An Operant Analysis of Joint Attention Skills

Per Holth

Abstract

Joint attention, a synchronizing of the attention of two or more persons, has been an increasing focus of research in cognitive developmental psychology. Research in this area has progressed mainly outside of behavior analysis, and behavior-analytic research and theory has tended to ignore the work on joint attention. It is argued here, on the one hand, that behavior-analytic work on verbal behavior with children with autism needs to integrate the research body on joint attention. On the other hand, research on joint attention should integrate behavior-analytic principles to produce more effective analyses of basic processes involved. An operant analysis of phenomena typically considered under the heading of joint attention is followed by examples of training protocols aimed at teaching joint attention skills, such as social referencing, monitoring, gaze following, and such skills interwoven with mands and with tacts. Finally, certain research questions are pointed out.

Keywords: Joint attention, Language training, autism.

During the last 25 years, there has been an increasing preoccupation with ‘joint attention’ as a crucial area in children’s ‘social-cognitive development.’ Research has focused on normative patterns of emergence of joint attention skills (e.g., Corkum & Moore, 1995) and on how such skills are related to later developing skills summarized as ‘symbolic abilities’ (Hobson, 1993; Mundy, Sigman, & Kasari, 1993), ‘language abilities’ (Baldwin, 1995; Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Bruner, 1975; Tomasello, 1988), and ‘general social-cognitive processes in children’ (Baron-Cohen, 1995; Bruner, 1975; Mundy, 1995; Tomasello, 1995). Moreover, it appears that children diagnosed with autism may display a syndrom-specific deficit in joint attention skills (e.g., Baron-Cohen, 1989, Mundy & Crowson, 1997; Sigman & Kasari, 1995; Sigman, Kasari, Kwon, & Yirmiya, 1992). It seems strange, then, that behavior analysts, and even those working in the field of autism, have not paid much attention to the work on joint attention.

Research within the cognitive-developmental tradition has typically focused on identifying characteristic patterns of responding in different groups of children and on the consistency of responding across situations and over time (cf., Moore & Dunham, 1995). In spite of the fact that the whole body of “cognitive” research on joint attention focuses on behavior that needs to be analyzed in great detail, this field appears to have developed almost completely apart from behavior analysis. Some researchers have even argued specifically against behavior-analytic interpretations in this area (e.g., Bruner, 1995; Tomasello, 1995).

Recently, however, other researchers have occasionally called for some joint ventures of traditional joint-attention researchers and behavior analysts in an effort to develop intervention programs that might effectively remedy joint attention deficiencies in children with autism (e.g., Mundy, 2001; Mundy & Crowson, 1997).

The general aim of the present article is to show that an operant analysis is basically relevant to research on joint attention and to the aim of developing procedures that might help remedy basic deficiencies in joint attention typically displayed by children with autism. Specific aims of the current presentation are (1) to decompose the concept of joint attention sufficiently to make it amenable to an operant analysis, (2) to show how certain well-established basic behavioral processes can be utilized in interventions that attempt to correct deficiencies in joint attention skills in children with autism, and (3) to outline additional basic operant research that is needed in order to account for the variables of which joint attention skills are a function.

The concept of ‘joint attention’

Issues treated under the heading of ‘joint attention’ range from the early work by Bruner and colleagues on gaze following (e.g., Scaife & Bruner, 1975) to issues related to children’s so-called development of a “Theory of Mind” (e.g., Baron-Cohen, 1991; Mundy, Sigman, Ungerer, & Sherman,
1986). Therefore, ‘joint attention’ may not be particularly useful as a technical term unless the diversity of phenomena currently referred to by the concept prove to covary as a unitary phenomenon. In order to evaluate the concept as such, we need to analyze the different behavioral phenomena now listed as examples of ‘joint attention.’ Before turning to an overview of some of the phenomena typically treated within this realm, let me just briefly consider some attempted definitions.

**Defining ‘joint attention’**

According to Baldwin (1995), “technically speaking, joint attention simply means the simultaneous engagement of two or more individuals in mental focus on one and the same external thing” (p. 132). Very often, however, a lot more may be implied by ‘joint attention.’ Sigman and Kasari (1995) distinguished between a narrow and a broad definition of ‘joint attention.’ The narrower definition refers simply to “looking where someone else is looking.” Their broader definition includes what they called “responsive and initiating behaviors as well as the checking of another person’s face that occurs while the infant is playing with something, when the infant has accomplished some task, the infant has pointed to something, or in an ambiguous situation” (p. 189). Further, according to Bruner (1995), “joint attention involves knowing that another is looking at and experiencing something in the visual world” (p. 7). Tomasello (1995) includes in his definition that “both participants are monitoring the other’s attention to the outside entity,” and that “the coordination that takes place in joint attentional interactions is accomplished by means of an understanding that the other participant has a focus of attention to the same entity as the self” (p. 105-107). Finally, Sarria, Gomez, & Tamarit (1996) observed that although joint attention “may be achieved through other sensory modalities, such as vocalizations or physical contact” (p. 49).

In *A Preliminary Manual for the Abridged: EARLY SOCIAL COMMUNICATION SCALES* (ESCS), Mundy, Hogan, & Doehring (1996) suggested that “the function of [joint attention] behaviors is to share attention with the interactive partner or to monitor the partner’s attention. They differ from Requesting bids in that they do not appear to serve an instrumental or imperative purpose.” However, it is clear according to Corkum & Moore (1995), “joint attention plays an integral part in both the protodeclarative and protoimperative gestures” (p. 64).

Although empirical studies usually rely on some narrower operational definition of joint attention such as those specified in the ESCS by Mundy et al. (1996), people working in this area seem to agree that the concept of ‘joint attention’ implies something in addition to those operationalized skills. This additional implication has been described as “knowing that another is looking at and experiencing something in the visual world” (Bruner, 1995, p. 7), “understanding that the other participant has a focus of attention to the same entity as the self” (Tomasello, 1995, p. 107), or “the recognition that mental focus on some external thing is shared” (Baldwin, 1995, p. 132). Such “knowing,” “understanding,” or “recognition” is, of course, harder to specify. Dunham and Moore (1995) predicted that research will, eventually, lead to the decomposing of joint attention into “the series of transformations that are presumably occurring in social cognition across developmental time” (p. 23). From a behavior-analytic perspective, a decomposition may be required in order to bring the phenomena under investigation within reach of its scientific principles.

**Phenomena treated under the heading ‘joint attention’**

In order to take on the task of pinpointing the behavioral phenomena of which such phenomena can consist, let us start by delineating some of the cruder categories that are typically conceived of as involving ‘joint attention.’ These include ‘gaze following,’ ‘social referencing,’ ‘protoimperative gestures,’ ‘protodeclarative gestures,’ and ‘monitoring.’

(a) Look or gaze following. Perhaps the simplest examples of joint attention skills are those referred to as ‘responsive joint attention,’ in which one looks where someone else is pointing or touching (following proximal point/touch) or is looking in the direction of someone’s gaze or beyond the end of someone’s index finger (following line of regard; see Mundy, Hogan, & Doehring, 1996). However, not even in these cases is
joint attention a purely formal class. As pointed out by Tomasello (1995), there are two common types of adult—child interactions that may look like joint attention, but which lack the criterion of “knowing,” “understanding,” or “recognition.” These are “onlooking” and “cued looking,” in which something may catch the attention of two people simultaneously, but without one person’s attention influencing the other person’s attention.

(b) Social referencing. When confronted with some novel stimulus, a child (typically) will look toward a familiar person and subsequently react to the novel stimulus in accord with the displayed expression of the familiar person. As in the case of gaze following, however, joint attention in such ‘social referencing’ requires that the child “understands” that the familiar person is attending to the same thing or event as the child attends to. If, for instance, the child simply looks to a familiar person as a kind of ‘comfort seeking,’ this will not count as ‘joint attention’ (cf. Baldwin, 1995).

(c) Protoimperative. Protoimperative gestures have been described as “gestures intended to make another person do something for one’s benefit” (Sarriá et al., 1996). However, a simple contingency between a gesture and a ‘beneficial effect’ can occur without ‘joint attention.’ Sometimes, the term ‘protoimperative’ has been preserved for cases that involve some type of “coordination of attention with other people” (Sarria et al., 1996). In accord with this, Tomasello (1995) wrote: “My interpretation of protoimperative pointing in the 12- to 14-month period, therefore is that the child is attempting not just to obtain the object but to change the adult’s intentions so as that they become aligned with its own” (p. 111).

(d) Protodeclarative. Bates, Camaioni, and Volterra (1975) defined the protodeclarative as a preverbal effort to direct other’s attention to an object or event. Tomasello (1995) interpreted the protodeclarative as having “the purely social motive of sharing attention to something” (p. 111).

(e) Monitoring. Gaze or attention monitoring can take place in a simple responsive manner, as it certainly does when we are just observing other people, as in a movie. However, such monitoring can be interactive and involve acting to influence the other person’s attention. It appears that an emerging criterion for using the term ‘joint attention’ or ‘true joint attention’ in all types of cases mentioned above is exactly the interactive monitoring of another person’s attention. In developmental psychology, researchers have tried to capture the essence of ‘joint attention’ in social-cognitive terms.

Why behavior analysts should study joint attention

There are, at least, three specific, good reasons why behavior analysts should be interested in studying performances typically grouped under the heading of “joint attention.” We will explore each in this section.

First, because children diagnosed with autism seem to display a syndrome-specific deficit in joint attention skills (e.g., Baron-Cohen, 1989, Mundy & Crowson, 1997; Sigman & Kasari, 1995; Sigman, et al., 1992), outcome studies of applied work with children with autism should include measures of relevant joint attention skills. Until intervention studies (e.g., Lovaaas, 1987; McEachin, Smith, & Lovaaas, 1993) include outcome measures that address the cardinal social and social-cognitive symptoms of the syndrome, their results will remain open to the criticism that no children have been shown to even nearly “recover” from autism. Even the best-outcome children of behavioral interventions, who do demonstrate major gains on measures of IQ and social development, may continue to exhibit equally important difficulties on specific social and cognitive skills (Mundy & Crowson, 1997).

Second, because research has linked joint attention skills to later developing ‘symbolic abilities’ (Hobson, 1993; Mundy, Sigman, & Kasari, 1993), ‘language abilities’ (Baldwin, 1995; Bates et al., 1979; Bruner, 1975; Tomasello, 1988), and ‘general social-cognitive processes in children’ (Baron-Cohen, 1995; Bruner, 1975; Mundy, 1995; Tomasello, 1995), the development of intervention technologies specifically aimed to produce joint attention holds the potential of a significant breakthrough in interventions for children with autism.
Third, some cognitive psychologists have insisted that joint attention is not amenable to a learning explanation and that behavior analysis is essentially irrelevant to this field of inquiry (e.g., Bruner, 1995; Tomasello, 1995). The demonstration of how basic behavior principles may be involved in the establishment of phenomena treated under the heading of ‘joint attention’ will be a substantial support to behavior-analytic view that the role of behavior principles in ‘psychological development’ has been vastly underestimated in the field of developmental psychology.

As noted by Schlinger (1993), “developmental psychologists have provided valuable information about child development. Unfortunately, such information lacks a strong unifying theoretical background and fails to impart practical knowledge that can enable psychologists to reliably change behavior in natural settings” (p. viii). In many recent psychology textbooks and articles, the discussion of learning, not to mention behavior analysis, is almost nonexistent, and even when sections on learning principles are included, authors rarely refer back to these principles in analyses of complex phenomena. It is quite astonishing to observe that not even the basic, well-documented principle of operant reinforcement is incorporated in developmental psychology. In order to more fully appreciate the potential of an operant analysis of joint attention phenomena, some knowledge of behavior principles and technical terms are required. These will be briefly outlined here.

Behavior principles and some technical terms

A widespread misconception suggests that behavior analysis confines itself to what can be directly observed and to responses that result from a direct conditioning history. For instance, according to Bruner (1973) there is a tradition in psychology that prefers to stop at the level of behavior, dispensing with notions like intention, “but it is a necessity for the biology of complex behavior, by whatever label we wish to call it” (p. 2). The field of “intention” is, of course, the very field of operant behavior. An operant analysis will never stop at the level of behavior, by whatever label we may wish to call it. On the contrary, an operant analysis will instantly move on to the variables of which the behavior is a function.

Operant reinforcement

Whenever, colloquially speaking, someone intends to obtain an effect, we could say that the behavior operates upon the environment (social or otherwise) to produce that effect. The operant reinforcement principle is likely to be, at least partly, familiar to researchers in developmental psychology: When behavior is followed by certain consequences, the frequency of such responses increase as a result. Basic behavioral research has amply demonstrated the robustness of the reinforcement principle in human as in other animal behavior. Even so, authors sometimes reject a behavior-analytic view as “less plausible” and assert that some of the coordinated action in ‘joint attention’ is just an unlikely candidate for the conditioning explanation. For instance, Tomasello (1995) contended that:

... while the conditioning explanation can never be ruled out completely, children’s spontaneous gaze alternations, and the way they are coordinated with their ongoing social interactions at around 12 months of age, makes less plausible the conditioning explanation and more plausible the view that the child understands that the adult is a separate person who has intentions and attention that may differ from its own. (p. 109)

However, that the child “understands that the adult is a separate person” only sums up the current structure of behavior-environment relations and obviously does not preclude the relevance of a conditioning history. ‘Understanding’ can be considered as a summary label for complex sets of performances, but it does not point to independent variables of which performance is a function.

Operant discrimination

In an operant analysis, ‘attention’ boils down to ‘stimulus control.’ Whenever, in colloquial terms, we
say that a child attends to something, behavior analysts will move on to specify the child’s behavior that is controlled by that something. Why, then, would anyone prefer such a technical vocabulary? There is good reason for the focus on relations between behavior and events in the environment. When that is accomplished, a large body of research literature on how to establish and change stimulus control becomes directly relevant to our work in this field. This literature includes work on simple stimulus control (e.g., Herrick, Myers, & Korotkin, 1959; Blough, 1958; Reynolds, 1961), compound stimulus control (cf. Dinsmoor, 1995; Donahoe & Palmer, 1994), simultaneous and successive discrimination (e.g., Loess & Duncan, 1952; Zentall & Clemet, 2001). More recently, there is a growing body of experimental research on more complex stimulus control, to be noted below.

**Conditioned reinforcement and behavior chains**

There is sometimes a preconception regarding what can reasonably function as a reinforcer. For instance, some have indicated that on the basis of the observations of “how social or sharing or reciprocal such attentional activity is . . . it was . . . inevitable that we grew uncomfortable with learning theory explanations of how eye-to-eye contact came into being, or how it shifted over to shared attention on common objects. With respect to the former, there were even studies indicating that eye-to-eye contact itself was reinforcing in learning tasks” (Bruner, 1995, p. 2). Similarly, according to Tomasello (1995), “in [the case of declaratives] the child simply shows or shares something with an adult, which would not seem amenable to a conditioning explanation as there are no apparent rewards involved.” Although he admits that “if human beings are rewarded by smiles and other signs of acknowledgement from adults, then they might be conditioned in their use of protodeclaratives as well,” he adds that “this stretches the conditioning explanation somewhat out of shape” (p. 111). Why this should be stretching the conditioning explanation somewhat out of shape is not explained, and I can think of no other reason for this suggestion than some sort of preconception of what can possibly function as reinforcers, for instance couched in terms of a drive-reduction theory (cf. Chomsky’s 1959 review of Skinner’s *Verbal Behavior*).

The consequences of behavior that can function as reinforcers can be either: (1) purely material things, (2) social stimuli, or (3) stimuli correlated with access to other (high-probability) activities (i.e., the Premack Principle). Further, some reinforcers function as such without requiring any type of prior “learning,” while others come to function as such only after they appear in certain types of relation to other reinforcers.

Although the details of the principles involved in the establishment of new, conditioned, reinforcers may still need to be explored in some detail (cf. Fantino & Logan, 1979) we do know a lot about how to establish new things or events as reinforcers. The standard procedure that is suggested in the literature of applied behavior analysis (e.g., Lovaas et al., 1981; Maurice, Green, & Luce, 1996) is a “pairing” of stimuli that one wants to establish as conditioned reinforcers with unconditioned or primary reinforcers. A safer, and possibly more effective, procedure is to establish the new, to-be-conditioned, reinforcer as an $S^D$ for a response that produces the unconditioned reinforcer (e.g., Dinsmoor, 1950; Keller & Schoenfeld, 1950; Lovaas, Freitag, Kinder, Rubenstein, Schaeffer, & Simmons, 1966; Skinner, 1938). Behavior chains will then build up, in which the reinforcing consequence of one behavioral element constitutes the occasion for other behavior which typically produces reinforcement.

Generally, the effectiveness of conditioned reinforcers will depend on the presence of the establishing operation (e.g., deprivation) that the primary reinforcement effect depends on. However, if the conditioned reinforcer obtains its effect through a similar relation to a number of different primary reinforcers, it will become a *generalized conditioned reinforcer*. The effectiveness of such reinforcers is less dependent upon each specific establishing operation upon which each of the unconditioned reinforcers may depend.

**Conditional discriminations**

The three-term contingency $S^D \rightarrow R \rightarrow S^R$ is, perhaps, the most robust behavior-analytic formula, but behavior analysis is not limited to it. The three-term contingency can be placed under conditional or contextual control: A response may be followed by a reinforcing event in the presence of a particular stimulus,
but this relation may hold only in the presence of some additional stimulus (e.g., Sidman, 1986). For instance, you may dial a telephone number in front of you to produce the voice of some interesting person, but only when the dialing tone is present first. Similarly, in many social settings, responding to certain features in the environment will be reinforced by other people, but only when they, too, attend to those features of the environment as well. Features of conditional discrimination training have been studied in great detail over the last 20 years and have been reported in the literature on stimulus equivalence (e.g., JEAB, 1996; Sidman, 1994).

**Joint control**

Sometimes, behavior depends on the simultaneous, or joint, control by two different stimuli over a single response (e.g., Lowenkron, 1998). For instance, if somebody requests a 14 mm socket, you may repeat “14mm” as you scan a number of sockets until you see one which controls the same response (saying “14 mm”), such as one with “14mm” printed on it, before you stretch out and pick up that socket. In social interactions, when you try to locate an object or event to which another person attends, it may be helpful if you are told a name or otherwise given a description of that object or event. You may then visually scan the environment until you see something that controls the same verbal response in you. In the absence of a verbal description, you may simply respond, at least in part, like you observe the other person to do, and scan the environment in that other person’s visual field until you see something that controls that same response (e.g., smiling or frowning) in you.

**Conjugate reinforcement**

Reinforcement is not just an on-or-off issue. In what has been termed ‘conjugate reinforcement,’ there is a contingent relation between the intensity (e.g., frequency) of the response and the intensity of some continuously available stimulus, and changes in the intensity of the continuously available stimulus functions as a reinforcer. (e.g., Lovitt, 1967; Rovee-Collier & Gekoski, 1979). Much of what may function as social reinforcers, such as other persons’ attention may often not be an on-or-off matter, but a matter of intensity typically related to an intensity of responding. For example, when guiding someone else’s attention, we may be sensitive to small changes in the direction of the person’s looking in the right or wrong direction.

**Continuous repertoires**

Sometimes reinforcement is contingent upon a correspondence between response dimensions and stimulus dimensions, as in what has been referred to as continuous fields (Skinner, 1953). Such continuous fields may lead to continuous repertoires in which intermediate values on the stimulus dimension control intermediate values on the response dimension, and extreme values on the stimulus dimension control corresponding extreme values on the response dimension (Wildemann & Holland, 1972). Crude gaze following, for example, may result from the direct training of only a limited number of different exemplars.

**Observing responses**

Organisms are, of course, not only passively exposed to stimuli. They operate on the environment as if “gathering information” relevant to the issue of how to respond next. However, “gathering information” may not accurately describe the function of such behavior. In an experiment on observing behavior by Dinsmoor (1983), pigeons were exposed to a multiple schedule in which pecks on one key were extinguished in the presence of a red light and reinforced according to a variable ratio schedule in the presence of a green light. Under such circumstances, key pecking that produce either green or red light (that is correlated with reinforcement and nonreinforcement, respectively) will be maintained. In terms of information value, the red light and the green light should be equal. However, if responses on the observation key only produces green light when the reinforcement schedule operates, pecking the observation key is maintained, whereas if only the red light is produced when the extinction schedule is operating, responding to the observation key is not maintained (Dinsmoor, 1983). To the extent that this finding can be extrapolated to human behavior, monitoring the behavior of other persons is best maintained when some properties of their behavior serve as
positive discriminative stimuli, i.e., occasions for doing something that produces a reinforcing event. Correspondingly, such monitoring may not be well maintained when distinct properties of the other person’s behavior mainly function as negative discriminative stimuli ($S_D^-$) in the presence of which behavior is not reinforced.

**An operant analysis of joint attention performances**

In colloquial terms, joint attention can be said to involve the detection of what another person attends to. A fairly general principle seems to be that people particularly attend to things, events, or properties that are novel. Such preference for novel stimuli is well documented even in infants and utilized in experiments on so-called “recognition memory” using habituation procedures (e.g., Bornstein, 1976), paired-comparison (novelty preference) procedures (e.g., Fantz, 1964), or novelty discrimination procedures (e.g., Werner & Siqueland, 1978). Further, when it comes to verbal skills and listening skills, people tend to report on deviations from standard patterns of events and to listen to such reports with more interest than in reports on routine events or things that do not change, except when invariability itself is novel. Obviously, novelty does not exist by itself but only as a property of the history of each person with respect to particular things and events. Although some types of events are likely to be novel to most people, a detailed knowledge of what is novel to, and likely to exert stimulus control over some perceptual behavior of, a particular person will require a more detailed knowledge of the history of that person.

**An operant analysis of gaze following.**

Mundy et al. (1996) distinguished between a lower and a higher level of responding to joint attention. Lower level behavior consists of orienting head and eyes in accord with another person’s proximal point or touch. Higher level behavior involves following someone’s line of regard (beyond the index finger if pointing is involved) to some object or event.

In behavior-analytic terms, the lower-level behaviors can occur as standard discriminated operants – the product of a standard three-term contingency: The adult’s pointing or touching is the occasion upon which looking in that direction is typically followed by reinforcing consequences. These reinforcing consequences may very well be purely visual.

Higher-level performances can involve very much more complex skills. As Bruner (1995) pointed out, the child’s action must not only be started by the adult’s gaze, but it must also stop “when the infant finds a visual target out there” (p. 7). A relatively simple version of a higher-level behavior could consist of a two-component behavior chain in which an adult’s gaze in a particular direction serves as a discriminative stimulus for the child’s turning the head/eyes to look in that general direction. In that vicinity, something irregular is happening which functions simultaneously as a conditioned reinforcer for turning the head/eyes and as an $S^0$ for visual focusing and further looking. However, if the child, eventually, focuses on something for reasons totally apart from what the adult was initially attending to, this may not fulfill the stricter criteria of true joint attention adhered to by some authors. In true joint attention, the child must focus (and stop scanning) not just dependent upon seeing something that singularly reinforces the child’s seeing, but which is likely also to have functioned as an $S^0$ for the adult’s look. In traditional terms, then, we want to know the basis on which the child determines what the adult is attending to. More specifically, the child may focus (and stop scanning) when looking is jointly controlled by the adult’s gaze and some novel or irregular thing or event.

**An operant analysis of ‘social referencing’**.

Simple forms of social referencing may be built up similarly to observing behavior as studied in the laboratory (e.g., Dinsmoor, 1983). Again, if true joint attention is involved in the sense that the child can be said to “understand” that the familiar person attends to the same event as the child does, the principle of joint control must be involved. Hence, social referencing with joint attention requires that the child behaves in accord with the behavior of the familiar person towards a novel stimulus, but only contingent upon an event
of joint control in which the child could be said to infer that the familiar person’s behavior is controlled by the same novel event as the child’s own behavior.

An operant analysis of the ‘protoimperative’.

A definition of protoimperative gestures as “gestures intended to make another person do something for one’s benefit” (Sarriá et al., 1996) may correspond closely to Skinner’s preliminary definition of a mand as “a verbal operant in which the response is reinforced [through the mediation of other persons] by a characteristic consequence and is therefore under the functional control of relevant conditions of deprivation or aversive stimulation” (Skinner, 1957, pp. 35-36).

Protoimperatives (or mands) can also occur without features of joint attention. A child may simply persist in doing what has previously produced reinforcers through the mediation of other persons without otherwise being sensitive to whether or not anyone attends at the moment. However, protoimperatives usually work more smoothly and reliably when the child engages in observing behavior that establishes another person’s attention to what the child is pointing at.

An operant analysis of the ‘protodeclarative’. Whereas protoimperatives correspond to Skinner’s (1957) definition of a mand, protodeclaratives may correspond to a rudimentary version of a tact, which is established by reinforcement “with many different reinforcers or with a generalized reinforcer” (Skinner, 1957, p. 83). Specifically, “the purely social motive of sharing attention to something” (Tomasello, 1995, p. 111), may imply that behavior is typically reinforced by social consequences, such as other persons’ nods, smiles, visual orienting, uttering “yes,” “oh,” “look at that,” or other relevant comments that in Skinner’s (1957) terminology constitute intraverbals. Thus, joint attention is central to the ‘protodeclarative’ in the sense that the joint attention of other persons constitutes the reinforcement that characterizes this function.

An operant analysis of monitoring

Instead of just responding to discrete instances of other persons’ looking or pointing, a child may “keep an eye on” someone in order to detect such instances. As lower-level joint attention, such continuing observing behavior, or vigilance, may be automatically and abundantly reinforced when the child is monitoring parents or others who may be particularly qualified at focusing on events that may reinforce the child’s perceptual behavior. On the other hand, true joint attention in such monitoring would seem to require a contingency of the type that characterizes the ‘protodeclarative’ or tact and involve similar social reinforcers.

In sum

The present operant interpretation of joint attention skills points out seven basic factors. (1) In social interactions that involve visual joint attention, the visual orienting of one person is under discriminative control of the pointing or visual orienting of another person. (2) Such discriminative control may be conditional upon other stimuli. For instance, such point or gaze following may be particularly likely in the presence of certain facial expressions, when someone says “Look!” or when you have asked for directions. (3) In a three-dimensional world, a great many different objects, events, or properties of objects and events may exist in the direction of someone’s look, so that identifying the particular stimuli at which someone else is focusing must be jointly controlled by the direction of the look and something else. (4) Both the extent to which someone follows another person’s orienting, and the extent to which one operates to get others to follow one’s own orienting depends on previous consequences of such behavior. (5) When one directs the attention of someone else, small changes in the right direction may function as reinforcers, and when following someone else’s direction, a novel stimulus may, typically, function as a reinforcer. (6) In both cases, the reinforcers may have gained in strength because they are typical precursors of the moment of joint attention which, in turn, constitutes an occasion upon which other behavior (e.g., verbal behavior) is likely to be reinforced. (7) A limited number of exemplars of successfully following and directing others’ attention may
suffice to produce a continuous repertoire of such joint attention skills.

Advantages of an operant analysis: Implications for the applied field

Generally, the focus on accessible variables in behavior analysis makes it directly applicable to practical issues. Having found that “joint visual attention is not spontaneously demonstrated by infants until about 10 months of age” and that, “given the appropriate feedback infants are able to acquire a gaze-following response from about 8 months on,” Corkum and Moore (1995, p. 78) concluded that “learning is a possible mode of acquisition for joint visual attention.” Such acceptance of “learning as a possible mode” may be a first step towards an analysis of the variables of which joint attention skills are a function. If joint attention skills are amenable to an operant analysis, learning protocols aimed at the establishment of such skills appear to be a rather straightforward matter. Here are some examples:

1. Social referencing: Establishing normal social stimuli as reinforcers

If social stimuli that function as reinforcers for behavior in most people, including children, do not do so for behavior in children with autism, a crucial step may be to establish such events as reinforcers. The following outline of a training procedure will focus on establishing others’ nodding and smiling as reinforcers.

Training: Trainer and child are seated face-to-face at opposite sides of a table. Spread approximately 10 small edible reinforcers around the table. Any attempt from the child to take pieces from the table should be blocked. When the child sits quietly, nod and smile before you let the child take one item. If the child does respond, repeat the nod and smile, and prompt the child to take one item from the table. Then, as long as you do not nod and smile, block any attempts the child may make to take things from the table, and when you nod and smile, let the child take another item, and so on. Let the time vary between each time you nod and smile. When the child takes items from the table only immediately following your nods and smiles, this constitutes a simple version of social referencing. Further, it is appropriate to say that your nods and smiles function as an $S_D$ for the child’s response in taking items from the table, which is also a reliable indication that your nods and smiles will function as a conditioned reinforcer for any behavior in the child that produce your nods and smiles as consequences. An early change in the child’s behavior will be an obvious increase in the child’s visual attention to your face. Your nods and smiles can then be utilized to establish useful social behavior in the child, such as calling your name and, later, directing your attention to other objects and events.

Clearly, the simple procedure described above is only a start, and a large number of problems remain and will have to be solved. First, your nods and smiles are likely to function as reinforcers only when those edible reinforcers are visible at the table. Second, although your nods and smiles now function as a conditioned reinforcer, they will not be generalized: They will only function as reinforcers as long as those edibles are reinforcing, i.e., as long as relevant deprivation is maintained. Third, nods and smiles by others than the trainer may still go unnoticed by the child. Fourth, in addition to nods and smiles, other persons’ uttering “yes,” “oh,” “look at that,” and other relevant comments (intraverbals) will also need to be established as conditioned reinforcers in order to establish a general interest in the normal social consequences of engaging in standard “communication.”

2. Establishing monitoring

Additional monitoring may be established effectively by having the child actively guiding someone else’s behavior through several steps that are necessary in order for the child to make that other person locate and deliver a reinforcer. An example of a sequence of relevant tasks is the following: First, attach envelopes, say 5-6, on a horizontal line on the wall. Let the child sit and watch from a distance of 3-4m that someone puts some snack (or other potential reinforcer) into one of the envelopes. Then, tell the child to instruct you on where to find the snack for him or her. Start pointing to some random envelope and have the child prompted to guide you by pointing further to the left, further to the right, or by saying “stop” as your pointing
finger moves before the envelope in which the reinforcer is located. In a second version of the task, have the envelopes arranged in a vertical row and have the child guide you by pointing further up, further down, or saying “stop.” Next, combine the tasks by having envelopes pasted over a large area on the wall. In more advanced version of the task, the child may be taught to specify in more detail how you should move from your current pointing position, such as “next one to the right and two up!” Let the child confirm that you have the right position by saying “stop” or “yes, that’s the one” before you pick up the reinforcer, and to make “mistakes” occasionally, so that the child cannot successfully relax his/her monitoring of your behavior.

3. Establishing gaze or point following

Pretaining: Trainer and child are seated face-to-face at opposite sides of a table. The trainer shows the potentially reinforcing stimulus to the child, asks the child to turn around (or otherwise makes sure that the child cannot observe), puts a potentially reinforcing stimulus under one of two opaque cups turned upside down on the table. Next, the trainer says “ready” and makes sure that the child observes the cups and chooses one of them by pointing to it. The trainer lifts up the cup and, if the reinforcer is located under the cup to which the child pointed, the child is allowed to grab it. If the reinforcer is located under the other cup, the child is just allowed to observe it before it is removed by the trainer and a new trial is started. The pretraining continues until the child turns around within a couple of seconds when asked to do so, and turns back and chooses one of the cups within 5s when the trainer says “ready.”

Training: Use the same arrangement as during pretraining, except that the trainer moves his/her face as close to the cup that contains the reinforcer as is necessary to make the child look at the trainer’s face before being allowed to choose one of the cups. Repeat this until the child observes the trainer’s face and consistently (e.g., four successive times) chooses the object with the reinforcer placed under it. Next, the trainer fades his face away from the cup on successive trials until the child observes the trainer’s face and chooses the “correct” cup even when the trainer sits laid back and just looks at the cup under which the reinforcer is placed.

General ideas for training extensions: Hide the reinforcer behind different objects in different places, use different reinforcers and different trainers.

4. Establishing mands with joint attention

Mands (protodeclaratives and declaratives) without joint attention are evident when they typically occur indiscriminate of a listener’s attention. For instance, even a well developed for, such as “Can I have that chocolate?” may occur and simply be repeated even in the absence of evidence that any listener attends.

Training: Let the child observe you putting potential reinforcers away so that they are not accessible to the child without your participation. When the child produces the first mand, do not deliver the terminal reinforcer, but prompt an attention-getting response in the child, such as calling your name, which will typically be followed by your appropriate listener behavior, such as answering “yes,” and visually orienting in the child’s direction. Then, let this constitute the occasion upon which child’s mand is typically reinforced.

5. Establishing tacts

Joint attention appears to be particularly important to verbal behavior under stimulus control, such as tacts. Skinner (1957) defined the tact as “a verbal operant in which a given response form is evoked (or at least strengthened) by a particular object or event or a property of an object or event.” According to Skinner’s technical analysis, the unique relation to a discriminative stimulus, rather than to a specific establishing operation, is obtained by (1) many different reinforcers or (2) generalized reinforcers. However, additional analyses are required in order to work out an effective intervention plan regarding how to strengthen a tact repertoire in persons who demonstrate a distinct lack in that domain. First, a normal tact repertoire is not likely to be practiced and maintained if normal listeners’ responses do not function as generalized (conditioned) reinforcers. Hence, procedures for establishing generalized conditioned
reinforcement need to be based on naturalistic observations of the specific events that are likely to constitute listener reactions to tacts in the speaker’s natural language. A large literature on ‘joint attention’ seems particularly relevant to this issue. Another person’s joint attention, in the form of visual orienting, nodding, smiling, and uttering different types of “relevant comments,” constitutes the reinforcement that characterizes the tact function. Hence, unless those responses that can be summarized as joint attention from another person actually function as reinforcers for a child’s behavior, there is no basis for a development of tacts.

Numerous attempts to establish conversational skills (such as tacts and intraverbals) in children with autism appear to have succeeded mainly within the limits of an artificial training setting in which verbal behavior has been reinforced by characteristic consequences that typically produce mands. A successful tact training program, then, must ensure that the consequences that typically follow and maintain tacts in the natural environment do, in fact, function as reinforcers. Hence, training along the lines described in previous sections (1) Social referencing: Establishing normal social stimuli as reinforcers, and (2) Establishing gaze and point following, may turn out as pivotal (e.g., Burke & Cerniglia, 1990; Koegel, Koegel, Harrower, & Carter), or prerequisites, for successful tact training. Once normal consequences do function properly as reinforcers, it is possible that exposure to naturalistic conditions may suffice to foster commenting and other conversational skills. However, we may want to speed up such a development through additional training. First, there may be a large number of “names” of objects and events that may initially be most expediently established through traditional discrete trial training (e.g., as described in Lovaas et al., 1981).

Second, in order to produce a high frequency of “learn units” (e.g., Greer & McDonough, 1999), it may be preferable to establish child-initiated training during many different naturally occurring circumstances by initially reinforcing tacting abundantly whenever it occurs.

Third, it may be wise to teach the kinds of verbal skills that are most likely to be reinforced by standard listeners. What does seem more likely to be reinforced in natural settings over time is commenting on things or events that are novel in some way. A deficiency in this area may be particularly evident in many children with autism. As one parent wrote to an internet discussion group on applied behavior analysis for children with autism, “Does anyone have any ideas on how to develop a program on teaching a child to comment? My son . . . does not make comments. A purple cow could walk by and he wouldn’t mention it.” A series of tasks that may teach the necessary skills in discriminating novel stimuli may start with simple “What’s missing?” tasks (e.g., Lovaas et al., 1981) and similar training focused on “What’s added?”, “What’s changed?” and “What’s strange?” In order to increase the rate of spontaneous commenting in natural settings, instructions may be faded by increasing the time and the distance from instructions to opportunities to respond. Novel stimulus constellations can be arranged in other rooms and gradually in more distant places so that the child is given opportunities to respond in the absence of immediate instructions.

Research questions derived from an operant analysis

According to an operant analysis, it is entirely possible that normal development of behavioral repertoires in children exposed to normal environments relies fundamentally on normal social reinforcing stimuli to function as such from very early on. If even the parents’ visual attention and smiling does not function as reinforcers for the behavior of an infant, important early forms of social skills related to joint attention may not develop. A number of important research questions follow from this interpretation:

1. Do other’s visual attention, nodding and smiling normally function as a reinforcer from birth, or does the reinforcing effect of such stimuli develop later, possibly mainly as a result of operant conditioning procedures? This could be investigated by using a conjugate schedule (see Rovee-Collier & Gekoski, 1979) in which the degree of visual orientation towards the child, nodding, and/or smiling of a human face on a monitor is changed contingent on the rate of sucking a non-nutritive nipple.

2. Is there a difference even at birth in the extent to which visual attention, nods, and smiles function as reinforcers for the behavior of typically developing children as compared with children with autism? The conjugate schedule procedure just mentioned could, in principle, be used to investigate whether such a
difference between children with autism and normally developing children exists even shortly after birth. Diagnosis and a possible initiation of corrective measures within the first few weeks after birth would seem like an interesting option.

(3) If the reinforcing effect of other persons’ visual attention, smiling and nodding typically depends on other, primary reinforcers, what are those primary reinforcers, and what are the relevant procedures to which children are typically exposed in their natural environments?

To the extent that other’s visual attention, nodding and smiling normally function as a reinforcer already from birth on, social interaction may function as what has been termed an autocatalytic process (e.g., Skinner, 1953) in which the reinforcing effect of such social events gain in strength because such events, in addition to being reinforcing in the first place, also constitute occasions upon which additional behavior is likely to produce more of the same. It is possible that the reinforcing effect may typically increase over time because such social events are typically correlated with a higher rate of positive reinforcement, such as other’s compliance with requests (mands). Further, it seems likely that other’s visual attention combined with nodding and smiling is typically correlated with a low frequency of aversive social stimuli. Research in this field could include initial naturalistic observation of parent—child interactions and proceed with systematic exaggeration of these features of children’s environments.

(4) Can conditioned reinforcers established through contrived contingencies be maintained as reinforcers at near-normal rates of back-up (primary) reinforcement? When other persons’ nodding and smiling do not have a reinforcing effect, such an effect can be produced by differentially reinforcing some behavior in the presence of such nodding and smiling. However, such an arrangement may work only when visible, or in the presence of “therapists” with a history of using it. Moreover, such an arrangement with obviously contrived reinforcers is potentially stigmatizing, particularly as the child advances to otherwise more normalized social environments. Hence, explicit conditioning of normal social reinforcers can only lead to a lasting normalized social skills repertoire if the conditioned reinforcing effect is maintainable at near-normal rates of primary reinforcement. How far can contrived contingencies of primary reinforcement be faded towards a non-conspicuous level without losing their effect?

**Conclusion**

The literature on joint attention has identified elements of social interaction that appear to be crucial for normal social functioning in general, and for verbal behavior in particular. Joint attention deficits seem to characterize children with autism, and a thorough operant analysis seems required in order to identify variables of which joint attention skills are a function. Hopefully, the current operant interpretation will spark off experimental analyses from which more advanced and effective intervention plans can be developed.

**References**


Behavior and Development, 1, 79-98.


Author’s Note

The current article is based upon a paper entitled “TACTs and the role of 'Joint Attention': An Operant Analysis,” delivered as an invited address at the 6th International Congress on Behaviorism and the Sciences of Behavior, August 2002.

Please address correspondence to:

Per Holth
The Behavioral Center
P.O. Box 1565
Vika, 0118
Oslo, Norway
e-mail: per.holth@atferd.unirand.no

Advertising in the Journal of Early and Intensive Behavior Intervention

Advertising is available in JEIBI. All advertising must be paid for in advance. Make your check payable to Joseph Cautilli. The ad copy should be in our hands at least 3 weeks prior to publication. Copy should be in MS Word or Word Perfect, RTF format and advertiser should include graphics or logos with ad copy.

The prices for advertising in one issue are as follows:

1/4 Page: $50.00 1/2 Page: $100.00 vertical or horizontal Full Page: $200.00

If you wish to run the same ad in both issues for the year, you are eligible for the following discount:

1/4 Pg.: $40 - per issue
1/2 Pg.: $75 - per issue -vertical or horizontal
Full Page: $150.00-per issue.

For more information, or place an ad, contact Halina Dziewolska by phone at (215) 462-6737 or e-mail at: halinadz@hotmail.com