

7 Joint Attention, Communication, and Knowing Together in Infancy

Malinda Carpenter and Kristin Liebal

What is joint attention? Are we in joint attention when I see you looking at something and I follow your gaze and turn to look at it too? When we attend a talk or a performance, are we in joint attention with all the other audience members? Is an infant in joint attention when he alternates gaze between his mother and a toy? Some authors would answer “yes” to all these questions; others, however, would be more hesitant. Given how crucially important joint attention is in human social life, it is quite surprising that as a field we have not yet come to a full agreement on what, exactly, joint attention is. The term joint attention is used in different ways by different scholars from philosophers to developmental and comparative psychologists to neuroscientists to roboticists. Even scholars within the same field sometimes include different behaviors as evidence of joint attention and require different levels of underlying understanding for those behaviors. This is a problem not just for scientific consistency but also because it perpetuates debates that may not be necessary—for example, arguments about whether or not chimpanzees engage in joint attention. Often, the central disagreements in these debates essentially come down to the use of different definitions of joint attention (see Carpenter & Call, in press). Given rapidly increasing interest in joint attention, we think it is a good time to start moving toward a common agreement about what counts as joint attention behavior, and why.

In this chapter, we thus first argue for the adoption of a relatively conservative definition of joint attention—one that requires the coordination of attention in joint attention to be truly *joint*. To this end, we distinguish “true” joint attention—shared attention—from other related types of behaviors by highlighting the importance of both partners’ *knowing together* that they are attending to the same thing. We then explore when in development human infants can first be said to engage in joint attention of this type (for our answer to the question of whether other animals such as

chimpanzees can engage in joint attention, see Carpenter & Call, in press). After concluding that infants participate in joint attention in this way by around one year of age, we address in detail the question of *how* they might be able to do this, given the theoretical complexities involved in knowing something together with someone (mutual knowledge). We propose a simple approach to sharing attention (and other psychological states like goals, in joint action) that we think avoids some of the problems with the classical philosophical approach to mutual knowledge. We conclude with a brief discussion of the significance of joint attention and related behaviors.

What Joint Attention Is (and Is Not)

The classic definition of joint attention involves a triadic interaction in which two individuals coordinate attention to an object of mutual interest (e.g., Bakeman & Adamson, 1984). Across time, this definition has been pulled in different directions by different researchers. Some focus most on the fact that the two individuals are looking at the same thing. These researchers often define joint attention as something like “the intentional co-orientation of two or more organisms to the same locus,” with “at least one of the organisms” doing something intentionally so as to end up focusing on the same thing as the other (Leavens & Racine, 2009, p. 241). For them, gaze following is typically a good example of joint attention behavior. In contrast, other researchers focus more on the coordination aspect of joint attention and the *sharing* of attention. These researchers argue (as we also do below) that attending to the same thing that one’s partner is attending to is not enough for joint attention: in addition, it is crucial that both partners *know together* that they are attending to the same thing (e.g., Hobson, 2005; Tomasello, 1995). Different proponents of this view differ with regard to how, exactly, one might know something together with someone else; however, in this view both partners are (at least eventually) equally involved, and, we would argue, must actively share attention about the thing. This sharing is what makes joint attention joint, rather than just parallel, attention.

To illustrate the difference between these two approaches, imagine a gaze-following situation in which one individual sees another turn to look at something and, as a consequence, turns to look at it too. Both individuals are now simultaneously attending to the same thing, so many people call this joint attention. However, where is the jointness here? The individuals may be looking at the same thing, but they are not necessarily

doing so *together* (see also Moll & Meltzoff, this volume). Gaze following can be done in a unilateral, even exploitative manner. Importantly, the looker does not even need to know that the follower is present, much less that they are both looking at the same thing. Now imagine being in the audience at a movie. You and the other audience members are all watching the movie, and this is discernable to all concerned based on visual orientation, proximity, the context, and so forth. Yet, again, there is not necessarily any jointness here. If you watch the movie and go home afterwards without talking to anyone, in what way have you shared the experience with the other audience members? Doing something together is more than just doing it at the same time. That is, now imagine that at the movie, sitting on one side of you is the friend you came to the movie with, and sitting on the other side of you is a stranger. Your knowledge about the two people's cognition may be identical—you know what each of them can see, hear, and so forth (and they know this about you)—but your experiences with them are quite different. With your friend, the joint attention “channel” is open, and everything that happens between the moment you enter and exit the theater together is shared in a way that it is not with the stranger. As these examples show, many situations that typically have been thought of as joint attention situations can actually involve individual, parallel attention rather than truly joint, shared attention.

The most commonly used behavioral evidence of joint attention is gaze alternation between the object of interest and the eyes/face of the social partner. However, gaze alternation alone is not enough to establish the existence of joint attention. There are many situations in which one might look back and forth between an object and a social partner without coordinating attention with him or her: gaze alternation can be a sign of alternating or checking attention rather than of coordinating or sharing attention (Tomasello, 1995). Even a mutual look (i.e., eye contact) is not necessarily enough since two individuals might happen to check each other's focus of attention at the same time and accidentally make eye contact. A mutual “sharing look” is the best evidence, but few studies have looked for these (see Hobson & Hobson, 2007, for a notable exception).

Most researchers report that human infants begin engaging in joint attention by nine to twelve months of age. However, given all the difficulties that we have just discussed, what evidence do we really have that infants this young are engaging in true joint attention? Next, we review studies that are relevant to this question. There are two important criteria

to look for: (1) the motivation to share attention in the first place and (2) that the participants know together that they are sharing attention. We thus first report evidence of interactions whose sole purpose is to share attention about objects or events. Developmentally, the first such interactions are typically joint attention looks (i.e., episodes of joint attentional engagement), followed by declarative gestures such as shows and points (e.g., Carpenter, Nagell, & Tomasello, 1998). However, since there is little if any direct research on the degree of sharedness of joint attention looks or shows, in our review we focus on infants' production of declarative pointing, for which there is much direct evidence. Then, we review experiments aimed at determining whether infants recognize whether they have shared attention with others about something—the knowing together component of joint attention.

Evidence for Sharing Attention and “Knowing Together” in Infants

The theory is that with declarative gestures, sharing attention with the other person is the goal of the gesture—the end in itself (e.g., Gómez, Sarriá, & Tamarit, 1993).¹ However, there are other possible interpretations of these behaviors, so it is important to test empirically what infants are trying to do when they perform them. Liszkowski, Carpenter, Henning, Striano, and Tomasello (2004) directly tested several competing hypotheses about why infants point in declarative situations by eliciting declarative points from twelve-month-olds and manipulating an adult's reaction to those points. To test the “rich” view that infants point to share attention and interest with others, in one condition the adult reacted by engaging in joint attention with infants about the object (i.e., alternating gaze between infants and the object and commenting about the object interestedly). To test a leaner view that, instead, infants simply (and egocentrically) just want the adult's rewarding attention and emotions on themselves (Bates et al., 1975; Moore & D'Entremont, 2001), in another condition the adult responded with positive emotion to the infants only, ignoring the object completely. To test the possibilities that infants just wanted the adult to look at the object, or wanted nothing at all from her and were simply pointing for themselves, in two other conditions the adult reacted by just looking at the object or by ignoring infants' point, respectively. Infants' different patterns of responding across conditions indicated that they were only satisfied with the adult's response in the first, joint attention condition: when she reacted in any other way, infants were more likely to respond by repeating their point to the object within trials and

were also more likely to give up and stop pointing for the adult altogether across trials. Liszkowski and colleagues thus concluded that infants' points were an invitation to the adult to share attention and interest about the object.

Further support for this idea comes from a study by Liszkowski, Carpenter, and Tomasello (2007b), who showed that it is important to infants both (1) that the adult shares attention to the *specific* referent they are pointing to—and not just to some other random object in that general direction—and (2) that the adult shares infants' own attitude of interest to that object. In that study, the adult varied whether in response to twelve-month-olds' points he attended to the correct versus an incorrect referent nearby and whether he expressed interest or no interest in the object. We found again that infants were more satisfied when the adult reacted with a full joint attention response—attending to the correct object *and* sharing interest. If the adult misunderstood the referent of infants' point and “shared” attention and interest to a different object nearby, infants repeated their point to the original referent to direct his attention to it. If the adult reacted (even to the correct object) in an uninterested way, infants were more likely to stop pointing across trials.

Liszkowski, Albrecht, Carpenter, and Tomasello (2008) further showed that twelve-month-olds' declarative points are “premeditated” in the sense that infants take into account the recipient's attentional availability *before* pointing, waiting to point until the adult is visually attending to them and able to see their point (see also Franco & Butterworth, 1996, and Franco & Gagliano, 2001, for evidence of this in slightly older infants). As in Liszkowski and colleagues' other studies, this study also showed that infants work to repair misunderstandings and provide clarification when their message is not understood: when the adult reacted to infants' points by saying “Huh, what?,” infants repeated their point. Infants thus are willing to work to achieve joint attention with others. Indeed, just the fact that they turn away from interesting sights at all in order to share them with others highlights their very strong motivation to engage in joint attention, and strongly supports the idea that the sharing of attention and interest is an important end in itself.

Finally, evidence that this sharing takes place on a mental level, about mutually imagined objects (instead of just involving the adult physically turning and facing something), comes from the finding that twelve-month-olds can point declaratively about absent referents—objects that were previously present but which have now disappeared. Liszkowski, Carpenter, and Tomasello (2007a) showed that not only do infants point to the

location where a puppet previously had been but that they do so selectively and appropriately depending both on whether the adult had previously seen the puppet and how she had previously reacted to it (i.e., with interest or not). In summary, by twelve months of age, when (on average) infants have just begun pointing declaratively, there is much empirical evidence to support the idea that the function of infants' declarative points is to share attention and interest about objects and events with others.

There are also several recent studies that suggest that one-year-old infants know what they have shared or experienced together with whom. For example, Liebal, Carpenter, and Tomasello (2010) had infants share several toy ducks with one experimenter and then several teddy bears with another experimenter. Then, at test, infants entered a room with just one of the experimenters in which there were two pictures on the wall: one of a duck and one of a teddy bear. Infants most often pointed to the picture of the type of toy they had shared with the particular experimenter they were pointing for. They thus accurately tracked their shared experiences with specific individuals and used this when deciding what to point to for others declaratively.

Further evidence that one-year-old infants understand what they have shared with whom comes from studies on infants' use of shared experience to aid in their comprehension of others' ambiguous communicative acts. For example, Liebal, Behne, Carpenter, and Tomasello (2009) have shown that fourteen-month-olds interpret an adult's ambiguous point to an object ("There!") as relevant to their just previously shared activity with that adult. A control condition in which infants shared the activity with one adult and then another adult pointed ruled out the possibility that infants were responding egocentrically, based on what they themselves were just previously doing, instead of what they had just shared with the first adult (see also Ganea & Saylor, 2007, and Saylor & Ganea, 2007, for further evidence of this).

Moll, Richter, Carpenter, and Tomasello (2008) took this a step further by having fourteen-month-olds share (i.e., interact excitedly about) three objects with an adult in sequence, one of them in a special way (they encountered it several times on the way to the testing room). When later the adult gestured excitedly toward the three objects together on a tray and ambiguously requested, "Wow, look, can you give it to me please?!", infants gave her the object they had shared in a special way. (Note that since infants had shared all three objects with the adult, this study shows that they know not just *what* they have shared with others but also *how* exactly they have shared it.) To test the possibility that infants simply gave

the adult that object because it was special for them, individually (not because it was the one they had shared together), Moll and colleagues included a control condition in which infants shared the three objects with the adult exactly as before, one in a special way, but then a *different* adult ambiguously requested one of the objects. If infants were just choosing the special object because it was most interesting to them, they should have chosen it in this condition too, but they did not—they chose that object at chance levels. To test the possibility that infants gave the adult the special object because it was the object that was apparently special for *her*, in another control condition infants watched as the adult experienced the objects individually (again, one in a special way) and then requested one of the objects from them ambiguously. Again, in this condition infants chose the special object at chance levels. In summary, infants responded not based on what they themselves knew individually about the objects, nor on what the adult knew individually, but instead on what they knew together. Thus, there is evidence that by fourteen months of age, infants both have the motivation to share attention and interest with others as an end in itself and, importantly, know what they know or have shared together with others. They thus have the two most important components of truly joint joint attention.

From a philosophical point of view, the fact that infants are capable of knowing something together with someone else is quite remarkable (and perhaps, in some views, quite implausible) given all the theoretical complexities involved in mutual knowledge. We thus need to take a closer look at what “knowing together” involves to see how it could be possible for infants. This is what we do next.

The Mutual Knowledge Problem

The classic, philosophical approach to mutual knowledge involves recursive mind reading. In order to know something together with someone, the argument goes, each partner must know that you know it, and that you know that I know it, and that you know that I know that you know it, and so on infinitely (e.g., Lewis, 1969; Schiffer, 1972). The mind-reading part of this may not be a problem for one-year-old infants: there is plenty of evidence that they know what others know, at least at some level (e.g., Liskowski, Carpenter, & Tomasello, 2008; Tomasello & Haberl, 2003). The problem lies in the recursive part. How can a one-year-old infant possibly process recursive inferences of such complexity? Although to our knowledge there has been no research on this, it seems very unlikely that infants

(of any age) are capable of even just a couple of levels of the recursive thinking that this approach requires (in fact, many people doubt that even adults achieve mutual knowledge in this way in normal, everyday circumstances—the processing demands are just too high; e.g., Clark & Marshall, 1981). It is possible that there are simpler, more direct ways of achieving something similar. For example, Gómez (1994, p. 73) argues that eye contact establishes “attention contact,” that is, “attending to the attention of a person who, in turn, is attending to our attention,” much like two mirrors reflecting each other infinitely (see also Peacocke, 2005).

However, even if so, another (and more damaging) problem with the classic recursive approach is that it does not seem to us to be really joint in any meaningful way. As figure 7.1 shows, in contrast to the typical “joint attentional triangle,” it is basically two individual perspectives that never meet in the middle: each participant just assesses the knowledge state of the other (she knows that I know...) individually. This approach involves solitary, parallel assumptions about what other “normal” people must be able to see or hear instead of socially, actively, jointly coordinated perspectives about a common topic (see, e.g., Calabi, 2008, for more on this). It is not knowing *together*; it is each knowing what the other knows at the same time, and that is not quite the same thing.

Another popular approach to explaining mutual knowledge does capture the togetherness needed. That is, in the context of joint action and collective intentionality (which are highly relevant here as we will see below), Searle (e.g., 1990) has proposed that there is an irreducible, primitive “we-intention” (or, if we take the liberty of translating it to joint attention, “we-attention”), and Campbell (2005) has proposed a different type of nonreductionist, “relational” view of joint attention. However, a common

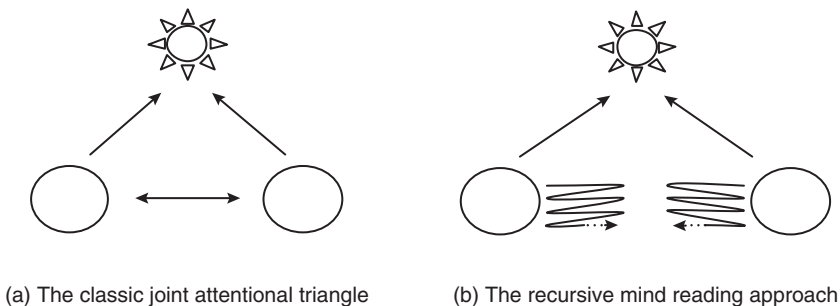


Figure 7.1

Depictions of different approaches to joint attention.

criticism of Campbell's and especially Searle's accounts is that they do not really spell out how this "we-ness" is achieved (see, e.g., Pacherie, 2007; Peacocke, 2005).

We thus need another approach, one that solves all of these problems at once: an approach that involves actual coordination of attention and knowledge, is simple enough for infants, and is achieved by a concrete mechanism we can "grab onto." We propose the outlines of just such an approach here. The details are not yet fully worked out, but we hope that it will spark fresh discussions about different levels or types of joint attention and about mutual knowledge and how it is achieved.

Sharing through Communication

We propose that to know something *together* with someone, both partners need to actively and openly *share* it. We further propose that the easiest and surest way to share something with someone is via *communication*. Whether it is verbal communication (e.g., "Isn't that great?") or just a meaningful, expressive look (something that is well within the capability of one-year-old infants; see figure 7.2 for an example) does not matter. Communication makes knowing something together instantaneous and effortless and simple enough for infants. It provides an indication (or confirmation or acknowledgment) that attention is shared and thus removes any doubt about whether the other saw or heard the thing too. One could even go so far as to argue that two individuals are not in truly joint or shared attention until they both signal—until they make it mutually manifest or public—to each other that they are (see Csibra, unpublished; Gilbert, 2007; Sperber & Wilson, 1986; Taylor, 1980, for more on the public or mutually manifest nature of communication).

Let us illustrate by going back to the movie example we discussed earlier. You know that the stranger sitting next to you is watching the movie, and he knows that you are watching the movie too. You each know that you are attending to the same thing. However, we would argue, you are not truly in joint attention yet—you do not yet know it *together*. What would make the difference—what would turn it into truly joint or shared attention—is if at some point you and the stranger turned to look at each other to smile about something one of the actors said or to remark on how good the movie was or the like. Only then have you truly shared something about the experience. Your look to the stranger is not just alternating attention, and it is not simply to see whether he is attending to your attention (which at that moment would be off the movie screen anyway). Its

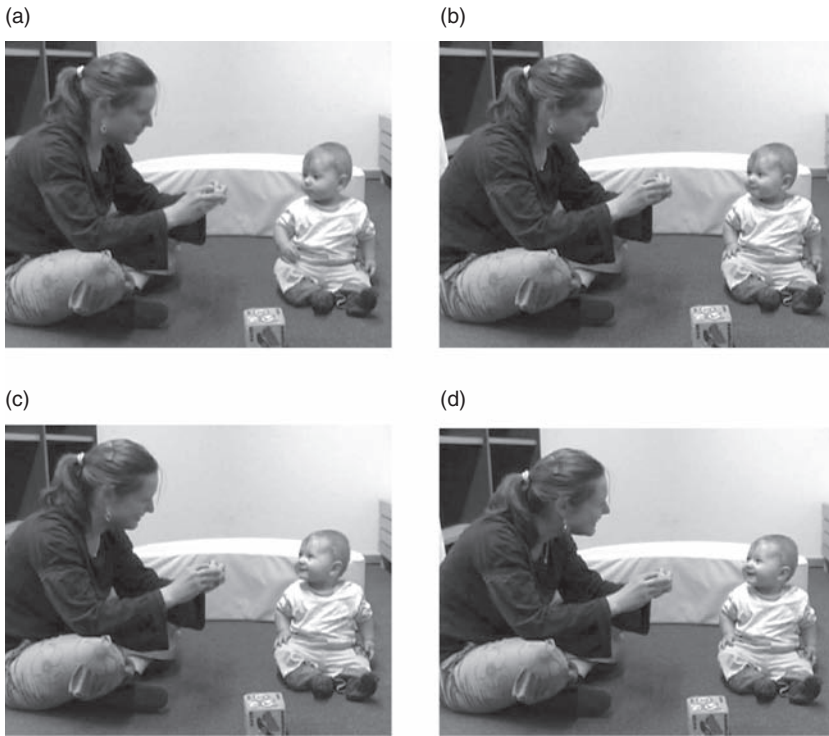


Figure 7.2

Sequential frames from a video of a nine-month-old initiating joint attention with his mother: (a) He watches as his mother makes a rubber duck squeak; (b) he smiles at the sound of the duck; and (c) and (d) he looks to his mother's face communicatively to share attention and interest with her.

function is to communicate something about what just happened, and in so doing to share attention to it. It changes the experience into something qualitatively different.

Communication turns a mutually experienced event into an *interaction*, into something joint. Before you looked to the stranger to share attention, there was something there—something more than the simplest type of parallel attention we discussed above in the context of gaze following since you each knew that you each knew that you were attending to the same thing (or at least could have figured this out if asked). There thus is some level of common or mutual attention just under the truly joint, shared level that we are talking about here (Michael Tomasello, personal communication). It is interesting that the contrast between this “not-yet-

shared” version of attention and the shared version reflects well the contrast between the two main approaches to joint attention in the literature. That is, one view usually describes joint attention in a “dry” and cognitive manner, simply in terms of who is looking at (and perhaps thinking about) what (e.g., Butterworth, 1995). Often the “knowing together” requirement is absent from these accounts, but even if it is there, something less than active sharing appears good enough to fulfill this requirement. The other view of joint attention emphasizes the richness of the interaction, the affective exchanges involved, and the direct, active participation of each partner. It captures the sharedness and the pleasure gained from the interaction (see, e.g., Hobson & Hobson, 2008; Mundy & Willoughby, 1996; Reddy, 2008; Trevarthen & Hubley, 1978). There are many hints in the descriptions of joint attention in these latter accounts that the authors recognize the inherent communicativeness and conversational nature of these interactions. However, here we are exploring a series of stronger claims, that is, not just that joint attentional interactions are in many ways like conversations but that the sharing of attention in “true” joint attention involves communication, that the sharing of psychological states in general involves communication (for more on this point, see below), and that this type of communication may offer a simple solution to the philosophical problem of mutual knowledge. So now let us get into some specifics concerning what we mean by communication here, again with a focus on joint attention in infancy.

Communicative Looks in Joint Attention

As we have noted above, the two most universally accepted joint attention behaviors in infants are (1) episodes of joint attentional engagement (i.e., episodes in which the two partners make eye contact about an object of mutual interest) and (2) declarative pointing and showing. Almost everyone agrees that declarative pointing and showing are communicative acts in one-year-olds, but few have considered infants’ looks to the face of the adult during episodes of joint attentional engagement to be communicative acts as well. We argue that they are—if they are certain kinds of looks. The distinction between sharing looks and checking and other types of looks (e.g., Carpenter, Tomasello, & Savage-Rumbaugh, 1995; Clyman, Emde, Kempe, & Harmon, 1986; Hobson & Hobson, 2007) is crucial here. Checking looks, which are just used to gather information unilaterally—to see or monitor where the other person is looking—correspond more to the “dry,” cognitive, recursive approach to joint attention. Here we are talking

about communicative looks, including sharing looks. We see infants' communicative looks as basically no more or less complicated than their first intentionally communicative gestures, which emerge developmentally around the same time and which themselves are supported by these looks.² These looks represent "real" communication (although, of course, it is very basic): they are intentional, they are referential, and they have content—they convey a message about the object or event (e.g., "Isn't that great?!"). Indeed we would go so far as to argue that, just like infants' pointing gestures, which have the reference and attitude components of speech acts (Liszkowski et al., 2007b), infants' communicative looks can have a topic-comment structure, as Bruner described decades ago:

The first and perhaps simplest form of comment is, I think, giving indication that a topic is being shared in joint action, and it is principally revealed in the child's management of gaze direction.... [The] child when involved in a transaction over some object or activity, looks up at some juncture and makes eye-to-eye contact with the mother, often smiling as well. The topic is the joint activity, the comment is the establishment of "intersubjective" sharing in connection with that activity, after which the activity goes on. (Bruner, 1975, p. 281)

These communicative looks can express different messages in different circumstances, depending on how the joint attention episode begins. An important (though sometimes slippery) distinction here is the distinction between "top-down" and "bottom-up" joint attention (Brinck, 2003; Kaplan & Hafner, 2006; Tomasello, 2008), here modified slightly. That is, depending on whether the referent of joint attention is introduced by one of the partners (top-down) or whether it calls attention to itself, to both partners at the same time (bottom-up), different types of communicative looks are involved.

In the *top-down* situation, the person who wishes to initiate joint attention actively directs the other person's attention to something. In this situation, three types of communicative looks are usually involved. The first type of look is an *initiation look* by the initiator to the recipient, which serves to get the recipient's attention. This look is an "invitation to interact" (Brinck, 2008; Bruner, 1983) and opens the channel of communication between the two partners. It signals the initiator's communicative intention ("I'm trying to tell you something"/"This is for you") and thus is usually accompanied by ostensive cues like saying "Hey!" and/or calling the recipient's name. The second type of communicative look is a *reference look* toward the object or event that the initiator wants to call attention to. It signals the initiator's referential intention (in Tomasello's, 2008,

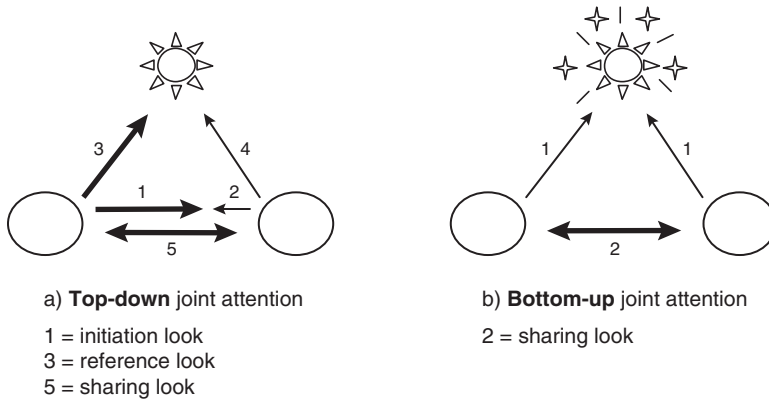


Figure 7.3

The sequence of looks in each type of joint attention (communicative looks are indicated by bold arrows and identified below).

terms) and is usually accompanied by a gesture like a point or nod toward the object. These two looks thus serve to open the joint attentional interaction and establish the topic or referent of it.

The third type of look in the top-down joint attention situation is the *sharing look*. If the initiation look serves to open the joint attentional triangle, the sharing look serves to close it (see figure 7.3). Whereas the initiation look is relatively one-sided (i.e., although its function is to make eye contact, it is produced by just one of the partners), the sharing look is bidirectional, with both partners participating equally (see the all-important double arrow in figure 7.1a and figure 7.3). This look is what turns parallel or recursive or not-yet-shared attention into truly joint, shared attention.

There is a lot packed into a sharing look. It is a confirmation or acknowledgment that attention is shared (“Yes, I see it too!”), as well as a comment on the just-established topic. It is in this comment that most of the communicative content of the look lies. The messages expressed in the comment can vary widely, but in the prototypical case the comment expresses an attitude about the referent that each partner hopes will be shared, in the sense of agreed with, by the other. Whether or not it is successful (since the participants’ attitudes might differ), this alignment of attitudes seems to be the goal of much joint attentional interaction. The attitude expressed is typically positive (this look is often accompanied by a smile) and can be glossed as something like “Wow, cool, huh?!” if both participants happen

to look at each other simultaneously or, depending on the precise timing of the sharing looks, something like the following (quick—almost simultaneous) conversational sequences:

[initiator:] “Isn’t it great?!” [recipient:] “Yeah!”

or

[recipient:] “Wow!” [initiator:] “Yeah, cool, huh?!”

Other messages are clearly also possible: they can be negative (“Ew, that’s gross, don’t you think?!”) or express something more complex, for example, “Yeah, it’s like the one we saw before!,” as in the Liebal et al. (2010) study on selective pointing to things that are relevant to one’s previous shared experience with one’s partner. In any case, it is interesting that these comments seem to include within them the assumed or hoped-for viewpoint of the other person toward the object of mutual attention, as expressed above by words like “Isn’t” and “huh?!.” It is also noteworthy that the referent is assumed as well, as captured in the above sentences by the use of a pronoun or no noun at all. This again suggests that there is some type of common or mutual but not-yet-shared attention that is brought about by the initial communicative act that drew the recipient’s attention to the initiator’s intended referent. This not-yet-shared attention can be the foundation on which the communication in the sharing look is based (cf. Calabi, 2008).

In summary, in the top-down case, three communicative looks are usually needed, to get the recipient’s attention, to establish the topic (i.e., identify the referent), and to comment on this referent, in the hopes of sharing attitudes about it.

In the *bottom-up* joint attention situation, in contrast, the referent draws attention to itself because of its salience—for example, there is a loud noise or sudden appearance or a strange and salient sight. In this situation, the referent is given by the context so no reference look (or gesture) is needed. Typically, only one communicative look is needed in the bottom-up situation: the sharing look to the partner (although the sharing look might be slightly more complicated in this situation than in the top-down situation because some initial communicative intent must also be contained in this look, since the channel of communication is being opened at the same moment as the “triangle” is being closed with the sharing look). The message contained in the sharing look in this case is thus something like “Hey, did you see that?! [i.e., not as a serious request for information but as an exclamation] Wow!” Again, as in the top-down case, a similar attitude

is assumed or hoped for, and the referent is also assumed (and confirmed and acknowledged with the sharing look), this time on the basis of the salience of it in the participants' mutual environment.

There is at least one other common situation that does not fit so well into the top-down and bottom-up categories. When infants are playing with their caregivers in a typical (in some cultures) joint attentional engagement situation (i.e., on the floor surrounded by toys), infants often produce sharing looks. These looks can happen right after something interesting happens, in a bottom-up-type situation (e.g., the block tower they are building falls down), but they also sometimes occur at seemingly random moments in the middle of play, for no apparent reason. There is thus no obvious bottom-up trigger for the look, and infants are not clearly attempting to establish a specific referent in these cases, so it is not quite like the top-down case either. Perhaps at these moments, infants are just looking to the adult to comment on the fact that they are sharing a fun experience ("Isn't this nice?!")—"to share the experience of sharing" in Hobson and Hobson's words (Hobson & Hobson, 2008, p. 79).

We just wish to make one final note about sharing looks. There is so much packed into these looks, and yet subjectively they feel utterly simple and directly perceived. Although it is remarkably difficult to provide an objective, operational definition of sharing looks (see Hobson & Hobson, 2007, for a valiant attempt), subjectively, it is immediately obvious when you are the recipient of one (e.g., Calabi, 2008; Reddy, 2008). (Note that, not coincidentally, you get a similar feeling when someone is trying to communicate with you—you know it instantly, as do even very young infants; Csibra, 2010.)

Sharing of Other Psychological States

We have proposed that an easy way of knowing something together with someone in joint attention is via communication: in effect, each partner "tells" the other, with just a meaningful look, that attention is shared. This analysis can, and, we think, should, be applied to joint action and other joint endeavors as well. That is, the same mutual knowledge requirement is present for the shared goals inherent in joint action: for truly joint joint action, it is not enough to act in parallel with someone or to be just mutually responsive. Instead, both partners must know together that they have the goal to act together (Bratman, 1992). Again, we would argue that this knowing together can be achieved most simply and directly by communication, by something as simple as a meaningful look (see Tuomela, 2005,

for a related view). In fact, communication plays an even greater variety of important roles in joint action than it does in joint attention—it is instrumental at every stage of the joint action, from the initial invitation and acceptance of the invitation (the establishment of the shared goal),³ to the negotiation of roles and resolution of any coordination problems, to the conclusion of the joint action (e.g., by leave-taking) at the end (Carpenter, 2009).

There is an interesting follow-on effect of communication that has been discussed in other contexts but that is relevant here as well. Communication makes things public and thus creates commitments and obligations (see, e.g., Gilbert, 2007; Sperber & Wilson, 1986; Taylor, 1980; Tuomela, 2005). The commitments and obligations inherent in joint (vs. parallel) action are clear (Gilbert, 1990)—even three-year-old children feel them (Gräfenhain, Behne, Carpenter, & Tomasello, 2009). There are surely similar effects in joint (vs. parallel or not-yet-shared) attention. For example, once you have shared something with someone, you cannot deny having seen or heard it, and once you have expressed a certain attitude about something in a joint attention situation, you cannot (easily) suddenly switch your attitude about it with that person (see Carassa & Colombetti, 2009).

Open Questions and Directions for Future Research

We realize that we have just scratched the surface with this account and that there are many aspects of it that still need to be worked out. For instance, we hope that philosophers will continue to think about sharing and how it might be achieved. One important question in this regard is whether there are any ways of truly sharing psychological states that do not involve communication. If so, are they within the capability of one-year-old infants? One possible candidate that comes to mind is imitation: Eckerman and colleagues have proposed that toddlers can share a topic nonverbally by imitating each other (see, e.g., Didow & Eckerman, 2001). However, we agree with Užgiris (1984, p. 25) that this type of imitation is inherently communicative. She argues that social imitation “is a means of communication with the partner. The basic message that imitation conveys is mutuality or sharing of a feeling, understanding, or goal.” It “serves to affirm a shared state” (see also, e.g., Didow & Eckerman, 2001; Nadel, Guérini, Pez , & Rivet, 1999; for similar views). Obviously, the question of what it means to communicate is very relevant in this context as well, and deserves much further consideration.

We also hope that psychologists will direct more research toward investigating different levels of joint attention. There are a number of claims that nonhuman apes and very young infants can engage in joint attention and/or joint action (e.g., Leavens & Racine, 2009; Pika & Zuberbühler, 2008; Striano & Bertin, 2005; Tanner & Byrne, 2010). However, there is thus far little evidence that either of these things is truly joint in these groups, in the way that we have characterized jointness here (see Carpenter & Call, in press). Future research should thus address whether the knowing together component of joint attention and joint action is present in these groups. Finally, it will also be helpful for future research to look for relations between communication, joint attention, and joint action in infants (as well as other theoretically related behaviors like social imitation, as Hobson & Hobson, 2007, have done with older children).

Conclusion

It is very telling that in philosophical discussions of the problems inherent in mutual knowledge, joint attention, joint action, and coordination problems, communication has to be explicitly and artificially stripped out of the situation for there even to be a problem at all (see the examples given in, e.g., Campbell, 2005; Schelling, 1960). When participants are able to communicate in certain ways, it is almost trivially easy to coordinate, share, and be sure (or make sure) that they are attending to or knowing or intending the same thing.

Here we have argued that communication can be a simple solution to the problem of mutual knowledge inherent in joint attention and other joint endeavors. More specifically, we have argued for the adoption of a conservative definition of joint attention and joint action, one that requires that participants know together with each other that they are attending or acting jointly. We have proposed communication—even as just a look—as a simple solution to the problem of how infants and, of course, adults could know something together with someone else: the simultaneity and at the same time mutual reactivity of participants' communicative, sharing looks to each other tells each of them immediately that attention or a goal is shared.

The importance of this type of communication, and of sharing and aligning attitudes and behavior in general, cannot be overstated in human development. Humans, apparently more than any other animals, have a strong motivation to share psychological states and experiences with others and to align themselves with and be like others (Carpenter, 2010;

Tomasello, Carpenter, Call, Behne, & Moll, 2005). The strength of this motivation is seen from the costs we are willing to bear to do these things. For example, even as infants and young children, we align our behavior with that of others by copying them even though this might result in less efficient performance on our part (e.g., Nagell, Olguin, & Tomasello, 1993), we choose to act jointly with others even though we then feel obligated to share the reward with them at the end (Hamann, Warneken, Greenberg, & Tomasello, submitted), and we miss out on parts of interesting sights in order to turn away to share them with others in joint attention. It is no wonder this motivation is so strong and so early to emerge in development; it underlies much of what is special about human social interaction and culture. Thus, to return to our initial question from a different perspective: What is joint attention? As the “primordial sharing situation” (Werner & Kaplan, 1963), it is no less than infants’ entry into a more deeply social world of interconnecting attitudes and experiences.

Acknowledgments

We thank Axel Seemann, Richard Moore, Jonathan Beier, Stephen Butterfill, and Federico Rossano for helpful comments and discussions. We are especially indebted to Michael Tomasello and Emily Wyman. They do not agree completely with our views, but our extensive discussions with them have greatly enriched our thinking on this topic.

Notes

1. Note that although Bates, Camaioni, and Volterra (1975) coined the term “protodeclarative gestures,” this term is typically now used in a different, social-cognitively richer way than in their original account: as gestures used to direct others’ attention to objects for the purpose of sharing attention and interest in them.
2. See also the literature on anticipatory smiles: infants’ communicative looks to adults while already smiling (for an example of an anticipatory smile, see figure 7.2). Anticipatory smiles increase significantly right around the time that participation in joint attentional engagement begins (and may in fact be the same thing as sharing looks). In support of the idea that they are communicative is the finding that anticipatory smiles are predicted by intentional gestural and vocal communication and intentional, means–end behavior (see, e.g., Jones & Hong, 2001; Venezia, Messinger, Thorp, & Mundy, 2004).
3. Note that in joint action there is also something similar to the top-down and bottom-up distinction we discussed above in relation to joint attention. The top-

down case involves one partner inviting the other to participate in a joint action; the bottom-up case involves some salient event (e.g., a heavy object falling on someone) that both partners immediately and simultaneously realize requires a joint action.

References

- Bakeman, R., & Adamson, L. (1984). Coordinating attention to people and objects in mother–infant and peer–infant interactions. *Child Development, 55*, 1278–1289.
- Bates, E., Camaioni, L., & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill-Palmer Quarterly, 21*, 205–224.
- Bratman, M. E. (1992). Shared cooperative activity. *Philosophical Review, 101*, 327–341.
- Brinck, I. (2003). The pragmatics of imperative and declarative pointing. *Cognitive Science Quarterly, 3*, 429–446.
- Brinck, I. (August, 2008). Joint attention in verbal and nonverbal communication. Presentation at the workshop on joint attention: Perspectives from philosophy, psychology, and social neuroscience. Bochum, Germany.
- Bruner, J. (1975). From communication to language—A psychological perspective. *Cognition, 3*, 255–287.
- Bruner, J. (1983). *Child's talk: Learning to use language*. New York: Norton.
- Butterworth, G. (1995). Origins of mind in perception and action. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 29–40). Hillsdale, NJ: Erlbaum.
- Calabi, C. (2008). Winks, sighs and smiles? Joint attention, common knowledge and ephemeral groups. In N. Psarros, H.-B. Schmid, & K. Schulte-Ostermann (Eds.), *Concepts of sharedness* (pp. 41–58). Frankfurt: Ontos Verlag.
- Campbell, J. (2005). Joint attention and common knowledge. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds* (pp. 287–297). Oxford: Clarendon Press.
- Carassa, A., & Colombetti, M. (2009). Joint meaning. *Journal of Pragmatics, 41*, 1837–1854.
- Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science, 1*, 380–392.
- Carpenter, M. (2010). Social cognition and social motivations in infancy. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development*, 2nd ed. (pp. 106–128). Oxford: Wiley-Blackwell.

Carpenter, M., & Call, J. (in press). How joint is the joint attention of chimpanzees and human infants? To appear in H. S. Terrace & J. Metcalfe, *Agency and joint attention*. Oxford: Oxford University Press.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4, Serial No. 255).

Carpenter, M., Tomasello, M., & Savage-Rumbaugh, S. (1995). Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees. *Social Development*, 4, 217–237.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.

Clyman, R. B., Emde, R. N., Kempe, J. E., & Harmon, R. J. (1986). Social referencing and social looking among twelve-month-old infants. In T. B. Brazelton & M. W. Yogman (Eds.), *Affective development in infancy* (pp. 75–94). Norwood, NJ: Ablex.

Csibra, G. (unpublished manuscript). *Why human infants follow gaze: A communicative-referential account*.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind & Language*, 25, 141–168.

Didow, S. M., & Eckerman, C. O. (2001). Toddler peers: From nonverbal coordinated action to verbal discourse. *Social Development*, 10, 170–188.

Franco, F., & Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language*, 23, 307–336.

Franco, F., & Gagliano, A. (2001). Toddlers' pointing when joint attention is obstructed. *First Language*, 21, 289–321.

Ganea, P., & Saylor, M. M. (2007). Infants' use of shared linguistic information to clarify ambiguous requests for objects. *Child Development*, 78, 493–502.

Gilbert, M. (2007). Mutual recognition, common knowledge, and joint attention. In T. Rønnow-Rasmussen, B. Petersson, J. Josefsson, & D. Egonsson (Eds.), *Hommage à Wlodek: Philosophical papers dedicated to Wlodek Rabinowicz* (pp. 1–21). www.fil.lu.se/hommageawlodek.

Gilbert, M. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, 15, 1–14.

Gómez, J. C. (1994). Mutual awareness in primate communication: A Gricean approach. In S. T. Parker, R. W. Mitchell, & M. L. Boccia (Eds.), *Self-awareness in animals and humans* (pp. 61–80). Cambridge: Cambridge University Press.

- Gómez, J. C., Sarriá, E., & Tamarit, J. (1993). The comparative study of early communication and theories of mind: Ontogeny, phylogeny, and pathology. In S. Baron-Cohen, H. Tager-Flusberg, & D. J. Cohen (Eds.), *Understanding other minds: Perspectives from autism* (pp. 397–426). New York: Oxford University Press.
- Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young children's understanding of joint commitments. *Developmental Psychology, 45*, 1430–1443.
- Hamann, K., Warneken, F., Greenberg, J. R., & Tomasello, M. (submitted). Collaboration encourages equal sharing in children but not chimpanzees.
- Hobson, J. A., & Hobson, R. P. (2007). Identification: The missing link between joint attention and imitation? *Development and Psychopathology, 19*, 411–431.
- Hobson, R. P. (2005). What puts the jointness into joint attention? In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds* (pp. 185–204). Oxford: Clarendon Press.
- Hobson, J. A., & Hobson, R. P. (2007). Identification: The missing link between joint attention and imitation? *Development and Psychopathology, 19*, 411–431.
- Hobson, P., & Hobson, J. A. (2008). Engaging, sharing, knowing. Some lessons from research in autism. In J. Zlatev, T. P. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 67–88). Amsterdam: Benjamins.
- Jones, S. S., & Hong, H. (2001). Onset of voluntary communication: Smiling looks to mother. *Infancy, 2*, 353–370.
- Kaplan, F., & Hafner, V. V. (2006). The challenges of joint attention. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems, 7*, 135–169.
- Leavens, D. A., & Racine, T. P. (2009). Joint attention in apes and humans. Are humans unique? *Journal of Consciousness Studies, 16*, 240–267.
- Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.
- Liebal, K., Behne, T., Carpenter, M., & Tomasello, M. (2009). Infants use shared experience to interpret pointing gestures. *Developmental Science, 12*, 264–271.
- Liebal, K., Carpenter, M., & Tomasello, M. (2010). Infants' use of shared experience in declarative pointing. *Infancy, 15*, 545–556.
- Liszkowski, U., Albrecht, K., Carpenter, M., & Tomasello, M. (2008). Infants' visual and auditory communication when a partner is or is not visually attending. *Infant Behavior and Development, 31*, 157–167.
- Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science, 7*, 297–307.

- Liszkowski, U., Carpenter, M., & Tomasello, M. (2007a). Pointing out new news, old news, and absent referents at 12 months of age. *Developmental Science*, *10*, F1–F7.
- Liszkowski, U., Carpenter, M., & Tomasello, M. (2007b). Reference and attitude in infant pointing. *Journal of Child Language*, *34*, 1–20.
- Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition*, *108*, 732–739.
- Moll, H., Richter, N., Carpenter, M., & Tomasello, M. (2008). Fourteen-month-olds know what “we” have shared in a special way. *Infancy*, *13*, 90–101.
- Moore, C., & D’Entremont, B. (2001). Developmental changes in pointing as a function of attentional focus. *Journal of Cognition and Development*, *2*, 109–129.
- Mundy, P., & Willoughby, J. (1996). Nonverbal communication, joint attention, and early socio-emotional development. In M. Lewis & M. W. Sullivan (Eds.), *Emotional development in atypical children* (pp. 65–87). Mahwah, NJ: Erlbaum.
- Nadel, J., Guérini, C., Pez  , A., & Rivet, C. (1999). The evolving nature of imitation as a format for communication. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy* (pp. 209–234). Cambridge: Cambridge University Press.
- Nagell, K., Olguin, R., & Tomasello, M. (1993). Processes of social learning in the tool use of chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology*, *107*, 174–186.
- Pacherie, E. (2007). Is collective intentionality really primitive? In M. Beaney, C. Penco, & M. Vignolo (Eds.), *Mental processes: Representing and inferring* (pp. 153–175). Cambridge: Cambridge Scholars Press.
- Peacocke, C. (2005). Joint attention: Its nature, reflexivity, and relation to common knowledge. In N. Eilan, C. Hoerl, T. McCormack, & J. Roessler (Eds.), *Joint attention: Communication and other minds* (pp. 298–324). Oxford: Clarendon Press.
- Pika, S., & Zuberb  hler, K. (2008). Social games between bonobos and humans: Evidence for shared intentionality? *American Journal of Primatology*, *70*, 207–210.
- Reddy, V. (2008). *How infants know minds*. Cambridge, MA: Harvard University Press.
- Saylor, M. M., & Ganea, P. (2007). Infants interpret ambiguous requests for absent objects. *Developmental Psychology*, *43*, 696–704.
- Schelling, T. C. (1960). *The strategy of conflict*. Harvard University Press.
- Schiffer, S. R. (1972). *Meaning*. Oxford: Clarendon.
- Searle, J. R. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 401–415). Cambridge, MA: MIT Press.

- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford: Basil Blackwell.
- Striano, T., & Bertin, E. (2005). Coordinated affect with mothers and strangers: A longitudinal analysis of joint engagement between 5 and 9 months of age. *Cognition and Emotion, 19*, 781–790.
- Tanner, J. E., & Byrne, R. W. (2010). Triadic and collaborative play by gorillas in social games with objects. *Animal Cognition, 13*, 591–607.
- Taylor, C. (1980). Critical notice: *Linguistic behavior* by Jonathan Bennett. *Dialogue, 19*, 290–301.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Erlbaum.
- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–735.
- Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what's new for other persons. *Developmental Psychology, 39*, 906–912.
- Trevarthen, C., & Hubley, P. (1978). Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In A. Lock (Ed.), *Action, gesture, and symbol* (pp. 183–229). London: Academic Press.
- Tuomela, R. (2005). We-intentions revisited. *Philosophical Studies, 125*, 327–369.
- Užgiris, I. C. (1984). Imitation in infancy: Its interpersonal aspects. In M. Perlmutter (Ed.), *The Minnesota symposia on child psychology: Vol. 17. Parent-child interactions and parent-child relations in child development* (pp. 1–32). Hillsdale, NJ: Erlbaum.
- Venezia, M., Messinger, D. S., Thorp, D., & Mundy, P. (2004). The development of anticipatory smiling. *Infancy, 6*, 397–406.
- Werner, H., & Kaplan, B. (1963). *Symbol formation*. Hillsdale, NJ: Lawrence Erlbaum.

