CHILD DEVELOPMENT PERSPECTIVES

The Microstructure of Action Perception in Infancy: Decomposing the Temporal Structure of Social Information Processing

Gustaf Gredebäck[1] and Moritz M. Daum[2]

[1]Uppsala University and [2]University of Zurich

ABSTRACT—***In this article, we review recent evidence of infants' early competence in perceiving and interpreting the actions of others. We present a theoretical model that decomposes the timeline of action perception into a series of distinct processes that occur in a particular order. Once an agent is detected, covert attention can be allocated to the future state of the agent (priming), which may lead to overt gaze shifts that predict goals (prediction). Once these goals are achieved, the consequence of the agents' actions and the manner in which the actions were performed can be evaluated (evaluation). We propose that all of these processes have unique requirements, both in terms of timing and cognitive resources. To understand more fully the rich social world of infants, we need to pay more attention to the temporal structure of social perception and ask what information is available to infants and how this changes over time.***

KEYWORDS—*action prediction; action priming; attention; cognitive mechanisms; infancy; eye tracking*

Gustaf Gredebäck, The Uppsala Child and Babylab, Uppsala University; Moritz M. Daum, University of Zurich.

Both authors contributed equally to this article.

The research described in this article was funded by ERC-StG CACTUS 312292.

[The copyright line for this article was changed on 09 May 2016 after original online publication.]

Correspondence concerning this article should be addressed to Gustaf Gredebäck, Department of Psychology, The Uppsala Child and Babylab, Uppsala University, Box 1225, 751 42 Uppsala, Sweden or Moritz M. Daum, Department of Psychology, Binzmühlestr. 14/Box 21, CH-8050 Zürich, Switzerland. Electronic mail may be sent to gustaf.gredeback@psyk.uu.se or daum@psychologie.uzh.ch.

DOI: 10.1111/cdep.12109

We live in a dynamic world in which things change along dimensions such as time and space. People act on the world, people interact with each other, and objects move. Accordingly, when growing up in this dynamic environment, we must identify relevant agents and understand what these agents are doing. A series of processes occurs as we observe social events. Some processes take place over a few 100 ms (1), while others extend over many seconds (2). Knowing that different ways of measuring perceptual and cognitive processes operate on different time scales is not new (3), nor is the notion of sequential social processes (4). However, little attention has been devoted to studying the microstructure of action perception and action understanding, and to investigating how the combination of these processes facilitates infants' understanding of their social world.

Before any of these processes can be set in motion, social agents need to be detected and identified. From birth, infants are sensitive to other people's movements (5) and goal-directed actions (6), examples of how we initially detect social agents. If the detected agent performs an action that cues directionality, then babies can shift attention covertly in that direction (action priming); in some instances, babies can estimate action goals before the action is completed (action prediction). These processes help us monitor and understand ongoing actions performed by others. Once the observed action is completed, the outcome, or the manner in which the action was performed, can be related to our expectations about the observed action (action evaluation).

In this article, we integrate research on the component processes of action perception into a coherent framework, taking a holistic perspective on action perception early

Child Development Perspectives

in life and describing its temporal microstructure. In our walk along the timeline of action perception, we start with action priming, followed by action prediction and action evaluation.

ACTION PRIMING

Once an agent has been identified, observers adjust their attention in accordance with the direction of the other. When a person observes an agent look to the left, that person shifts his or her attention in the same direction (7). This sensitivity to the direction of others' actions is expressed not only by overt gaze shifts but by adjusting one's attention in the absence of an overt response; we attend more to a location in space even if we do not look at it. To capture this latter form of attention modulation, Posner (8) introduced a priming paradigm in which a centrally displayed *cue* (eyes looking to the left in the previous example) is followed by a peripheral *target*. Detecting a target is faster when the target appears along the direction of the cue (congruent condition; target appears on the left side in the previous example) compared with the opposite direction (incongruent condition; target appears on the right in the previous example). Posner proposed that this *priming effect* results from a *covert shift of attention* (internal orienting of visual attention preceding observable eye movements) in the direction indicated by the cue (8). Such priming effects have been observed in adults for a range of stimuli, including arrows, referential human manual gestures (9), and shifts of gaze (7).

Similar priming effects exist early in life. Infants detect targets that appear at locations that are congruent with the direction of human referential actions more rapidly than targets that appear in noncued locations. More specifically, when infants identify the direction toward which a gaze shifts (10) or a hand grasps (9, 11) or points (12, 13), they shift their attention in the same direction, resulting in the priming effect. This attentional shift in the direction of others' actions does not develop uniformly. Instead, the priming effect (in infancy) is strongest for actions that infants can perform. Priming is present at 3 months for gaze shifts (10), emerges at 5–7 months in response to a static illustration of a grasping hand (9), and appears around 12 months for an image of a pointing hand (12). When the hand is moving, priming with respect to pointing hands occurs at 4½ months (13). Nonsocial stimuli that move in the same way as the hand do not produce the same priming effect (11, 13).

Studies of the neural correlates of action priming have recorded similar effects. In electroencephalography studies, the priming effect, with respect to gaze, is indexed by temporal parietal event-related potential (ERP) components N290 and P400, both known to index social processing in early infancy (14). The observation of static hands depicting a grasping or pointing gesture yielded similar findings (15). With respect to grasping, the P400 component differentiates congruent and incongruent trials for 5- and 6-month-olds who are proficient at grasping, but not 4- or 5-month-olds who are less proficient at grasping. With respect to pointing, the P400 component starts to differentiate congruent and incongruent pointing between 6 and 8 months (16), and nears the adult response at 12 months (17).

Early in life, motion is apparently necessary for priming. For example, movement per se is the determining factor underlying priming of gaze direction in newborns (18), and a moving hand primed pointing gestures for 4½-month-olds (13). ERP correlates of priming from static hands emerge later, at 5 months for grasping and at 9 months for pointing (15, 16).

Action priming occurs rapidly. In adults, the largest action priming effect is evident when the central cue (e.g., a hand) precedes the peripheral target by about 300 ms (7, 19), illustrating that agent identification and action priming take place < 300 ms after onset of a social cue that directs attention. In infants, we know less about the exact duration of these processes; typically, much longer delays are used, such as 1,000 ms (9, 10) or 720 ms (12). In one recent study, action priming occurred as quickly as 100 ms after stimulus onset at the age of 4½ and 6½ months. While priming occurred across a broader time range in younger infants (100–500 ms), the time range was narrower in the 6½-month-olds where priming occurred with a stimulus onset asynchrony of 100 ms but not 500 ms (20). This is consistent with research on adults, which suggests that priming effects dissipate rapidly (e.g., 7, 19). Furthermore, ERP components differentiate congruent and incongruent reaching and pointing at ~400 ms after stimulus onset (15, 16). Taken together, these observations suggest that the time needed for infants to identify the central cue and shift attention to the side is somewhere between 100 and 500 ms. The latency periods of infants' saccadic eye movements range between 400 and 500 ms (21), suggesting that the latency to fixation on the primed object is driven largely by the delay of the oculomotor system.

This initial modulation of attention prepares infants for likely events, but does not in itself yield an overt response or further process higher order properties of actions, like goals or intentions. In some instances in which more information is available, action goals can be predicted using overt gaze shifts.

ACTION PREDICTION

Once an agent is detected and covert attention has been deployed in the direction cued by the agent, further social information processing can occur. When the appropriate information is available (often actions that correspond with the observer's own motor repertoire; 22), both the goal of an action (23) and the end point of an ongoing action (24) can be predicted using overt gaze shifts (22).

When observing someone else reach for and manipulate an object, adults disengaged from the hand just after the hand starts to move (at 38 ms; 25) and fixate on the goal of the ongoing action just before the hand reaches this location (at 150 ms; 25).

Fixating the goal before the hand arrives and disengaging rapidly from the hand while initiating action are two ways adults predict what will happen next.

In what we think is the only study (26) to measure disengagement time (when the individual's gaze moves away from the hand toward the goal) in infants, 10-month-olds moved their gaze from a moving hand ~300 ms after the hand started to move toward the goal. Most often, infants moved their eyes once, landing close to the goal 100–200 ms before the hand reached this position. This study highlights the rapid nature of social perception and explains the temporal structure of predicting action; babies needed only 300 ms of movement information to disengage and accurately predict goals.

We argue that the predictive disengagement from the hand toward the goal is supported by a covert relocation of attention (as described in the priming section earlier). Seeing the hand configuration and the direction of the arm prior to the initiation of the actual reaching action may be sufficient to cue directionality and prepare the action-perception system for upcoming actions. This connection between covert priming and overt prediction is not supported directly by empirical findings, but should be viewed as a core component of this proposal.

Once the hand starts to move, many sources of information are available to initiate predictive eye movement. Several studies have demonstrated that an infant's ability to perform the observed action is essential (26–28); infants generally predict goals only for actions they can perform. Other factors that contribute to an infant's ability to make predictive eye movements include prior visual experience with the observed events (24), saliency of the goal (29), and the effect of a distal goal on infants' interpretations of immediate, here-and-now actions (e.g., reaching for an object that will be displaced later versus reaching for an object that will be placed inside a container later, 23), individual or joint actions (30, 31), and interaction style during conversations (32, 33).

So far, we have reviewed three processes (agent identification, action priming, and action prediction) that depend on each other and are organized hierarchically. The two central components, action priming and action prediction, are rapid processes that operate on a time scale of a few hundred ms. They occur while the observed action is being performed and are constrained by the time and information available. Once an observed action is completed, plenty of time remains to evaluate the manner in which the action was conducted (34) and the goal achieved by the action (35). In the next section, we target the final process in the timeline of social information processing, action evaluation.

ACTION EVALUATION

Action evaluation typically takes place after an observed action has been terminated and the goal is achieved. Action evaluation is usually measured via infants' reactions to events that are in congruence with, or that violate, their expectations. This violation of expectancy is often referred to as a surprise reaction that results in increased attention toward unexpected events compared to expected events. This change in attention can be assessed via several dependent variables, such as looking time (e.g., 35, 36), pupil diameter (e.g., 37, 38), and ERPs (e.g., 39).

Much of our knowledge about infants' social perception stems from studies of action evaluation. For example, this research has demonstrated that infants encode the goals of both complete (35) and incomplete grasping actions (34, 40). Among several behaviors, infants recognize the goal-directedness of successful and failed reaching actions (41), and they recognize the goals of action sequences (42).

When compared with action priming and prediction, action evaluation tasks are not temporally demanding. As an example, infants in the seminal study by Woodward (35) were given up to 120 seconds (s) to detect an action and respond with surprise and enhanced looking time. However, this long response time was not necessarily related to the prolonged processing requirements of action evaluation, but rather to limits inherent to looking-time paradigms.

Recent advances allow us to pinpoint the temporal structure of action evaluation with higher resolution. For example, pupil diameter differs not only with respect to changes in luminance, but also with alterations in attention generated by changes in

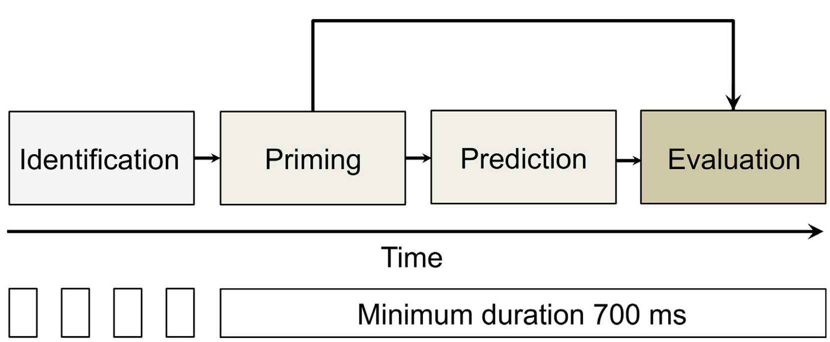

Figure 1. Suggested timeline of action perception and the minimum duration of the three component processes (priming, prediction, and evaluation) targeted in this review.

cognitive load, arousal, or action evaluation (38). When 4-, 6-, and 12-month-olds observed feeding actions that were unexpected (the food was not brought to the mouth but to the food recipient's hand), their pupils dilated more than when they observed expected feeding actions (the food was brought to the recipient's mouth; 37). These responses occurred within 3 s of the onset of an image or critical event in a movie (37).

Similarly, electroencephalographic responses to observed events indicate the cognitive processing of an observed event and the observer's reaction to whether an event is perceived as expected or unexpected. For example, in one study (39), babies were shown a sequence of images depicting actors picking up food and bringing it toward their mouths or their foreheads. Beginning at 9 months, infants showed a negative component in their ERPs at ~700 ms after stimulus onset that was significantly larger when the infants observed an unexpected compared to an expected action outcome. This finding suggests that infants do not need much time to evaluate observed events and relate them to prior expectations. Similar to action priming and action prediction, core components of action evaluation occur within the first second (39). At the same time, most measures of action evaluation require longer time spans for infants' reactions to be recorded. These findings fit well with the notion that action evaluation in infants occurs after the other processes mentioned earlier; action priming occurs after 100–500 ms and action prediction occurs within 300–500 ms of stimulus onset, whereas the earliest sign of action evaluation occurs 700 ms after the event (see Figure 1). Given that the microstructure of action perception can be segmented into distinct processes that are mutually dependent, the input for action evaluation may be richer than for the other two processes. The outcome of the action is observable and urgent time constraints are not imbedded in the process (the time available for action priming and action prediction is, by default, more limited).

CONCLUSION

In this article, we sought to decompose the timeline of action perception and propose a series of four distinct component processes that occur in sequence: agent identification, action priming, action prediction, and action evaluation. This proposal is derived from separate studies of each process and a belief that sequential social perception processes interact to create rich, complex, and time-dependent perceptions of the social world. The model proposed here has not been tested directly. We know that the four processes differ with respect to the amount of information that is, or is not, available. They all have different input and temporal requirements, and vary in the mechanisms used to process social information. To understand more fully the rich social world of infants, we need to pay more attention to the temporal structure of social perception and ask what information is available to infants and how this changes over time.

REFERENCES

1. Gredebäck, G., Johnson, S. P., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology*, *35*, 1–19. doi:10.1080/87565640903325758
2. Daum, M. M., Attig, M., Gunawan, R., Prinz, W., & Gredebäck, G. (2012). Action seen through babies' eyes: A dissociation between looking time and predictive gaze. *Frontiers in Cognition*, *3*, 370. doi:10.3389/fpsyg.2012.00370
3. Aslin, R. N. (2007). What's in a look? *Developmental Science*, *10*, 48–53. doi:10.1111/J.1467-7687.2007.00563.X
4. Reid, V. M., & Striano, T. (2007). The directed attention model of infant social cognition. *European Journal of Developmental Psychology*, *4*, 100–110. doi:10.1080/17405620601005648
5. Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 809–813. doi:10.1073/pnas.0707021105
6. Craighero, L., Leo, I., Umiltà, C., & Simion, F. (2011). Newborns' preference for goal-directed actions. *Cognition*, *120*, 26–32. doi:10.1016/j.cognition.2011.02.011
7. Langdon, R., & Smith, P. (2005). Spatial cueing by social versus nonsocial directional signals. *Visual Cognition*, *12*, 1497–1527. doi:10.1080/13506280444000805
8. Posner, M. I. (1994). Attention: The mechanisms of consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *91*, 7398–7403.
9. Daum, M. M., & Gredebäck, G. (2011). The development of grasping comprehension in infancy: Covert shifts of attention caused by referential actions. *Experimental Brain Research*, *208*, 297–307. doi:10.1007/s00221-010-2479-9
10. Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*, 131–134. doi:10.1111/1467-9280.00024
11. Wronski, C., & Daum, M. M. (2014). Spatial orienting following directional cues in infancy. *Developmental Psychology*, *50*, 2020–2029. doi:10.1037/a0037155
12. Daum, M. M., Ulber, J., & Gredebäck, G. (2013). The development of pointing perception in infancy: Effects of communicative verbal cues on covert shifts of attention. *Developmental Psychology*, *49*, 1898–1908. doi:10.1037/a0031111
13. Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, *15*, 426–435. doi:10.1111/j.1467-7687.2012.01139.x
14. Senju, A., Johnson, M. H., & Csibra, G. (2006). The development and neural basis of referential gaze perception. *Social Neuroscience*, *1*, 220–234. doi:10.1080/1747091060098979
15. Bakker, M., Daum, M. M., Handl, A., & Gredebäck, G. (in press). Neurophysiological correlates of action perception at the onset of grasping. *Social Cognitive and Affective Neuroscience*. Advance online publication. doi:10.1093/scan/nsu119 Retrieved from http://scan.oxfordjournals.org/content/early/2014/09/05/scan.nsu119
16. Gredebäck, G., Melinder, A. M. D., & Daum, M. M. (2010). The neural basis and development of pointing comprehension. *Social Neuroscience*, *5*, 441–450. doi:10.1080/17470910903523327
17. Melinder, A. M. D., Konijnenberg, C., Hermansen, T., Daum, M. M., & Gredebäck, G. (in press). The developmental trajectory of pointing perception in the first year of life. *Experimental Brain*

Research. Retrieved from http://link.springer.com/article/10.1007/s00221-014-4143-2/fulltext.html

18. Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition*, *7*, 705–718. doi:10.1080/13506280050144399
19. Daum, M. M., & Gredebäck, G. (2011). Spatial cueing by referential human gestures, arrows and mechanical devices. *International Journal of Mind, Brain and Cognition*, *2*, 113–126.
20. Bertenthal, B. I., Boyer, T. W., & Harding, S. (2014). When do infants begin to follow a point? *Developmental Psychology*, *50*, 2036–2048. doi:10.1037/a0037152
21. Gredebäck, G., Örnkloo, H., & von Hofsten, C. (2006). The development of reactive saccade latencies. *Experimental Brain Research*, *173*, 159–164. doi:10.1007/s00221-006-0376-z
22. Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*, 878–879. doi:10.1038/nn1729
23. Gredebäck, G., Stasiewicz, D., Falck-Ytter, T., von Hofsten, C., & Rosander, K. (2009). Action type and goal type modulate goal-directed gaze shifts in 14-month-old infants. *Developmental Psychology*, *45*, 1190–1194. doi:10.1037/a0015667
24. Henrichs, I., Elsner, C., Elsner, B., Wilkinson, N., & Gredebäck, G. (2014). Goal certainty modulates infants' goal-directed gaze shifts. *Developmental Psychology*, *50*, 100–107. doi:10.1037/a0032664
25. Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*, 769–771. doi:10.1038/nature01861
26. Rosander, K., & von Hofsten, C. (2011). Predictive gaze shifts elicited during observed and performed actions in 10-month-old infants and adults. *Neuropsychologia*, *49*, 2911–2917. doi:10.1016/j.neuropsychologia.2011.06.018
27. Cannon, E. N., Woodward, A. L., Gredebäck, G., von Hofsten, C., & Turek, C. (2012). Action production influences 12-month-old infants' attention to others' actions. *Developmental Science*, *15*, 35–42. doi:10.1111/j.1467-7687.2011.01095.x
28. Melzer, A., Prinz, W., & Daum, M. M. (2012). Production and observation of contralateral reaching: A close link by 12 months of age. *Infant Behavior & Development*, *35*, 570–579. doi:10.1016/j.infbeh.2012.05.003
29. Henrichs, I., Elsner, C., Elsner, B., & Gredebäck, G. (2012). Goal salience affects infants' goal-directed gaze shifts. *Frontiers in Psychology*, *3*, 391. doi:10.3389/fpsyg.2012.00391
30. Fawcett, C., & Gredebäck, G. (2013). Infants use social context to bind actions into a collaborative sequence. *Developmental Science*, *16*, 841–849. doi:10.1111/desc.12074
31. Keitel, A., Prinz, W., & Daum, M. M. (2014). Perception of individual and joint action in infants and adults. *PLoS One*, *9*, e107450. doi:10.1371/journal.pone.0107450
32. Keitel, A., Prinz, W., Friederici, A. F., von Hofsten, C., & Daum, M. M. (2013). Perception of conversations—The importance of semantics and intonation in children's development. *Journal of Experimental Child Psychology*, *16*, 264–277. doi:10.1016/j.jecp.2013.06.005
33. Von Hofsten, C., Uhlig, H., Adell, M., & Kochukhova, O. (2009). How children with autism look at events. *Research in Autism Spectrum Disorders*, *3*, 556–569. doi:10.1016/j.rasd.2008.12.003
34. Daum, M. M., Vuori, M. T., Prinz, W., & Aschersleben, G. (2009). Inferring the size of a goal object from an actor's grasping movement in 6- and 9-month-old infants. *Developmental Science*, *12*, 854–862. doi:10.1111/j.1467-7687.2009.00831.x
35. Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34. doi:10.1016/S0010-0277(98)00058-4
36. Daum, M. M., Prinz, W., & Aschersleben, G. (2009). Means-end behavior in young infants: The interplay of action perception and action production. *Infancy*, *14*, 613–640. doi:10.1080/I52500 00903263965
37. Gredebäck, G., & Melinder, A. M. D. (2010). Infants' understanding of everyday social interactions: A dual process account. *Cognition*, *114*, 197–206. doi:10.1016/j.cognition.2009.09.004
38. Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science*, *7*, 18–27. doi:10.1177/1745691611427305
39. Reid, V. M., Hoehl, S., Grigutsch, M., Groendahl, A., Parise, E., & Striano, T. (2009). The neural correlates of infant and adult goal prediction: Evidence for semantic processing systems. *Developmental Psychology*, *45*, 620–629. doi:10.1037/A0015209
40. Daum, M. M., Prinz, W., & Aschersleben, G. (2008). Encoding the goal of an object-directed but uncompleted reaching action in 6- and 9-month-old infants. *Developmental Science*, *11*, 607–619. doi:10.1111/j.1467-7687.2008.00705.x
41. Brandone, A. C., & Wellman, H. M. (2009). You can't always get what you want: Infants understand failed goal-directed actions. *Psychological Science*, *20*, 85–91. doi:10.1111/j.1467-9280.2008.02246.x
42. Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition*, *95*, 1–30. doi:10.1016/j.cognition.2003.12.004