that infants develop a new cognitive structure that allows them to build such a representation of the self. This might also be of interest for researchers in robotics.

Study 2 contributes to the aim of clinicians to achieve an earlier diagnosis of problems in infant development such as the developmental disorder autism spectrum disorder (ASD). It was shown that infants’ difficulties in interactions with their parents are related to their interest in perfect intermodal matches when viewing self-generated behavior. Infants’ interest in perfect intermodal matches could be an indicator of ASD-associated behavior early in life. However, future research, especially in longitudinal studies, is needed to further investigate whether this idea is a promising one.

Study 3 has important implications for researchers in the field of developmental psychology who frequently use looking time tasks. Completers and non-completers of looking time tasks differ in terms of temperament. Non-completers’ temperament profile suggested that they did feel uncomfortable in the test situation, but were not limited in their attentional capacities. Additionally, there was evidence that infants’ dropout from looking time tasks is
stable over time. I suggested that these findings could be taken into account in high-cost, longitudinal study designs when researchers wish to use looking time tasks. However, infants’ dropout from looking time studies does not limit the representativeness of the final sample in terms of cognitive developmental status. This information is not only important for the interpretation of studies 1 and 2 of the current dissertation, but also for other looking time studies in the field of infancy research.

4.2 Limitations

Studies in the field of infancy research have a common shortcoming: the lack of probability sampling. A random selection with all infants having equal probabilities of being chosen to take part in studies cannot easily be implemented. In the studies presented in the current dissertation, infants were chosen from a database of families who had previously agreed to be contacted and asked if they would like to take part in infancy research. Parents who agree to take part in such research are usually highly educated – as can easily be seen from the description of the sample in study 1. Most of the participating parents had either a university degree or university entrance-level qualifications (total for both levels of qualification: mothers 64.6%, fathers 65.5%). Additionally, parents often reported that they were interested in taking part in such studies and learning about the results of infant studies in the future. The willingness to take part in a longitudinal study and therefore the agreement to be contacted for several test phases over a longer period of time can be anticipated only for highly motivated parents. Due to such reasons, samples in infancy research cannot be regarded as representative of the whole population. Ideally, one could consider screening for abilities such as intermodal match differentiation and self-recognition within regular examinations like the ”U-Untersuchungen” (early health screening examinations) which are mandatory in Germany. Another approach to overcome this limitation could be to set group criteria concerning the representativeness of the sample and to invite only infants whose families fulfill these group criteria. However, realistically, the limitation of a lack of
probability sampling cannot be overcome due to practical reasons during recruitment of participating families, experimental procedures (especially standardization) and analyses, although true probability sampling in infancy research would be ideal.

Another major shortcoming which has already discussed in the Discussion section of study 1 is the lack of information about quality factors of intermodal match differentiation tasks and self-recognition tasks, such as their validity and reliability. To date, no study has investigated the validity or reliability of intermodal match differentiation tasks, which should clearly be focused on in future research. For the mirror self-recognition task, investigation of the retest reliability suggests that only some non-passers of the task are probably false negatives (Asendorpf & Baudonniere, 1993; Courage, Edison, & Howe, 2004; Nielsen & Dissanayake, 2004). Overall, investigation of the retest reliability of mirror self-recognition tasks points to an acceptable reliability (Kärtner et al., 2012). Nevertheless, no study has provided findings on the internal reliability of such tasks.

Findings regarding intermodal match differentiation rely on the interpretation of infants’ looking times to one of two video presentations of different degrees of intermodal matches. Not only is this task highly artificial, but infants also experience it with different degrees of intermodal matches, and their ability to differentiate between the two views is additionally assumed to be mirrored in their preference for either one or the other view. If this assumption is correct, and infants’ ability to differentiate between different degrees of intermodal matches can thus be reliably measured with such looking time tasks, this measurement is still an indirect one. However, intermodal match differentiation tasks like those used in the present studies are commonly used, and easy to implement and analyze. Nevertheless, this limitation highlights the importance of assessing infants’ experiences with intermodal matches during parent-infant interactions, which is one of the future directions of the current dissertation. Parent-infant interactions are not as artificial as the intermodal match differentiation tasks used in the present studies, as infants experience such interactions every day.
Additionally, the studies presented in the current dissertation each have focus-specific limitations which are discussed in the respective discussion sections and do not need to be repeated here.

4.3 Future Studies

The schema developed in sections 1.2.2 and 1.2.3 offers the possibility to assort different kinds of presentations in intermodal match differentiation tasks into three categories; perfect match, less-than-perfect match, and no match (see also Table 1). Using this schema, it is now possible to assort findings of the studies on intermodal match differentiation in a next step. Thus, the schema not only offers an overview and a classification of forms of presentations in intermodal match differentiation tasks, but will also provide an overview of the findings among different age groups resulting from the studies that used the different presentations outlined in Table 1. The development of such an overview and review of previous studies on infants’ intermodal match differentiation is one of the future directions of the current dissertation.

The emergence of a representation of the self occurs independently of the experience of intermodal matches as a rather sudden cognitive change in the second year of life, as suggested by the results of study 1. This cognitive change, allowing representations of the self in the imagination in the second year of life, might be based on infants’ experiences in interactions with their parents. Study 2 highlighted the importance of infants’ preference for perfect intermodal matches and the relationship of this preference with well-functioning parent-infant interactions. Overall, the findings suggest the relevance of the experience of intermodal matches in parent-infant interactions in the course of the development of self-recognition. Parents who provide their infants with nearly perfect intermodal matches might help them to reach the milestone of self-recognition earlier than infants of parents who provide their infants with no intermodal matches (Bigelow, 2001). First empirical evidence for this idea comes from a study that compared children from a culture with low parental
matching behavior to a culture with high parental matching behavior: Infants from the first culture recognize themselves in the mirror later than those from the second culture (Keller, Kärtner, Borke, Yovsi, & Kleis, 2005). In the course of the longitudinal study presented in the current dissertation, a parent-infant interaction was also conducted at 6 months of age, but has not yet been analyzed. Further evaluation of parental behavior in terms of the degree of intermodal matching and its relationship to self-recognition will provide clearer insights into the relevance of infants’ experience of intermodal matches in interaction and communication with their parents in their everyday life.

The findings of study 1 point to an emergence of a representation of the self, which occurs independently of the experience of intermodal matches as a rather sudden cognitive change, allowing a representation of the self in the imagination in the second year of life. The ability to build a mental representation of the self and to recognize oneself in the mirror is known to emerge in association with other abilities like object permanence (Bertenthal & Fischer, 1978), pretend play and personal pronoun use (Lewis & Ramsay, 2004), and self-conscious emotions like empathy (Bischof-Köhler, 1994) and embarrassment (Lewis et al., 1989). Additionally, it has been claimed that self-recognition is a prerequisite for the ability to infer the mental states of others (Gallup, 1998). In future large-scale longitudinal studies, it might be interesting and fruitful to take into consideration the above-mentioned previously found associations between infant abilities and the emergence of self-recognition, but also uncharted associations, in order to further examine the emergence of the self.

The development of self-recognition does not end with successful mirror and video self-recognition. At about 4 years of age, around 75% of children are able to recognize themselves when presented with a video of themselves with a delay of 2-3 minutes (Kristen-Antonow et al., 2015; Povinelli et al., 1996; Suddendorf, 1999). Delayed self-recognition from a video presentation is currently being analyzed in 54 four-year-old children from the longitudinal sample presented in the current dissertation. Of interest, for example, is the relationship
between the mirror self-recognition tasks at 18 months and the delayed self-recognition task at 4 years, and the relationship between video self-recognition at 26 months and the delayed self-recognition task at 4 years.

No study to date has investigated the quality factors of intermodal match differentiation tasks and self-recognition tasks such as their validity and reliability (see also section 4.2). This limitation points to obvious research questions for future studies. For instance, it could be interesting to assess the retest reliability of intermodal match preference and differentiation. Infants could be tested twice with a visual-proprioceptive intermodal match differentiation task (as used in study 1 at 6 months of age) with different delays between the first and second presentation of the task. Such a study design would require a change in the context of the task presentations, such as two different infant laboratories, different experimenters, and different apparatuses in order to avoid transfer effects from the first presentation of the task to the second.

Nowadays, mobile phones and digital cameras allow parents to easily take pictures and videos of their infants. Parents in our study reported that they use such devices and present the pictures and videos to their infants after they have been taken. Parents in our study also reported using video telephony such as Skype™ or FaceTime together with their infants in order to stay in contact with other family members who do not live nearby. It might be interesting to investigate whether this early experience and extensive practice of perfect intermodal matches via videos of the self as used for video telephony, and a high amount of immediate presentations of pictures of the self might influence the development of a representation of the self (see Bahrick et al., 1996 for a similar line of thought).
5 Conclusion

The current dissertation project was the first attempt to disentangle the variety of definitions of the different degrees of intermodal matches used in earlier theoretical frameworks and studies. A definition of the nature of perfect, less-than-perfect and no matches in infants’ intermodal match perception, differentiation and preference was presented along with a comprehensible schema. This schema offers the possibility to assort intermodal match preference tasks in order to compare these tasks and their findings.

This project provided important insights into the role of infants’ experience with intermodal matches in the first year of life for the development of their ability to recognize themselves in a mirror and on a video in the second and third year of life. Moreover, this is the first empirical, longitudinal approach to address this relationship and contributes to a large amount of theoretical frameworks on this relationship. I have highlighted the importance of taking into account more than one specific developmental pathway leading from experiences of self-generated intermodal matches to self-recognition. The development of a representation of the self relies on more than this. In this context, the proposed importance of experiences of intermodal matches in social interactions should be emphasized. The current dissertation project has implications for the field of infancy research, clinics, and even robot technology.
References


List of Figures

Figure 1. Schema about the nature of perfect matches, less-than-perfect-matches in intermodal match differentiation tasks. .......................................................... 15


Figure 1. Experimental setup in the visual-proprioceptive contingency task at 6 months of age. .......................................................... 38

Figure 2. Experimental setup in the visual-tactile contingency task at 9 months of age. .... 40

Figure 3. Experimental setup in the mirror-leg-recognition task at 18 and 26 months of age. For the mirror-face-recognition task at 18 months of age the mirror was completely uncovered and the chair was removed. The infants were allowed to explore the mirror freely. ............ 42

Figure 4. Experimental setup in the video-face-recognition task at 26 months of age. ...... 46


Figure 1. (a) Stimulus presented during habituation task. (b) and (c) Stimulus presented during novelty preference task, after infants reached the habituation criterion. ....................... 102
List of Tables

Table 1 Overview about forms of visual-proprioceptive presentations in intermodal match differentiation tasks resulting from manipulation of the sources of different degrees of intermodal matches (contingency, congruency, timing) .................................................. 17


Table 1 Number of Infants, Infants’ Sex, Infants’ Mean Age (Months;Days) with Standard Deviation in Parentheses (Days) and Age Range at 6, 9, 18, and 26 Months of Age Test Phases ................................................................. 36

Table 2 Minutes of Coding, Looking Behavior and Definition of Mark-Directed Behavior in the Mirror-Face-Recognition Task at 18 Months of Age (18-MFR), the Video-Face-Recognition Task at 26 Months of Age (26-VFR), and the Mirror-Leg-Recognition Task at 18 and 26 Months of Age (18-MLR and 26-MLR) ................................................................. 49

Table 3 Children’s Average Looking Time and Range of Looking Time to Mirror and Monitor in the 18-MFR, 18-MLR, 26-VFR, and 26-MLR ................................................................. 50

Table 4 Number of Children Who Completed and Percentage of Children Who Recognized Themselves in the Self-Recognition Task at 18 and 26 Months of Age .................................. 51

Table 5 Contingency Table of the Numbers of Children (with Percentage in Parentheses) Who Did or Did Not Recognize Themselves in the Mirror-Face-Recognition Task in Relation to whether the Children Did or Did Not Recognize Themselves in the Mirror-Leg-Recognition Task at 18 Months of Age ................................................................. 52

Table 6 Number of Infants Who Completed the BSID-III, Mean Score, Standard Deviations, Minimum Score, and Maximum Score for the Cognitive Scale of the BSID-III at 6, 9, 18, 26 Months of Age and Average Score ................................................. 53

Table 7 Correlation Matrix Displaying the Relations Between Contingency Preference Indices, Contingency Detection Indices in 6-VPC and 9-VTC, and Self-Recognition in the 18-MFR, 26-VFR, and 26-MLR. The Lower Left Part Represents Partial Correlations After Controlling for the Average Score of the Cognitive Scale of the BSID-III at 6, 9, 18, 26 Months of Age (Bayley, 2006) ................................................................. 55

Table 8 Results of Binary Logistic Regression Analysis of Mirror-Face-Recognition at 18 Months of Age ........................................................................................................................................... 56

Table 9 Results of Binary Logistic Regression Analysis of Video-Face-Recognition at 26 Months of Age ......................................................................................................................................... 57

Table 10 Results of Binary Logistic Regression Analysis of Video-Face-Recognition at 26 Months of Age ......................................................................................................................................... 58


Table 1 Summary of Associated Definitions of the Ages & Stages Questionnaire: Social-Emotional (ASQ:SE) Subscales, Item Examples, and Numbers of Items in the Subscales .. 81
Table 2 Correlation Matrix Displaying the Relations Between Proportion of Looking Time at the Delayed Video in the Contingency-Preference Task and the Scores of the Ages & Stages Questionnaire: Social-Emotional Scales and the Total Score, Respectively. The Lower Row Represents Partial Correlations After Controlling for Cognitive Development (Bayley Scales of Infant Development, 3rd ed.) ................................................................. 83


Table 1 Summary of associated definitions of the IBQ-R subscales, item examples and numbers of items in the subscales. ........................................................................................................ 104

Table 2 Numbers of infants (with percentages in parentheses) that completed or did not complete the preferential looking tasks at 6 and 9 months of age, the habituation and novelty preference task. ...................................................................................................................... 107

Table 3 Means (with standard deviations in parentheses) of the IBQ-R subscales for the total sample, and for Cs and NCs_fuss of the preferential looking task at 9 months of age and t-test results comparing the IBQ-R subscale means of Cs and NCs_fuss. ................................................. 107

Table 4 Contingency table of the numbers of infants (with percentages in parentheses) that did or did not complete due to fussiness the preferential looking task at 6 and 9 months of age. .................................................................................................................. 109
List of Publications

**PUBLICATIONS**


**TALKS**


**POSTER**


REVIEWS FOR:
Infant Behavior and Development
Social Development
Ruhr-Universität Bochum Research School PLUS
Curriculum Vitae

Sarah Klein-Radukic
Education, Psychology and Sociology
TU Dortmund University
Emil-Figge-Straße 50
44227 Dortmund
Germany

EDUCATION

Since 05/2012  Doctoral candidate
Supervisor: Prof. Dr. Norbert Zmyj, Prof. Dr. Axel Schölmerich
Faculty of Psychology, Applied Social and Developmental Psychology,
Ruhr-Universität Bochum

02/2011  Diplom
Supervisor: Prof. Dr. Jochen Musch
[„What do you think about minarets in Europe? The control of social desirability with the stochastic lie detector“], grade: sehr gut

10/2006 – 03/2011  Studies of Psychology,
Heinrich-Heine-Universität Düsseldorf

10/2005 – 03/2006  Studies of Mathematics and Pedagogy, Education science for Gymnasien and Gesamtschulen,
Universität Duisburg-Essen

06/2005  Abitur (grade: 1,8),
Don-Bosco-Gymnasium, Essen Borbeck

PROFESSIONAL EXPERIENCE

Since 07/2016  parental leave

05/2015 – 07/2016  Research assistant
Supervisor: Prof. Dr. Norbert Zmyj
Faculty of Education, Psychology and Sociology, Institute of Psychology, Developmental Psychology,
TU Dortmund University

05/2012 – 04/2015  Research assistant
Supervisor: Prof. Dr. Norbert Zmyj
Faculty of Psychology, Applied Social and Developmental Psychology,
Ruhr-Universität Bochum

04/2011 – 04/2012  Full-time lecturer
Supervisor: Prof. Dr. Heinrich Tröster
Faculty of Rehabilitation Sciences, Psychology in Rehabilitation,
TU Dortmund University
10/2009 – 12/2010  Student research assistant
Supervisor: Prof. Dr. Axel Buchner
Faculty of Mathematics and Natural Sciences, Department Psychology,
Cognitive and Industrial Psychology,
Heinrich-Heine-Universität Düsseldorf

PRIZES
1st poster prize Research Day, Ruhr-Universität Bochum, “Humanities and Social Sciences” section
Acknowledgements

An empirical work like the current dissertation is never the work of one person alone. The time has come to thank the people who have supported and encouraged me during the last years.

I am grateful for all the support I received from Norbert Zmyj. Thank you for being a great and encouraging mentor and for guiding me safely through herds of cats, stumbling blocks and pitfalls during this project.

I would like to thank Axel Schölmerich for providing helpful advice and support whenever necessary.

This project and dissertation would not have been possible without the reliability of the students and interns Marie-Christin Frerich, Irene Gettmann, Sina Hamester, Tessa Heinrich, Hannah Hermschulte, Nadja Herten, Sina Hulten, Stefanie Kleinschmidt, Nathalie Marcinkowski, Milena Meyers, Erich Molz, Susanne Röttgers, and Paula Siegmann. Thank you for your support with the invitation of the families, during data collection, and data analysis.

Special thanks to Katharina Kohl, my study buddy, for her moral support during the writing process and for her helpful comments on previous versions of this dissertation.

I owe my gratitude to my husband, for backing me all the way and for his persistence in recurrently finding encouraging words.

Finally, I would like to thank the infants and parents who participated in the studies; the interest and engagement shown by the families was fantastic.
Erklärung über den Umfang des eigenen Beitrags zu jeder anzurechnenden Publikation


Fragestellung: Norbert Zmyj
Datenerhebung: Sarah Klein-Radukic
Datenanalyse: Sarah Klein-Radukic
Kommentar und Korrektur vorläufiger Artikelversionen: Norbert Zmyj
Verfassen des Artikels: Sarah Klein-Radukic


Fragestellung: Norbert Zmyj
Datenerhebung: Sarah Klein-Radukic
Datenanalyse: Norbert Zmyj und Sarah Klein-Radukic
Kommentar und Korrektur vorläufiger Artikelversionen: Sarah Klein-Radukic
Verfassen des Artikels: Norbert Zmyj


Fragestellung: Sarah Klein-Radukic und Norbert Zmyj
Datenerhebung: Sarah Klein-Radukic
Datenanalyse: Sarah Klein-Radukic
Kommentar und Korrektur vorläufiger Artikelversionen: Norbert Zmyj
Verfassen des Artikels: Sarah Klein-Radukic
Erklärung über den Umfang des eigenen Beitrags zu jeder anzurechnenden Publikation
Eigenständigkeitserklärung


Sarah Klein-Radukic

Essen, 01. März 2017