The current review provides an original examination of the literature involving mood as a dependent variable in behavioral interventions for individuals with autism spectrum disorder (ASD). Twenty-nine single-subject experimental design studies were identified through systematic searches and evaluated in relation to the following variables: (a) participant and setting characteristics; (b) experimental design; (c) approach to mood assessment; (d) intervention procedures, and, (e) intervention effects on mood. The experimental rigor of the included studies was evaluated on the National Autism Center's Scientific Merit Rating Scale (2009). Results suggest that behavioral interventions can be successfully used to improve the mood of individuals with ASD, and that observable indices of mood should be more often measured in behavioral intervention research, particularly when improved QoL is a stated goal of intervention.
Mood as a Dependent Variable in Behavioral Interventions for Individuals with ASD: A Systematic Review

Devon Ramey
Olive Healy
Russell Lang
Laura Gormley
Nathan Pullen

1School of Psychology, Trinity College Dublin
2Texas State University; Clinic for Autism Research Evaluation and Support (CARES)
3Autism Spectrum Instructional Resources (ASpIRe)

Corresponding Author:
Olive Healy, Ph.D., BCBA-D
School of Psychology, Áras an Phiarsaigh
Trinity College Dublin, University of Dublin
College Green, Dublin 2, Ireland
olive.healy@tcd.ie
Abstract

The current review provides an original examination of the literature involving mood as a dependent variable in behavioral interventions for individuals with autism spectrum disorder (ASD). Twenty-nine single-subject experimental design studies were identified through systematic searches and evaluated in relation to the following variables: (a) participant and setting characteristics; (b) experimental design; (c) approach to mood assessment; (d) intervention procedures, and, (e) intervention effects on mood. The experimental rigor of the included studies was evaluated on the National Autism Center’s Scientific Merit Rating Scale (2009). Results suggest that behavioral interventions can be successfully used to improve the mood of individuals with ASD, and that observable indices of mood should be more often measured in behavioral intervention research, particularly when improved QoL is a stated goal of intervention.

Keywords: autism, happiness, intervention, mood, quality of life
MEASURING MOOD IN BEHAVIORAL INTERVENTIONS

Measuring Mood as a Dependent Variable in Behavioral Interventions for Individuals with ASD: A Systematic Review

Applied behavior analysis (ABA) is a science in which the principles of behavior are systematically applied to improve socially significant behaviors of individuals (Cooper, Heron, & Heward, 2007, p. 20). In 1968, Baer, Wolf, and Risley wrote their seminal article outlining the seven dimensions of ABA, which are: applied, behavioral, analytic, technological, conceptually systematic, effective, and generality. By applied, they argued that ABA interventions are focused on changing meaningful behaviors that will improve the overall QoL and well-being of individuals (Baer et al., 1968). Montrose M. Wolf (1978) later discussed the purpose of the Journal of Applied Behavior Analysis, in which he defended the idea of changing behaviors of social importance. Although the subjective concept of “social importance” presented challenges, Wolf called for the assessment of social acceptability when evaluating interventions and their outcomes (i.e., social validity measures). He believed these measures could provide information regarding the appropriateness of procedures in relation to how they improve the life quality of individuals (Wolf, 1978, p. 207). Similarly, Van Houten and colleagues’ (1988) paper identified six client rights that were intended to “direct both the ethical and appropriate application of behavioral treatment” (p. 381), including the right to services whose overriding goal is personal welfare. Considering this, it could be argued that improving the QoL of recipients of behavioral services has been a consistent and central theme of ABA since the early days of the field.

Over time, there have been notable improvements and advances in the quality of services provided to individuals with disabilities. Although QoL was an important concept in the founding of ABA, it was not until more recently that service providers began to use empirical measurements of QoL, and life quality has become a more direct focus of some intervention programs (McDougall, Evans, & Baldwin, 2010). In fact, it was not until the late 1980s and 1990s that QoL began to develop as a formal construct within the field of intellectual and developmental disabilities (Brown, 2017, p. 7). Researchers have proposed different theories over what variables and characteristics best constitute the concept of QoL, and there have been numerous definitions developed over the years (Felce & Perry, 1995). One definition of QoL widely accepted in the field of disability services was developed by Schalock and Verdugo (Arias et al., 2018, p. 123). Within this definition, QoL is described as having eight different domains – emotional wellbeing, physical wellbeing, material wellbeing, social inclusion, interpersonal relationships, self-
determination, rights, and personal development (Schalock & Verdugo, 2002). Regardless of the definition used, there is a consensus among practitioners that individual happiness is one of the key components of QoL (Carr, 2007; Parsons, Reid, Bentley, Inman, & Lattimore, 2012). The notion that happiness is related to QoL has clear surface validity. Simply, if an individual is happy, their QoL is likely better than if they were unhappy.

Individuals without disabilities can typically identify and describe their personal wellbeing, and their QoL is often measured through self-reports (Felce & Perry, 1995). However, the communication and social deficits associated with autism spectrum disorder (ASD) may make expressing private events, and emotional states in particular, difficult for these individuals (American Psychiatric Association [APA], 2013; Parsons et al., 2012). Further, the sole reliance on self-report measures is problematic with any population sample. Although individuals with ASD may be unable to verbally express their feelings, they are thought to still engage in similar indicators of mood (e.g., smiling, laughing) as individuals without disabilities (Reid & Green, 2006).

In behavioral terms, a person’s happiness or mood is considered a private event which cannot be observed directly (Green & Reid, 1996). Although mood is not directly observable, it could be argued that it is still a type of behavior, albeit a covert behavior (Johnston, 2013; Skinner, 1957). Although an individual’s mood cannot be directly observed, the overt (i.e., public) behaviors theorized to be associated with the individual’s feelings can be. Behavioral indicators of mood are sometimes called indices of happiness (or unhappiness), and they can be operationally defined, measured, and evaluated for likely correspondence with mood (Green & Reid, 1996). Like other behaviors, indices of happiness and unhappiness are influenced by environmental events and it has been suggested that practitioners should take this into consideration when applying behavioral interventions (Parsons et al., 2012).

For example, observable indices of mood could be defined and measured as dependent variables, and behavioral interventions could be developed to increase these indices (and overall QoL) of individuals with ASD (Dillon & Carr, 2007; Green & Reid, 1996; Iwata, 1991). If QoL is a focus of rehabilitation, then it would seem necessary to target and measure indices of happiness and unhappiness as an outcome to an agency’s services (Parsons et al., 2012). By measuring indicators of mood among service users, a provider could additionally assess the efficacy of interventions intended to increase the wellbeing of an individual with ASD; any interventions which result in an increase in indices of unhappiness could potentially be modified or eliminated (Lattimore, Parsons, &
MEASURING MOOD IN BEHAVIORAL INTERVENTIONS

Reid, 2009; Parsons et al., 2012; Stasolla et al., 2014b). Furthermore, the social acceptability of two different interventions could be compared by documenting the relevant changes in mood of the individual receiving such interventions; an intervention that increases the happiness of the individual could be argued to be more socially acceptable (Toole, Bowman, Thomason, Hagopian, & Rush, 2003). Socially valid interventions are more likely to be implemented with fidelity over time, which may contribute to overall intervention success (Kennedy, 2002).

Despite arguments for including a mood measurement to evaluate interventions provided to individuals with disabilities, mood is not a common dependent variable found within behavior analytic literature (Green & Reid, 1996). Behavior analysts rely on measurable and observable (i.e., overt) behaviors, which may explain the paucity of studies that have incorporated a measure of covert behavior such as mood. Of the published studies that have focused on indicators of mood in individuals with disabilities, a majority have targeted individuals with severe/profound intellectual disabilities (ID) or profound multiple disabilities (PMD) (see Dillon & Carr, 2007; Lancioni, Singh, O’Reilly, Oliva, & Basila, 2005, for reviews of this literature). Therefore, the current review provided an examination of the extant literature involving mood as a dependent variable in behavioral interventions for individuals with ASD. To the authors’ knowledge, this is the first review to date that has evaluated mood as a dependent variable in interventions with this population.

Method

The methods for this systematic literature review followed the guidelines outlined by the PRISMA statement (Moher, Liberati, Tetzlaff, & Altman, 2009).

Search Procedures

A four-step systematic search was conducted to identify the studies to be included in the current literature review. First, an online search was conducted across three electronic databases – ERIC, PsycINFO, and PubMed – to identify potential studies for inclusion. Within each database, the following terms were entered in the keywords field: “happiness”, “unhappiness”, “mood”, “emotion”, and “affect”. Each of these terms were combined with the term “autis*” or “Asperger” (i.e., happiness AND (autis* OR Asperger)). This online search was limited to English-written, peer-reviewed studies, but there was no restriction on year of publication. Next, the title of each study identified through the database search was screened; any study that did not meet inclusion criteria was eliminated.
from the third step of the screening process. For example, if the title outlined a pharmacological treatment, the study was automatically eliminated because this did not meet inclusion criteria. These first two steps of the search process resulted in 607 studies. Following the database search, these initial studies were further screened to determine which ones met the inclusion criteria outlined below. In a final step of the search process, the reference lists of all studies meeting inclusion criteria were reviewed to identify any additional studies for possible inclusion. This four-step search procedure occurred from March 2018 to April 2018, and a total of 672 studies were screened for inclusion (see Figure 1).

Figure 1. Flow diagram showing the four-step search procedure.

Inclusion and exclusion criteria. To be included in this review, each study met the following criteria. First, the study was required to include at least one participant with a primary diagnosis of ASD. Second, the study must have incorporated a direct measurement of mood as a dependent variable to evaluate the effects of an intervention. The mood measure must have been operationally defined in the study and recorded using a type of observational system, such as event recording or interval recording. Studies using the General Child Affect Rating Scale (GCARS; Dunlap & Koegel, 1980), interest/enthusiasm scale (Koegel & Egel, 1979), or Mood Scale and Related Interview Questions (MSRIQ; Carr, McLaughlin, Giacobbe-Grieco, & Smith, 2003) were included because these Likert-type scales offer scores that have been anchored to direct observation of overt behaviors.

Studies were excluded from the literature review for the following reasons: (a) the participant(s) had a comorbid diagnosis of a mood disorder; (b) the study did not employ a single-subject experimental design; (c) the study measured mood using only subjective questionnaires, inventories, and/or temperament measures; (d) the study focused on measuring and/or teaching empathy, emotion recognition, or emotion regulation; (e) the study focused on teaching affective behavior (e.g., how to smile, how to express dislike), or lastly, (f) the study investigated the effects of a pharmacological treatment or used brain imaging to study changes in mood. Furthermore, dissertations
and published literature reviews were also excluded from this review. This process of identifying studies for review is depicted in Figure 2.

*Figure 2. Flow diagram showing inclusion/exclusion of studies identified during the database search process.*

**Data Extraction**

All included studies were reviewed and summarized in terms of the following variables: (a) participant and setting characteristics; (b) experimental design; (c) approach to mood assessment; (d) intervention procedures, and, (e) intervention effects on mood. Within the summary of intervention effects on mood, the treatment efficacy of each study was determined by calculating the percentage of non-overlapping data (PND) score. The methodological rigor of each study was assessed using the Scientific Merit Rating Scale (SMRS) of the National Autism Center [NAC] (2009).

**Percentage of non-overlapping data (PND).** Treatment efficacy was determined for each study that presented a visual display of data for the mood variable by calculating a PND score. PND is a non-parametric statistic that is calculated by determining the percentage of intervention data points that do not overlap with baseline data points (Scruggs, Mastropieri, & Castro, 1987). For studies that implemented an alternating treatments design, the PND was calculated by comparing the two conditions on a point-by-point basis. For example, the first data point of one condition was compared with the first data point of the other condition, the second with the second, third with third, and so on (Wolery, Gast, & Hammond, 2010, p. 352). An intervention with a PND score greater than 90% is considered “highly effective”, while a score of 70–90% is considered “moderately effective” (Scruggs et al., 1987). A score of 50–70% is considered “mildly effective” and a score below 50% is considered “ineffective” (Scruggs et al., 1987). If a study did not provide a visual display for the mood variable, or any baseline data points, a PND score could not be calculated.

**Scientific Merit Rating Scale (SMRS).** The SMRS was developed by the NAC to objectively measure the strength of methods used within studies designed for participants with ASD (NAC, 2009). The SMRS was used to
evaluate the experimental rigor of each study in the current review based on the following characteristics: (a) research design; (b) measurement of the dependent variable; (c) measurement of the independent variable (i.e., procedural integrity); (d) participant ascertainment, and (e) generalization and maintenance of treatment effects. For each study, each of these variables was given a score ranging from zero to five (with 0 representing a “poor” score and 5 representing a “strong” score). The resulting scores for each variable were then combined to produce a composite SMRS score for each study (NAC, 2009). To calculate the composite SMRS score, the following formula was used: research design (0.30) + dependent variable (0.25) + participant ascertainment (0.20) + procedural integrity (0.15) + generalization (0.10) (NAC, 2009). A composite SMRS score of 3, 4, or 5 indicated that sufficient scientific rigor had been applied in the study; a composite SMRS score of 2 provided initial evidence that the study had been established, but more rigorous research was necessary. A composite SMRS score of 1 or 0 indicated that insufficient scientific rigor had been applied in the study (NAC, 2009, p. 23).

Reliability of Screening and Data Extraction Procedures

A second rater independently screened the abstracts of 220 (32.7%) of the 672 initial studies identified through searches to determine whether these studies met the inclusion criteria. The inter-rater reliability of the screening process was calculated by using the following formula to obtain a percentage: (number of agreements) ÷ (number of agreements + number of disagreements) x 100. Initial agreement as to whether a study should be included in this review was attained for 207 (94.1%) of the 220 identified studies. In the cases of non-agreement, consensus was reached between raters through discussion until 100% agreement was achieved.

The accuracy of the data extraction procedure was assessed on a seven-item checklist which contained the following: (a) Is this an accurate description of the participant(s) and setting?; (b) Is this an accurate description of the experimental design?; (c) Is this an accurate description of the mood measurement?; (d) Is this an accurate description of the intervention?; (e) Is this an accurate description of intervention outcomes?, and, (f) Is this an accurate description of the SMRS score? A second rater independently completed the checklist for each of the 29 included studies and the inter-rater reliability of the coding process was calculated. Items identically coded by both raters were marked as an agreement. The inter-rater reliability percentage was then calculated using the same formula above: (number of agreements) ÷ (number of agreements + number of disagreements) x 100. Initial
agreement for the coding process was 92% (range = 66.7-100%). Consensus was reached between raters through discussion until 100% agreement was achieved across all studies.

Results

Of the 672 initial studies screened for inclusion, 29 studies met the inclusion criteria and were subsequently reviewed for data extraction. The variables coded from each study are summarized in Table 1.

---

Insert Table 1 about here

---

Participant(s) and Setting

**Participant characteristics.** Table 1 presents the participant and setting characteristics for each study. Overall, there were a total of 99 participants who took part across the 29 included studies; 92 of these participants (92.9%) had a primary diagnosis of ASD. A majority of the studies (n = 24; 82.8%) involved children or adolescents (0–17 years), while five of the studies involved adults (18+ years; Carr et al., 2003; Kennedy, 1994; Koegel, Navab, Ashbaugh, & Koegel, 2016; Lattimore et al., 2009; Parsons et al., 2012). Of the participants, 75.8% were male (n = 75) and 24.2% were female (n = 24). Beyond the primary diagnosis of ASD, 22 (22.2%) of the total participants had a secondary diagnosis of mild to profound ID. Other diagnoses included the following: symptoms of attention-deficit/hyperactivity disorder (ADHD; n = 2); tuberous sclerosis (n = 2); developmental delays (n = 1); mixed developmental disorder (n = 1); cerebral palsy (n = 1); quadriplegia and hydrocephalus (n = 1); verbal and motor apraxia (n = 1); severe hearing loss (n = 1), and seizure disorder (n = 1).

**Setting characteristics.** Seven of the studies (24.1%) took place in a school classroom or school playground. Eight studies (27.6%) were carried out in a clinical or residential setting and five studies were conducted in the participants’ home (Arbogast & Fryling, 2015; Jull & Mirenda, 2011; Koegel, Stiebel, & Koegel, 1998; Koegel, Vernon, & Koegel, 2009; Stasolla et al., 2014b). Two studies were completed in an afterschool program or day camp setting (Gengoux, 2015; Jenkins & Reed, 2013). Six studies took place across multiple settings, including a supported work environment, group home, and/or residential facility (Parsons et al., 2012); residential facility and supported work environment (Lattimore et al., 2009); home and community (Vernon, Koegel,
Dauterman, & Stolen, 2012); home or school (Blakeley-Smith, Carr, Cale, & Owen-DeSchryver, 2009); home or community (Koegel, Werner, Vismara, & Koegel, 2005), and clinic or home (Vismara & Lyons, 2007). The setting was not specified in one study (Koegel & Egel, 1979).

Experimental Design

The experimental design for each study is described in Table 1. All studies utilized single-subject experimental designs (SSED). Eighteen of the studies (62.1%) implemented a multiple-baseline design or multiple-probe design; of these, 15 used a multiple-baseline across participants design (51.7%) and one used a multiple-baseline across behaviors design (Dunlap & Koegel, 1980). Two studies used a multiple-probe across participants design (Kennedy, 1994; Lattimore et al., 2009). Of the remaining 11 studies, seven implemented a reversal design (Jull & Mirenda, 2011; Koegel, O’Dell, & Dunlap, 1988; Koegel et al., 2009; Moes, 1998; Sigafoos et al., 2006; Stasolla et al., 2014b; Vismara & Lyons, 2007). However, Stasolla et al. (2014b) implemented a variation of the reversal design by using an ABB\textsuperscript{1}AB\textsuperscript{1} design, where B\textsuperscript{1} represented the same intervention with different reinforcement contingencies. The remaining four studies implemented an alternating treatments design (Arbogast & Fryling, 2015; Dunlap, 1984; Geiger et al., 2012; Parsons et al., 2012).

To further examine intervention effects, six studies embedded their experimental design with another design type. Three studies chose to embed a multiple-baseline across participants design with a reversal design for at least one participant (Koegel & Egel, 1979; Koegel et al., 2005; Lang et al., 2014). Dunlap and Koegel (1980) embedded a multiple-baseline across behaviors design with a reversal design for one participant, while Koegel et al. (2009) and Vismara and Lyons (2007) both embedded a reversal design with an alternating treatments design for at least one participant.

Mood Assessment

All the included studies used either in-vivo (i.e., observed directly) or in-vitro (i.e., observed through videotape) observation to measure the mood of participants during intervention. The majority of studies (n = 19; 65.5%) used in-vitro observation by coding video probes for the presence or absence of the mood indicator. Information regarding the mood measurement for each study is provided in Table 1.
Five (17.2%) of the identified studies in this review targeted mood as the primary dependent variable (Arbogast & Fryling, 2015; Carr et al., 2003; Jenkins & Reed, 2013; Koegel et al., 1988; Parsons et al., 2012). The other 24 studies (82.8%) employed other primary dependent variables, such as academic, play, social, and/or adaptive skills, that were the main targets of intervention. In these studies, the mood indicator was included as a secondary or ancillary measure.

**Operational definitions.** Operational definitions for the mood variable differed, with 18 studies (62.1%) opting to measure some type of affect and eight studies (27.6%) employing a measure of indices of happiness and/or unhappiness. Three studies defined their dependent variable specifically as “mood” (Carr et al., 2003, p. 35; Lang et al., 2014, p. 863; Sigafoos et al., 2006, p. 198), while Koegel and Egel (1979) defined “enthusiasm” as a measure of mood (p. 419). Lang and colleagues (2014) selected both “child affect” and “child mood” as measurements of the mood variable (p. 863-864), while Koegel et al. (2016) labeled the dependent variable mood as “affect and conversational interest” (p. 136). All the studies included within this review employed some type of behavioral indicator of mood in their operational definitions. For example, Geiger and colleagues (2012) defined “positive affect” as “smiling, laughing, clapping, and making positive statements about the activity” (p. 51). Table 1 provides a summary of the type of mood measurement used across all included studies and the operational definitions utilized by the authors.

**Measurement procedures.** The measurement procedures employed by each study are described in Table 1. Koegel and Egel (1979) originated a six-point Likert-type scale to rate the enthusiasm of their participants during observations. From this original enthusiasm scale, the GCARS was developed by Dunlap and Koegel (1980), and the researchers in 14 other studies (48.3%) employed the GCARS to score the mood of their participants. Similar to the GCARS, the MSRIQ was first implemented by Carr et al. (2003), and this scale was subsequently used in two studies included in this review (Blakeley-Smith et al., 2009; Lang et al., 2014). These observation scales rated the level of mood demonstrated by the participants with labels such as “Unhappy/Bad Mood”, “Neutral”, or “Happy/Good Mood”. The mood level scored for each participant was based on the observation of specific behavioral indicators. For example, in Koegel et al. (2005), child affect was coded by the following labels: “Negative affect” (Score of 0-1) defined as, “appears discontent (i.e., frowns, cries), avoids participation (i.e., tantrums, leaves activity), not enjoying self (i.e., frustrated, tense, impatient)”; “Neutral affect” (Score of 2-3)
defined as, “may engage in activity, but not interested or enthusiastic, not stressed or relaxed”, or “Positive affect” (Score of 4-5) defined as, “enjoying self (i.e., smiles, laughs), interested (i.e., actively involved), appears relaxed and comfortable”) (p. 97). Sigafoos et al. (2006) developed a mood scale for their study, which was similar to the MSRIQ, but with only three mood scores. With this scale, the following mood levels were marked at the end of each interval: “Bad mood” (Score of 1), “Neutral mood” (Score of 2), or “Good mood” (Score of 3; p. 198). Koegel and colleagues (2016) developed a three-point rating scale for “affect and conversational interest” which was based on the GCARS (p. 136). With this scale, the following mood labels were scored: “Low affect/interest” (Score of 0-1), “Medium affect/interest” (Score of 2-3), or “High affect/interest” (Score of 4-5; p. 137).

Eight studies (27.6%) used partial-interval recording or momentary time sampling (MTS) to code the presence or absence of operationally defined indices of mood (Arbogast & Fryling, 2015; Geiger et al., 2012; Jenkins & Reed, 2013; Lattimore et al., 2009; Parsons et al., 2012; Spector & Charlop, 2017; Stasolla et al., 2014a; Stasolla et al., 2014b). The final two studies implemented a frequency measure to count the number of discrete occurrences of each mood indicator (Kennedy, 1994; Vernon et al., 2012). Vernon and colleagues (2012) also opted to include a duration measure to calculate the number of seconds the participants engaged in each of these behavioral indicators.

**Interoobserver agreement (IOA).** IOA for the mood variable was measured in all 29 studies, but these data were not reported in Koegel et al. (2012). The researchers did, however, report Cohen’s Kappa coefficient for the affect measure which was $\kappa = .89$, suggesting almost perfect agreement (Koegel et al., 2012, p. 368). For the remaining 28 studies, the reliability of the mood measurement was reported as acceptable (i.e., > 80%; Cooper et al., 2007) in 27 of the studies. Koegel et al. (2016) additionally used Cohen’s weighted kappa to calculate the reliability of the affect/interest measure, reporting the coefficient as $\kappa = .32$, suggesting that there was only fair agreement.

**Intervention Procedures**

Intervention procedures are outlined in Table 1. Each study included in the current review was categorized as utilizing either a single behavioral intervention or a behavioral intervention package, or the study provided a comparison of the effects of two different interventions or conditions. The target behaviors and other dependent variables for each category of intervention are subsequently described.
Behavioral interventions. A total of 20 studies (69%) focused on increasing the academic, social, or adaptive skills of participants through specific behavioral interventions or intervention packages. Seven of these studies used antecedent-based interventions as the independent variable (Blakeley-Smith et al., 2009; Carr et al., 2003; Dunlap & Koegel, 1980; Gengoux, 2015; Kennedy, 1994; Koegel & Egel, 1979; Moes, 1998). Specifically, Blakeley-Smith and colleagues (2009) used task analyses and modified task demands to meet the competency of their participants (i.e., good environmental fit). Kennedy (1994) used an antecedent-based intervention in which the instructor provided high rates of social comments while fading task demands to target the work productivity of participants. Similarly, Dunlap and Koegel (1980) randomly rotated the delivery of tasks to increase the independent responding of participants. Carr et al. (2003) used a mood induction procedure to increase the motivation of participants. In their study, stimuli associated with a “good mood” were delivered to participants prior to task demands if at that time the participants were demonstrating indices of a “bad mood”. This mood induction procedure continued until the participants demonstrated a post-intervention mood rating of 3-5. Moes (1998) also targeted the motivation of participants by providing within- and across-activity choices. Lastly, Gengoux (2015) used priming procedures to prepare participants for future social activities, while prompting procedures were used by Koegel and Egel (1979) to increase the on-task behavior of participants in their study.

Three additional studies used a special type of antecedent-based intervention as their independent variable. Specifically, Baker (2000), Baker et al. (1998), and Koegel et al. (2005) embedded the environment with the preferred items or perseverative interests of participants. Perseverative interests were defined as preoccupations or obsessions that an individual continually seeks out (Carnett et al., 2014, p. 369). Both Baker (2000) and Baker et al. (1998) embedded the perseverative interests of participants into either board games (e.g., Bingo) or playground games (e.g., tag, follow-the-leader), respectively. On the other hand, Koegel et al. (2005) employed cooperative opportunities using preferred activities that both the participants and their neurotypical peers enjoyed. A form of embedded pivotal response training (PRT) was evaluated by Vernon et al. (2012), in which a motivating social interaction was embedded into the delivery of non-social preferred stimuli.

The final nine studies implementing a specific behavioral intervention or intervention package used a wide variety of procedures. Jenkins and Reed (2013) measured the effects of animal-assisted therapy in the form of horseback riding on several different dependent variables. Stasolla et al. (2014a) measured on-task behavior by
teaching participants how to use a self-monitoring system. The researchers reinforced the implementation of this self-monitoring system with a token economy. Koegel et al. (1998) used parent training to teach guardians how to implement functional communication training (FCT) procedures and design environmental arrangements. The results of the guardians’ implementation of these procedures were measured across participants. Similarly, Spector and Charlop (2017) trained the siblings of children with ASD to implement a Natural Language Paradigm (NLP) intervention. Social interactions were the focus of both Koegel et al. (2012) and Jull and Mirenda (2011). More specifically, Koegel and colleagues (2012) evaluated the use of “initiations training” to increase the social interactions of participants, and Jull and Mirenda (2011) assessed the effectiveness of parent-implemented contextually supported play dates to support child social interactions.

Behavioral treatment packages were used by three studies (Koegel et al., 2016; Lang et al., 2014; Lattimore et al., 2009). Koegel et al. (2016) focused on teaching the reframing of negative statements by first providing a brief training package, which included an explanation of reframing, modeling, and role-play. The researchers subsequently used video-feedback to practice reframing and the participants were taught a self-management procedure to monitor their use of this skill. To teach appropriate play skills, Lang and colleagues (2014) used the following strategies: prompting, positive reinforcement, response interruption and redirection (RIRD), and lag schedules of reinforcement. Finally, Lattimore et al. (2009) used intensive teaching in the form of chaining, prompting, programming common stimuli, and positive reinforcement to teach their participants the steps of a task analysis. A description of each behavioral intervention can be found in Table 1.

Comparing two interventions/conditions. Different types of teaching procedures or conditions were compared in nine (31%) of the included studies. Traditional discrete trial training (DTT) was compared with embedded instruction in two studies (Geiger et al., 2012; Sigafoos et al., 2006). Geiger and colleagues (2012) inspected these two teaching procedures further by including a concurrent-chains preference evaluation in which the participants had an opportunity to select which type of instruction they preferred (DTT or embedded). Constant task instruction was compared with varied task instruction in one study (Dunlap, 1984). In this study, there were two different types of varied task instruction evaluated: varied with acquisition targets and varied with both acquisition and mastered targets. Koegel et al. (1988) compared the following two conditions: reinforcement of successive improvements in motor speech (i.e., shaping), and reinforcement of any verbal attempt. Similarly, Stasolla et al.
compared the following two reinforcement conditions: reinforcing adaptive responding skills with or without challenging behavior and reinforcing adaptive responding skills without challenging behavior. These researchers focused on adaptive responding in the form of microswitch activations used by participants.

PRT procedures with and without embedded social interactions were compared in Koegel et al. (2009). Similar to the study conducted by Vernon et al. (2012), a motivating social interaction was included in the delivery of a non-social stimulus within the embedded PRT intervention. Vismara and Lyons (2007) also evaluated the effects of PRT with and without embedded perseverative interests with participants. Specifically, daily play activities such as puzzles, books, and card games were embedded with perseverative interests such as alphabet letters and numbers within the embedded PRT condition.

In Parsons et al. (2012) and Arbogast and Fryling (2015), two different conditions were compared in an alternating treatments fashion, and the direct effects of these conditions on indices of happiness and unhappiness were evaluated. Parsons and colleagues (2012) evaluated the effects of a “happy” condition versus an “unhappy” condition for each participant. These conditions were similar to the conditions utilized within an experimental functional analysis (e.g., a control condition and an attention condition), and were individualized for each participant. Finally, Arbogast and Fryling (2015) investigated the effects of an ABA condition versus a non-ABA condition on indices of happiness. The ABA condition included structured DTT procedures with differential reinforcement and prompting, while the non-ABA condition included natural interactions with an adult who was not an ABA therapist. A description of the comparative interventions or conditions can be found in Table 1.

**Generalization and maintenance.** A total of nine studies (31%) collected generalization and/or maintenance measures to evaluate the effects of intervention outside the training context (see Table 1). Generalization data for the mood variable were collected in four studies (Jenkins & Reed, 2013; Koegel et al., 2012; Koegel et al., 2016; Vernon et al., 2012) and maintenance data were collected in eight studies (Baker, 2000; Baker et al., 1998; Kennedy, 1994; Koegel et al., 2012; Koegel et al., 2016; Koegel et al., 1998; Stasolla et al., 2014a; Vernon et al., 2012).

**Intervention Effects on Mood**
The effects of intervention on mood were analyzed for each study (see Table 1). The results are summarized by the type of intervention procedures described above. Quantifiable data sets (i.e., mean affect scores, percentage of intervals) were reported in 75.9% \((n = 22)\) of the included studies. The remaining seven studies did not provide these data sets, but presented visual displays of the mood measurement as outcome data which could be analyzed (Dunlap & Koegel, 1980; Koegel et al., 1988; Kennedy, 1994; Koegel et al., 2012; Moes, 1998; Parsons et al., 2012; Spector & Charlop, 2017).

**Behavioral interventions.** For the 20 included studies that implemented a specific behavioral intervention or intervention package, the treatment led to increases in mood for all participants in 14 of these studies. Blakeley-Smith et al. (2009) found that by ensuring good environmental fit, the affect ratings for all participants increased. Similarly, the affect ratings for all four participants in Gengoux (2015) increased with the implementation of the priming intervention. Within Kennedy (1994), antecedent manipulations resulted in an increased frequency of positive social affect for all participants. Koegel and Egel (1979) reported that prompting procedures resulted in increased enthusiasm ratings of all three participants. However, this finding only maintained when the participants responded correctly.

In both Baker (2000) and Baker et al. (1998), the composite affect scores were higher for all participants when their perseverative interests were embedded into play activities. Moreover, Koegel et al. (2005) found that embedding social activities with reinforcing properties for both the participants and their typically-developing peers resulted in higher affect ratings for both participants. Vernon et al. (2012) recorded frequency as well as duration of each affective response and both measures increased for all three participants during the embedded PRT intervention.

Stasolla et al. (2014a) reported positive findings, concluding that the percentage of intervals with indices of happiness increased for both participants when the self-monitoring system and token economy were employed. All participants in Koegel et al. (2012) demonstrated mostly negative (sometimes neutral) affect during baseline and had increases in affect towards the positive range following “initiations training”. Koegel et al. (1998) measured the happiness of two participants in their study and found that the parent implementation of FCT led to higher happiness ratings for both participants. Spector and Charlop (2017) reported similar findings, demonstrating that sibling-mediated NLP resulted in more happiness for all three participants. Finally, Lang and colleagues (2014) found that
the intervention package employed to teach play skills resulted in increases in both the mood ratings and composite affect scores of all three participants when rated by their mothers who were blind to the conditions.

Carr and colleagues (2003) determined that the mood ratings of all participants could be systematically increased with the application of the mood induction procedure utilized in their study, and the procedure was successful for all participants in three applications or less. In effect, two-thirds of the sessions required only one or two applications of the mood induction procedure and the intervention was demonstrated to be successful in 30 minutes or less.

Mixed results were reported in five of the included studies. Moes (1998) reported that a choice intervention increased the composite affect score of three of four participants in their study. Jull and Mirenda (2011) found that child affect ratings were higher during contextually supported play dates for one of the two participants. A training package which included both video feedback and self-management increased the affect/interest scores for four of five participants in Koegel et al. (2016). Dunlap and Koegel (1980) measured a mood variable with one of the two participants in their study. During varied task instruction, the composite affect score of this participant increased with two of the three target tasks. Lattimore et al. (2009) demonstrated that indices of happiness were observed for two of the four participants within their study. During intervention, indices of happiness for these participants were noted in 17% and 100% of the observation intervals, respectively. However, the authors reported that indices of happiness were never observed for the remaining two participants, and indices of unhappiness were never observed for any of the four participants. Negative outcomes were reported in only one study (Jenkins & Reed, 2013). In this study, horseback riding led to a slight increase in indices of unhappiness, while indices of happiness were higher during baseline for all participants. The effects of the behavioral interventions on mood are described in Table 1.

Comparing two interventions/conditions. For the nine included studies that compared two or more interventions, seven of these identified a favorable treatment through a reversal or alternating treatments design. Dunlap (1984) found that varied task instruction (in comparison to constant task instruction) resulted in higher composite affect scores for all five participants. In addition, the authors discovered that varied task instruction with mastered targets led to greater increases in composite affect than varied task instruction with acquisition targets. For the single participant in Sigafoos et al. (2006), embedded instruction was found to increase the mood of the adolescent more so than traditional DTT. Alternatively, Arbogast and Fryling (2015) determined that an ABA
educational intervention condition produced more indices of happiness for both participants in comparison to a non-ABA condition. In Vismara and Lyons (2007), it was established that PRT embedded with the perseverative interests of participants led to higher affect ratings for all three children than PRT with non-perseverative stimuli. Koegel et al. (2009) also found that embedded PRT (in comparison to typical PRT) resulted in greater increases in composite affect scores for all participants. Koegel and colleagues (1988) found that composite affect scores were higher when reinforcing any verbal attempts than reinforcing successive approximations in motor speech. Finally, Stasolla et al. (2014b) reported that reinforcing adaptive responses without challenging behavior led to more intervals with indices of happiness than reinforcing adaptive responses with or without challenging behavior.

For the remaining two studies that compared two or more interventions or conditions, the results were mixed. In Geiger et al. (2012), embedded instruction (in comparison to traditional DTT) led to increases in affect for only one of two participants. More specifically, the percent of intervals with positive affect increased more during embedded instruction for one participant, while this percentage increased more during traditional DTT for the other participant. In this study, the percent of intervals with negative affect increased for both participants during traditional DTT (Geiger et al., 2012).

Parsons et al. (2012) found that the “happy” condition appeared to increase the indices of happiness for only two of their three participants. For these two participants, the percentage of indices of happiness increased, while the percentage of indices of unhappiness decreased, during the “happy” condition. During the “unhappy” condition for these two participants, the opposite was found (i.e., indices of happiness decreased while indices of unhappiness increased). The results for each comparative intervention or condition are outlined in Table 1.

**PND.** The PND scores are reported in Table 1. The PND score could not be calculated for 11 studies (37.9%) due to either a missing baseline or visual display of data (Blakeley-Smith et al., 2009; Carr et al., 2003; Dunlap, 1984; Geiger et al., 2012; Jull & Mirenda, 2011; Koegel & Egel, 1979; Koegel et al., 2016; Koegel et al., 1998; Lang et al., 2014; Lattimore et al., 2009; Spector & Charlop, 2017). Of the remaining 18 studies, eight were determined to be highly effective (Arbogast & Fryling, 2015; Kennedy, 1994; Koegel et al., 2012; Koegel et al., 2009; Koegel et al., 2005; Stasolla et al., 2014a; Stasolla et al., 2014b; Vernon et al., 2012), while two studies were determined to be moderately effective (Baker et al., 1998; Parsons et al., 2012). Four studies were considered mildly
effective (Baker, 2000; Dunlap & Koegel, 1980; Koegel et al., 1988; Moes, 1998), and four studies were found to be ineffective (Gengoux, 2015; Jenkins & Reed, 2013; Sigafoos et al., 2006; Vismara & Lyons, 2007).

**Generalization and maintenance.** Nine studies collected data on the generalization and/or maintenance of the mood measurement (Baker, 2000; Baker et al., 1998; Jenkins & Reed, 2013; Kennedy, 1994; Koegel et al., 2012; Koegel et al., 2016; Koegel et al., 1998; Stasolla et al., 2014a; Vernon et al., 2012). Generalization probes were collected in four studies. Jenkins and Reed (2013) found that the intervention did not produce changes in indices of happiness or unhappiness during home visit probes with the participants. Conversely, Koegel et al. (2012) found that the affect ratings generalized to settings where the interventionist was not present for all three participants. Vernon et al. (2012) also found that positive affect generalized to non-treatment settings for all participants. Koegel et al. (2016) measured the affect/interest of four of the five participants in their study and concluded that the results generalized to probes in which participants had a conversation with a typically-developing peer in the natural environment.

Maintenance data were collected in seven included studies. Composite affect scores remained high during a one-month and three-month follow-up in Baker (2000), and a one-month and two-month follow-up in Baker et al. (1998). Positive social affect continued to occur more frequently during a four-month follow-up in Kennedy (1994). Positive affect also occurred more frequently and for longer durations one week after intervention for all three participants in Vernon et al. (2012). A maintenance check was implemented for one of the participants in the study reported by Koegel et al. (2012) and it showed that the affect ratings maintained three months following intervention for this individual. Koegel et al. (2016) collected maintenance data for four of the five participants in their study and reported that increases in affect/interest continued during a two-month follow-up. Finally, Koegel et al. (1998) concluded that happiness ratings remained high during a six-month and one-year follow-up for one participant, and a four-month follow-up for the second participant. The results from the generalization and maintenance measures can be found in Table 1.

**Scientific Merit Rating Scale (SMRS) Score**

The SMRS was used to evaluate the experimental rigor of each study (see Table 1). Of the 29 studies included in the current review, one received an SMRS score of 4 (Spector & Charlop, 2017) and none received an SMRS score of 5. Nine studies (31%) received an SMRS score of 3, while a majority of studies ($n = 18; 62.1\%$)...
received an SMRS score of 2. One study received an SMRS score of 1 (Arbogast & Fryling, 2015) and none of the studies received a score of 0.

Discussion

Within recent literature, there has been a growing trend in the promotion and measurement of QoL in people with disabilities, and researchers have proposed that one of the main components of QoL is individual happiness (Lancioni et al., 2005). However, to date, there remains a paucity of studies that have specifically measured happiness as a dependent variable in behavior analytic research for people with disabilities (Green & Reid, 1996, 1999). Within the few published reviews that have focused on measuring indicators of mood, all participants have had a primary diagnosis of severe/profound ID or PMD (see Dillon & Carr, 2007; Lancioni et al., 2005). Therefore, this review offers an examination of the current state of literature involving mood as a dependent variable in behavioral interventions for individuals with ASD.

Of the 29 published studies included in this review, 21 studies (72.4%) reported positive effects on mood for all participants. Of the 20 studies that implemented a single behavioral intervention or intervention package, 14 reported positive effects on the mood of all participants. These results suggest that certain behavioral interventions may influence the mood of participants with ASD. In some cases, change in mood was directly targeted (Carr et al., 2003), and in other studies mood was conceptualized as a collateral behavior change not directly targeted by intervention components (Lang et al., 2014). Regardless of whether mood change was an intended intervention outcome, this review supports the conclusion that intervention may influence mood. However, these findings need to be considered with caution due to several limitations found within the literature.

First, because only SSED’s were examined in this review, the number of participants were limited in each study. Further, only nine of the reviewed studies (31%) included some type of generalization or maintenance measure. Therefore, generalization of this review’s findings to larger groups of individuals with ASD remains uncertain and the extent to which desirable changes persist over time requires additional research.

A second limitation is that the PND scores could not be calculated for over a third of the studies (n = 11; 37.9%); therefore, the efficacy of the interventions employed within these studies could not be evaluated. Furthermore, of the 18 studies with a PND score calculation, four were shown to be “mildly effective” and four were
found to be “ineffective” in affecting mood. Further, because changes in mood and affect were not always the primary intervention goal, research controls (e.g., IOA, minimum number of data points per phase) were focused on the other dependent variables. Therefore, definite conclusions regarding the influence of interventions on mood remain tentative and should be considered with caution.

A third limitation was that over half of the studies (n = 18; 62.1%) yielded an SMRS score of 2, which suggested that the methodological rigor of these studies was questionable, and more rigorous research is necessary to make definitive conclusions regarding the effects of behavioral interventions on mood. Only one study in the current review demonstrated an SMRS score of four or five (Spector & Charlop, 2017).

A final limitation found within the research was that only five of the studies directly targeted mood as the primary dependent variable (Arbogast & Fryling, 2015; Carr et al., 2003; Jenkins & Reed, 2013; Koegel et al., 1988; Parsons et al., 2012). All other studies measured mood as a secondary or collateral dependent variable (Ledbetter-Cho, Lang, Watkins, O’Reilly, & Zamora, 2017). It could be argued that had mood been the primary dependent variable in the remaining 24 studies, the intervention effects may have been more pronounced. However, the results of this current review indicate that more research targeting mood as the primary dependent variable is warranted.

Several important findings emerged from this review. Specifically, 19 of the studies (65.5%) used a Likert-type scale to code the mood of participants. One could argue that these types of scales are less rigorous than other direct measurement procedures such as direct observation employing frequency or interval recording. In addition, these mood scales frequently used mentalistic language such as “appears frustrated” (Dunlap, 1984) or “seems to be enjoying self” (Dunlap & Koegel, 1980). Mentalistic terms such as “frustration” and “enjoyment” are not phenomena that can be objectively measured, and they are considered incomplete explanations to human behavior (Hayes & Brownstein, 1986). Objective description and measurement of behavior is considered best practice in ABA. Rather than relying solely on emotional referents such as “frustrated”, an interventionist should describe and measure the behavior in observable terms (e.g., “the participant screamed, cried, threw body onto floor”). Although some studies used considerably less-rigorous measurement procedures, of the 28 studies that reported IOA data, 27 reported to have good reliability within their measurement procedures. This finding suggests that different types of measurement systems (e.g., rating scales, direct observation) can yield reliable measures of mood.
One of the most notable findings is that individualized indices of happiness were only used in two of the studies (Lattimore et al., 2009; Parsons et al., 2012). It is hypothesized that individuals with disabilities engage in similar behavioral indicators of mood as individuals without disabilities (e.g., smiling, laughing, frowning) and that the same measures of mood used with individuals without disabilities could be applied to individuals with disabilities (Green & Reid, 1996; Lancioni et al., 2005). This assumption stems from the fact that individuals who lack communication skills are less likely to purposefully display an emotion that they are not feeling, which is considered a more advanced social skill (Green & Reid, 1996, 1999). Nevertheless, individuals with ASD do tend to demonstrate idiosyncratic overt behaviors that differ from those of their typically-developing peers (Donnellan, Hill, & Leary, 2013). For example, Faso, Sasson, and Pinkham (2015) found that the emotional expressivity of six adults with ASD was rated as more intense and less natural than their neurotypical peers. Furthermore, laughter, which would typically be correlated with a good mood, could be classified as a form of vocal stereotypy in a person with ASD (Ahearn, Clark, & MacDonald, 2007). As a result, it would be appropriate to individualize the operational definition of mood for each participant, just like with any other target behavior.

It should be noted that all measurement systems described in this literature review only provided indirect evidence of the mood of the participants (Lang et al., 2014). From a behavioral perspective, mood is considered a private event that cannot be directly observed. Behavioral indicators of mood are only presumed to represent the private event in question, and due to this correlational relationship, conclusions regarding the internal feelings of an individual cannot be made with certainty (Green & Reid, 1996). For instance, an individual with ASD may engage in conventional indicators of happiness (e.g., smiling, laughter) as a form of stereotypy or echolalia, but he/she may not feel happy. In fact, research has suggested that individuals with ASD may engage in stereotypy more frequently and for longer durations during times of distress than during times of elation (Willemsen-Swinkels, Buitelaar, Dekker, & van Engeland, 1998). To overcome this issue, future research should include physiological measures (e.g., electrodermal activity) or modified self-reports for individuals with ASD to validate the behavioral observations of mood (Adams & Oliver, 2011; Lancioni et al., 2005).

Based on this literature review, a few recommendations can be made regarding future research involving behavioral indicators of mood as a dependent variable in interventions for individuals with ASD. First, more research should focus on mood as the primary dependent variable as to determine the direct effects various
interventions have on the wellbeing of individuals with ASD. Second, the operational definitions of mood for each participant should be individualized so that idiosyncratic indicators of mood are accounted for across participants. Like other target behaviors, behavioral definitions of mood need to be individually tailored for individuals who may express their feelings in unconventional ways. Third, more rigorous research designs and observational methods should be employed in studies measuring mood changes as a result of behavioral interventions. Generalization and maintenance measures should also be considered for any intervention that is targeting mood. Finally, the observational method of mood should be paired with a self-report or physiological measure to further validate the outcomes of intervention.
On behalf of all authors, the corresponding author states that there is no conflict of interest.
MEASURING MOOD IN BEHAVIORAL INTERVENTIONS

References


Language Impairments, 2*, 1-22.

youth and young adults with chronic conditions and disabilities. *Remedial and Special Education, 31*, 252-260.

* Moes, D. R. (1998). Integrating choice-making opportunities within teacher-assigned academic tasks to facilitate
the performance of children with autism. *Journal of the Association for Persons with Severe Handicaps*,
23, 319-328.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and


and unhappiness among adults with autism: Potential targets for behavioral assessment and intervention.
*Behavior Analysis in Practice, 5*, 15-25.

(Ed.), *Antecedent assessment & intervention: Supporting children & adults with developmental disabilities

American Association on Mental Retardation.

systematic reviews and meta-analyses: A systematic review of reporting characteristics. *Evidence-Based
Communication Assessment and Intervention, 2*, 163-187.


*Studies included in this review.
Records identified through database searching
(n = 12,536)

Additional records identified through references
(n = 65)

Records after duplicates removed
(n = 12,601)

Records screened
(n = 12,601)

Records excluded
(n = 11,929)

Articles assessed for eligibility
(n = 672)

Articles excluded, with reasons
(n = 643)

Studies included in qualitative synthesis
(n = 29)
Figure 2

Total initial studies
N = 672

No ASD diagnosis; comorbid diagnosis of mood disorder
N = 619

Not a single-subject experimental design
N = 418

Used questionnaires, surveys, inventories, or temperament measures
N = 305

Measuring and/or teaching ToM, empathy, emotion recognition or regulation; teaching affective behavior
N = 287

Pharmacological treatment or brain imaging study
N = 282

Literature review, discussion, or dissertation
N = 177

No data presented; no operational definition for mood; mood not measured
N = 29
Figure Captions

Figure 1. Flow diagram showing the identification and screening process of studies identified during database search process.

Figure 2. Flow diagram showing inclusion/exclusion of studies identified during database search process.
<table>
<thead>
<tr>
<th>Study</th>
<th>Participant(s) &amp; Setting</th>
<th>Experimental Design</th>
<th>Type of Mood Measurement</th>
<th>Intervention</th>
<th>Outcomes for Mood</th>
<th>SMRS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arbogast &amp; Fryling (2015)</td>
<td>N = 2 (1 male); 6-9 yo (M = 7:6); both with autism</td>
<td>ATD</td>
<td><strong>Indices of happiness</strong> – 10s partial interval recording for any facial expression or vocalization typically associated with happiness in people without disabilities (e.g., smiling &amp; laughing)</td>
<td>Compared two conditions (a) Non-ABA condition: access to preferred items/activities &amp; an adult not associated with ABA therapy; &amp; (b) ABA condition: preferred items/activities &amp; an ABA therapist providing ABA-based interventions (e.g., DTT)</td>
<td>Indices of happiness were higher in the ABA condition (M = 44.8%; range = 23-90%) than the non-ABA condition (M = 14.1%; range = 3-25%) for both participants</td>
<td>1</td>
</tr>
<tr>
<td>Baker (2000)</td>
<td>N = 3 (2 males); 5-6 yo (M = 5:10); all with autism</td>
<td>MB across participants</td>
<td><strong>Child affect</strong> – composite affect score obtained using the GCARS (consisting of two Likert subscales for interest &amp; happiness)</td>
<td>Incorporating participants’ PI into a Bingo game, prompts &amp; prompt fading</td>
<td>Child affect was higher during intervention (M = 4; range = 3.6-4.3) than baseline (M = 2.2; range = 2-2.5) for all participants</td>
<td>3</td>
</tr>
<tr>
<td>Baker et al. (1998)</td>
<td>N = 3 (1 male); 5-8 yo (M = 7:2); all with autism</td>
<td>MB across participants</td>
<td><strong>Child affect</strong> – composite affect score obtained using the GCARS (consisting of two Likert subscales for interest &amp; happiness)</td>
<td>Incorporating participants’ PI into common playground games (e.g., tag, follow-the-leader), prompts &amp; prompt fading</td>
<td>Child affect was higher during intervention (M = 4.3; range = 3.5-4.8) than baseline (M = 2.2; range = 1.7-2.9) for all participants</td>
<td>2</td>
</tr>
<tr>
<td>Blakeley-Smith et al. (2009)</td>
<td>N = 6 (3 males); 4-13 yo (M = 8:4); 4 with autism, 2 with Asperger’s</td>
<td>MB across participants</td>
<td><strong>Affect</strong> – affect rating obtained using the MSRIQ (consisting of one Likert scale)</td>
<td>Task analyses, modifying task demands to meet participants’ competency level (i.e., good environmental fit)</td>
<td>Affect ratings were higher during intervention (M = 2.9; range = 2.8-3.1) than baseline (M = 1.4; range = 0.8-2.4) for all participants</td>
<td>2</td>
</tr>
<tr>
<td>Study</td>
<td>N</td>
<td>Setting</td>
<td>Intervention Details</td>
<td>Mood Description</td>
<td>Child Affect Description</td>
<td>Positive affect comparison</td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
<td>---------</td>
<td>----------------------</td>
<td>-----------------</td>
<td>--------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Carr et al. (2003)</td>
<td>N = 8* (5 males); 29-48 yo (M = 35.5); 5 with autism; all with mild-profound ID</td>
<td>MB across participants</td>
<td>Mood – pre-post mood rating obtained using the MSRIQ (consisting of one Likert scale)</td>
<td>Mood scale included (a) Bad mood: yelling, pouting, appears irritable, angry, or frustrated, not enjoying things (0-1); (b) Neutral: may smile or frown occasionally, overall seems neutral (2-3); Good mood: smiles, laughs, enjoying things (4-5)</td>
<td>A pre-intervention mood rating of 0-1 (bad) resulted in a 15-min mood induction procedure (i.e., stimuli associated with good mood delivered prior to task presentation), a post-intervention mood rating of 3-5 (neutral/good) resulted in termination of intervention &amp; presentation of task demand, a post-intervention mood rating of less than 3 resulted in the repetition of the mood induction procedure</td>
<td>For all participants, post-intervention mood ratings of 3-5 were obtained following only one (M = 33.5% of sessions; range = 16.7-62.5%), two (M = 37.8% of sessions; range = 6.3-57.2%), or three (M = 28.7% of sessions; range = 21.4-33.3%) presentations of the intervention PND: N/A; MC: NR; G: NR</td>
</tr>
<tr>
<td>Dunlap (1984)</td>
<td>N = 5 (4 males); 4-10 yo (M = 7.7); all with autism</td>
<td>ATD</td>
<td>Child affect – composite affect score obtained using the GCARS (consisting of four Likert subscales for enthusiasm, interest, happiness, &amp; general behavior)</td>
<td>Happiness subscale included (a) Unhappy: cries, pouts, tantrums, appears sad, angry, or frustrated, not enjoying self (0-1); (b) Neutral: may smile or frown occasionally, overall seems neutral (2-3); &amp; (c) Happy: smiles, laughs, enjoying self (4-5)</td>
<td>Compared three conditions (a) Constant task condition: massed practice approach, prompts &amp; prompt fading, differential reinforcement; (b) Varied-acquisition-task condition: random rotation of tasks, prompts &amp; prompt fading, differential reinforcement; &amp; (c) Varied-with-maintenance-task condition: random rotation of tasks with maintenance tasks, prompts &amp; prompt fading, differential reinforcement</td>
<td>Child affect was higher during the varied-with-maintenance-task condition (M = 3.6; range = 3.0-4.0) than both the constant task (M = 2.5; range = 1.7-2.9) &amp; varied-acquisition-task (M = 3.1; range = 2.6-3.7) conditions for all participants PND: N/A; MC: NR; G: NR</td>
</tr>
<tr>
<td>Dunlap &amp; Koegel (1980)</td>
<td>N = 2 (both females); 5-7 yo (M = 6.3); both with autism</td>
<td>ATD</td>
<td>Child affect – composite affect score obtained for one participant only using the GCARS (consisting of four Likert subscales for enthusiasm, interest, happiness, &amp; general behavior)</td>
<td>Happiness subscale included (a) Unhappy: cries, pouts, tantrums, appears sad, angry, or frustrated, not enjoying self (0-1); Neutral: may smile or frown occasionally, overall seems neutral (2-3); Happy: smiles, laughs, enjoying self (4-5)</td>
<td>Random rotation of tasks, prompts, reinforcement</td>
<td>Child affect was higher during intervention (neutral/positive) than baseline (neutral/negative) for two of the three tasks for the participant PND: Mildly effective; MC: NR; G: NR</td>
</tr>
<tr>
<td>Geiger et al. (2012)</td>
<td>N = 2 (both males); 4 yo; both with autism</td>
<td>ATD</td>
<td>Affective behavior – 15s partial interval recording for positive affect (e.g., smiling, laughing, clapping, positive statements) &amp; negative affect</td>
<td>Compared two conditions (a) Traditional DTT: presentation of S0, prompts, reinforcement; &amp; (b) Embedded DTT: incorporating participants’ perseverative interests into S0, prompts, reinforcement</td>
<td>Positive affect was higher in the traditional DTT condition (M = 48%; range = 17-100%) than the embedded DTT condition (M = 34%; range = 18-66%) for participants throughout the intervention</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Sample Size</td>
<td>Setting</td>
<td>Child affect</td>
<td>Priming of game/activity to be used during next day probe in the inclusive setting, reinforcement to teach game, prompts to initiate play during probe</td>
<td>Participant 1, while positive affect was higher in the embedded DTT condition ($M = 44%$; range = 5-81%) than the traditional DTT condition ($M = 17%$; range = 0-50%) for Participant 2; Negative affect was higher in the traditional DTT condition ($M = 11%$; range = 0-40%; $M = 9%$; range = 0-77%) than the embedded DTT condition ($M = 3%$; range = 0-11%; $M = 1%$; 0-14%) for both participants</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>--------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gengoux (2015)</td>
<td>N = 4 (3 males); 6-7 yo (M = 6:11); all with autism</td>
<td>MB across participants</td>
<td><strong>Child affect</strong> – affect rating obtained using the GCARS (consisting of one Likert scale)</td>
<td>Priming of game/activity to be used during next day probe in the inclusive setting, reinforcement to teach game, prompts to initiate play during probe</td>
<td>Child affect was higher during intervention ($M = 4$; range = 2-5) than baseline ($M = 3.2$; range = 1-5) for all participants</td>
<td></td>
</tr>
<tr>
<td>Setting: Summer day camp or afterschool program</td>
<td>Affect scale included (a) <strong>Negative affect</strong>: appears discontent (i.e., frowns, cries), avoids participation (i.e., tantrums, leaves activity), not enjoying self (i.e., frustrated, tense, impatient) (0-1); <strong>Neutral affect</strong>: may engage in activity but not interested or enthusiastic, not stressed or relaxed (2-3); <strong>Positive affect</strong>: enjoying self (i.e., smiles, laughs), interested (i.e., actively involved), appears relaxed &amp; comfortable (4-5)</td>
<td>PND: N/A; MC: NR; G: NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jenkins &amp; Reed (2013)</td>
<td>N = 7* (6 males); 6-14 yo (M = 9:6); all with ASD, 2 with tuberous sclerosis, 1 with verbal &amp; motor apraxia</td>
<td>MB across participants</td>
<td><strong>Indices of happiness &amp; unhappiness</strong> – 10s MTS for any facial expression or vocalization typically associated with happiness (e.g., smiling &amp; laughing) or unhappiness (e.g., frowning &amp; crying) in people without disabilities</td>
<td>Therapeutic horseback riding</td>
<td>Indices of happiness were higher during baseline ($M = 6%$; range = 0-37%) than intervention ($M = 2.8%$; range = 0-25%) for all participants; Indices of unhappiness were slightly higher during intervention ($M = 0.2%$; range = 0-6%) than baseline ($M = 0.2%$; range = 0-5%) for all participants</td>
<td></td>
</tr>
<tr>
<td>Setting: Afterschool program</td>
<td>*Only 4 participants included in the intervention</td>
<td></td>
<td></td>
<td>PND: Ineffective; MC: NR; G: NR; Intervention did not produce changes in indices of happiness or unhappiness during home visit probes</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>N = 2 (both males); 4-5 yo (M = 5:2); both with autism</td>
<td>Setting: Home</td>
<td>Child affect – affect rating obtained using the GCARS (consisting of one Likert scale)</td>
<td>Parent-implemented contextually supported play dates with TD peer</td>
<td>Child affect was higher during intervention (M = 3.5; range = 3-4) than baseline (M = 2.15; range = 2-2.3) for only one participant PND: N/A; MC: NR; G: NR</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------------------------------</td>
<td>---------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Jull &amp; Mirenda (2011)</td>
<td></td>
<td>Reversal</td>
<td>Affect scale included (a) Negative affect: appears discontent (i.e., frowns, cries), avoids participation (i.e., tantrums, leaves activity), not enjoying self (i.e., frustrated, tense, impatient) (0-1); Neutral affect: may engage in activity but not interested or enthusiastic, not stressed or relaxed (2-3); Positive affect: enjoying self (i.e., smiles, laughs), interested (i.e., actively involved), appears relaxed &amp; comfortable (4-5)</td>
<td>Parent-implimented contextually supported play dates with TD peer</td>
<td>Child affect was higher during intervention (M = 3.5; range = 3-4) than baseline (M = 2.15; range = 2-2.3) for only one participant PND: N/A; MC: NR; G: NR</td>
<td></td>
</tr>
<tr>
<td>Kennedy (1994)*</td>
<td>N = 3 (2 males); all 20 yo; 1 with autism, 1 with cerebral palsy, 1 with quadriaparesis &amp; hydrocephalus, all with moderate-profound ID</td>
<td>MP across participants</td>
<td>Positive social affect – frequency per minute for smiling laughing, nodding “yes”, &amp; positive verbalizations</td>
<td>Antecedent manipulations (i.e., low demand, high comment), gradual increase of demands if problem behavior remained low</td>
<td>Positive social affect was higher during intervention than baseline for all participants PND: Highly effective; MC: Maintained during 4-mo follow-up; G: NR</td>
<td></td>
</tr>
<tr>
<td>Koegel &amp; Egel (1979)</td>
<td>N = 3 (2 males); 6-12 yo (M = 10:1); all with autism</td>
<td>Setting: NR</td>
<td>Enthusiasm – enthusiasm rating obtained using one Likert scale</td>
<td>Prompting (verbal prompting for 2 participants, prompting package for third participant)</td>
<td>Enthusiasm was higher in intervention (M = 4.1; range = 4-5) than baseline (M = 0.8; range = 0-1) for all participants when their responding remained correct PND: N/A; MC: NR; G: NR</td>
<td></td>
</tr>
<tr>
<td>Koegel et al. (2012)</td>
<td>N = 3 (2 males); 5-6 yo (M = 5.4); 2 with autism, 1 with Asperger's</td>
<td>Nonconcurrent MB across participants</td>
<td><strong>Child affect</strong> – affect rating obtained using the GCARS (consisting of one Likert scale)</td>
<td>Facilitated social play with &amp; without initiations training (i.e., prompts to begin game or join game)</td>
<td>Child affect was higher during intervention (neutral/positive) than baseline (neutral/negative) for all participants</td>
<td>PND: Highly effective; MC: Maintenance data only collected for one participant, maintained during 3-mo follow-up; G: Generalized to settings where interventionist was not present</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>Setting: School</td>
<td></td>
<td>Affect scale included (a) <strong>Negative affect</strong>: bored or uninvolved, not eager to participate, appears sad, angry, or frustrated, not attending, noncompliant (i.e., inappropriate verbal or motor behavior, runs away), disruptive behaviors (e.g., tantrums, aggression) (0-1); <strong>Neutral affect</strong>: uninterested in activity, may smile or frown occasionally, overall seems neutral, may fidget, inattentive, generally compliant but not readily (2-3); <strong>Positive affect</strong>: readily attends, enjoying self, smiles, laughs, or shows other positive emotional behavior, alert &amp; involved, compliant (4-5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koegel et al. (2016)</td>
<td>N = 5 (3 males); 20-37 yo (M = 27.9); all with ASD</td>
<td>MB across participants</td>
<td><strong>Affect and conversational interest</strong> – affect/interest rating obtained using one Likert scale</td>
<td>Initial training, video-feedback, conversational probes with self-management</td>
<td>Affect/interest ratings were higher during intervention (M = 2.8; range = 2.4-3.1) than baseline (M = 2.2; range = 1.9-2.3) for four of the participants</td>
<td>PND: N/A; MC: Maintenance data only collected for four participants, maintained during 2-mo follow-up; G: Generalized to TD peer in the natural environment</td>
</tr>
<tr>
<td></td>
<td>Setting: University clinic</td>
<td></td>
<td>Affect/interest included (a) <strong>Low affect/interest</strong>: appears discontent, not enjoying self (0-1); <strong>Medium affect/interest</strong>: may smile or frown occasionally, overall seems neutral (2-3); <strong>High affect/interest</strong>: smiles, laughs, enjoying self (4-5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koegel et al. (1988)</td>
<td>N = 4 (3 males); 3-11 yo (M = 8.1); 3 with autism, 1 with developmental delays</td>
<td>Reversal</td>
<td><strong>Child affect</strong> – composite affect score obtained using the GCARS (consisting of four Likert subscales for enthusiasm, interest, happiness, &amp; general behavior)</td>
<td>Comparing two conditions (a) <strong>Motor speech</strong>: reinforcement of successive improvements in motor speech (i.e., shaping); &amp; (b) <strong>Verbal attempts</strong>: reinforcement of any verbal attempts</td>
<td>Child affect was higher in the verbal attempts condition (neutral/positive) than the motor speech condition (neutral/negative) for all participants</td>
<td>PND: Mildly effective; MC: NR; G: NR</td>
</tr>
<tr>
<td></td>
<td>Setting: University clinic</td>
<td></td>
<td>Happiness subscale included (a) <strong>Unhappy</strong>: cries, pouts, tantrums, appears sad, angry, or frustrated, not enjoying self (0-1); (b) <strong>Neutral</strong>: may smile or frown occasionally, overall seems neutral (2-3); &amp; (c) <strong>Happy</strong>: smiles, laughs, enjoying self (4-5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Sample Size</td>
<td>Setting</td>
<td>Method</td>
<td>Child Affect</td>
<td>Results</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>---------</td>
<td>--------</td>
<td>--------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Koegel et al. (1998)</td>
<td>N = 3 (2 males); 4-5 yo (M = 5); 2 with autism, 1 with mixed developmental disorder</td>
<td>MB across participants</td>
<td><strong>Child happiness</strong> – happiness rating obtained for two participants using the GCARS (consisting of one Likert scale) Happiness scale included (a) Unhappy: cries, pouts, tantrums, appears sad, angry, or frustrated, not enjoying self (0-1); (b) Neutral: may smile or frown occasionally, overall seems neutral (2-3); &amp; (c) Happy: smiles, laughs, enjoying self (4-5)</td>
<td>Parent training to teach parents how to implement environmental arrangements &amp; FCT, prompts &amp; prompt fading used by clinician</td>
<td>Happiness was higher during intervention (M = 3.3; range = 3-3.5) than baseline (M = 2.3; range = 2-2.5) for both participants PND: N/A; MC: Maintained during 6-mo &amp; 1-yr follow-up for Participant 1 &amp; 4-mo follow-up for Participant 2; G: NR</td>
<td></td>
</tr>
<tr>
<td>Koegel et al. (2009)</td>
<td>N = 3 (all males); 3:2-3:5 yo (M = 3:3); all with autism</td>
<td>Reversal; embedded ATD for one participant</td>
<td><strong>Child affect</strong> – composite affect score obtained using the GCARS (consisting of two Likert subscales for interest &amp; happiness) Happiness subscale included (a) Unhappy: cries, pouts, tantrums, appears sad, angry, or frustrated, not enjoying self (0-1); (b) Neutral: may smile or frown occasionally, overall seems neutral (2-3); &amp; (c) Happy: smiles, laughs, enjoying self (4-5)</td>
<td>Comparing two conditions (a) Non-embedded: PRT paradigm followed; &amp; (b) Embedded: PRT paradigm followed, social interaction embedded into reinforcement</td>
<td>Child affect was higher in the embedded condition (M = 4.3; range = 3.5-5) than the non-embedded condition (M = 2.5; range = 1-3) for all participants PND: Highly effective; MC: NR; G: NR</td>
<td></td>
</tr>
<tr>
<td>Koegel et al. (2005)</td>
<td>N = 2 (1 male); 8-9 yo (M = 8:6); both with autism</td>
<td>MB across participants; embedded reversal for one participant</td>
<td><strong>Child affect</strong> – affect rating obtained using the GCARS (consisting of one Likert scale) Affect scale included (a) Negative affect: appears discontent (i.e., frowns, cries), avoids participation (i.e., tantrums, leaves activity), not enjoying self (i.e., frustrated, tense, impatient) (0-1); Neutral affect: may engage in activity but not interested or enthusiastic, not stressed or relaxed (2-3); Positive affect: enjoying self (i.e., smiles, laughs), interested (i.e., actively involved), appears relaxed &amp; comfortable (4-5)</td>
<td>Embedding social activities with mutually reinforcing properties for both the participant &amp; TD peer, cooperative arrangements set up by adults</td>
<td>Child affect was higher during intervention (M = 4.1; range = 4-5) than baseline (M = 2.9; range = 2-3) for both participants PND: Highly effective; MC: NR; G: NR</td>
<td></td>
</tr>
<tr>
<td>Lang et al. (2014)</td>
<td>N = 3 (2 males); 3:6-3:10 years old (M = 3:8); all with autism</td>
<td>MB across participants; embedded reversal for two participants</td>
<td><strong>Child mood</strong> – mood rating obtained using the MSRIQ (consisting of one Likert scale)</td>
<td>Play intervention (prompts, RIRD, &amp; reinforcement) with &amp; without lag schedules of reinforcement</td>
<td>Child mood was higher during intervention (M = 3.7; range = 3-4) than baseline (M = 2.8; range = 2-3) for all participants; Child affect was also higher during</td>
<td></td>
</tr>
</tbody>
</table>
Mood scale included (a) **Bad mood**: yelling, pouting, appears irritable, angry, or frustrated, not enjoying things (0-1); (b) **Neutral**: may smile or frown occasionally, overall seems neutral (2-3); **Good mood**: smiles, laughs, enjoying things (4-5)

**Child affect** – composite affect score obtained using the GCARS (consisting of two Likert subscales for interest & happiness)

Happiness subscale included (a) **Unhappy**: cries, pouts, tantrums, appears sad, angry, or frustrated, not enjoying self (0-1); (b) **Neutral**: may smile or frown occasionally, overall seems neutral (2-3); & (c) **Happy**: smiles, laughs, enjoying self (4-5)

<table>
<thead>
<tr>
<th><strong>Lattimore et al. (2009)</strong></th>
<th>N = 4 (all males); 31-44 years old (M = 37); all with autism &amp; profound ID</th>
<th><strong>Indices of happiness &amp; unhappiness</strong> - 15s partial interval recording for individualized indices of three participants</th>
<th>Intensive teaching - repeated trials, task analyses &amp; chaining, prompts, programming common stimuli, reinforcement</th>
</tr>
</thead>
</table>
| **Setting**: Residential facility & supported work environment | MP across participants | Happiness – Participant 1: any facial expression or vocalization typically associated with happiness in people without disabilities (e.g., smiling & laughing); Participant 2: patting person on the back, laughing, smiling; Participant 3: making “wee/mee” vocalizations, smiling | During intervention, indices of happiness were observed in two participants only (M = 58.5%; range = 17-100%) & indices of unhappiness were not observed in any participants
PND: N/A; MC: NR; G: NR |
| **Intervention** | | Unhappiness – Participant 1: any expression or vocalization typically associated with unhappiness in people without disabilities (e.g., frowning, crying); Participant 2: hitting head, pressing finger on eye, signing “finish”, turning over furniture, physical with staff; Participant 3: Yelling, biting hand, hitting head, head butting others | 2 |

Intervention (M = 3.8; range = 3-5) than baseline (M = 2.9; range = 2.5-3) for all participants
PND: N/A; MC: NR; G: NR
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Setting/Environment</th>
<th>Behavioral Intervention</th>
<th>Mood Measure</th>
<th>Comparison of Conditions</th>
<th>PND/Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moes (1998)</td>
<td>N = 4 (all males); 5-9 yo (M = 7:6); all with autism</td>
<td>University clinic</td>
<td>Reversal</td>
<td>Child affect – composite affect score obtained for one participant only using the GCARS (consisting of three Likert subscales for enthusiasm, interest, &amp; happiness)</td>
<td>Child choice for order of activities, order of specific items/questions, &amp; stimulus materials, prompts &amp; reinforcement for correct responses</td>
<td>Child affect was higher during intervention (neutral/positive) than baseline (neutral/negative) for three participants PND: Mildly effective; MC: NR; G: NR</td>
</tr>
<tr>
<td>Parsons et al. (2012)</td>
<td>N = 3 (all males); 22-41 years old (M = 33:4); all with autism &amp; severe ID, 1 with hearing loss, 1 with a seizure disorder</td>
<td>Supported work environment, group home, &amp;/or residential facility</td>
<td>ATD</td>
<td>Indices of happiness &amp; unhappiness - 10s partial interval recording for individualized indices of three participants</td>
<td>Comparing two conditions (a) Happy: Participant 1 – providing a sketch pad to draw, Participant 2 – providing a break in the lounge, Participant 3 – provided with a leisure item to hold, social interaction; &amp; (b) Unhappy: Participant 1 – sitting at table with no materials present, no social interaction, Participant 2 – demands given with prompting &amp; verbal feedback, Participant 3 – not provided with leisure item to hold, no social interaction</td>
<td>Indices of happiness were higher in happy condition than unhappy condition for two participants; Indices of unhappiness were higher in unhappy condition than happy condition for two participants PND: Moderately effective; MC: NR; G: NR</td>
</tr>
<tr>
<td>Sigafoos et al. (2006)</td>
<td>N = 1 (male); 12 yo; autism &amp; severe ID</td>
<td>School</td>
<td>Reversal</td>
<td>Mood – scored at the end of each interval as “bad”, “neutral”, or “good”</td>
<td>Comparing two conditions (a) Embedded instruction: manding opportunities embedded within activities, prompts, reinforcement; &amp; (b) DTT: imitation or receptive trials, prompts, reinforcement</td>
<td>Mood was higher during the embedded instruction condition (M = 2.3; range = 2-3) than the DTT condition (M = 1.5; range = 1-2.3) for the participant PND: Ineffective; MC: NR; G: NR</td>
</tr>
<tr>
<td>Study</td>
<td>N</td>
<td>Setting</td>
<td>Design</td>
<td>Description</td>
<td>Outcomes</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Spector &amp; Charlop (2017)</td>
<td>N = 3 (all males); 6-9 yo (M = 7); all with ASD</td>
<td>Nonconcurrent MB across participants</td>
<td><strong>Happiness</strong> – 10s partial interval recording for smiling (i.e., corners of mouth turned upwards)</td>
<td>Sibling-mediated Natural Language Paradigm (NLP)</td>
<td>Happiness was higher during intervention than baseline for all participants PND: N/A; MC: NR; G: NR</td>
<td></td>
</tr>
<tr>
<td>Stasolla et al. (2014a)</td>
<td>N = 2 (both males); 7:6-8:6 yo (M = 8); both with ASD &amp; presenting symptoms of ADHD</td>
<td>Nonconcurrent MB across participants</td>
<td><strong>Indices of happiness</strong> – 10s partial interval recording for smiling, laughing, excited body movements with or without vocalizations, singing</td>
<td>Self-monitoring, token economy</td>
<td>Indices of happiness were higher during intervention (M = 89.9%; range = 66-100%) than baseline (M = 29.7%; range = 20-38%) for both participants PND: Highly effective; MC: Maintained during 1-mo follow-up; G: NR</td>
<td></td>
</tr>
<tr>
<td>Stasolla et al. (2014b)</td>
<td>N = 3 (all males); 8-10 yo (M = 9:4); all with ASD &amp; severe-profound ID</td>
<td>Reversal</td>
<td><strong>Indices of happiness</strong> – 15s partial interval recording for smiling, laughing, excited body movements with or without vocalizations</td>
<td>Comparing two conditions (a) Reinforcing adaptive responses with or without challenging behavior; &amp; (b) Reinforcing adaptive responses without challenging behavior</td>
<td>Indices of happiness were higher during intervention B (M = 83.3%; range = 59.96%) than both intervention A (M = 78.3%; range = 48-92%) &amp; baseline (M = 26.7%; range = 5-63%) for all participants PND: Highly effective; MC: Maintained during 1-mo follow-up; G: NR</td>
<td></td>
</tr>
<tr>
<td>Vernon et al. (2012)</td>
<td>N = 3 (all males); 2-4 yo (M = 3:2); all with autism</td>
<td>MB across participants</td>
<td><strong>Child positive affect</strong> – frequency &amp; duration of visible &amp;/or audible indications of happiness &amp; enjoyment (e.g., smiling, laughing)</td>
<td>Embedding social interaction into parent-delivered PRT</td>
<td>Child affect was higher during intervention (M = 18.7 occurrences; M = 150.5s) than baseline (M = 3.1 occurrences; M = 8.5s) for all participants PND: Highly effective; MC: Maintained during 1-wk follow-up; G: Generalized to non-training sessions</td>
<td></td>
</tr>
<tr>
<td>Vismara &amp; Lyons (2007)</td>
<td>N = 3 (all males); 2-3 yo (M = 2:8); all with autism</td>
<td>Phase reversal with embedded ATD</td>
<td><strong>Child affect</strong> - affect rating obtained using the GCARS (consisting of one Likert scale)</td>
<td>Comparing two conditions (a) PI stimuli: parent implementation of PRT with PI stimuli; &amp; (b) NP stimuli: parent implementation of PRT with NP stimuli</td>
<td>Child affect was higher during the PI condition (M = 3.8; range = 1-5) than the NP condition (M = 2.9; range = 1-4) &amp; baseline (M = 2.6; range = 1-4) for all participants PND: Ineffective; MC: NR; G: NR</td>
<td></td>
</tr>
</tbody>
</table>
Neutral affect: uninterested in activity, may smile or frown occasionally, overall seems neutral, may fidget, inattentive, generally compliant but not readily (2-3); Positive affect: readily attends, enjoying self, smiles, laughs, or shows other positive emotional behavior, alert & involved, compliant (4-5)

SMRS = Scientific Merit Rating Scale; ATD = alternating treatments design; ABA = applied behavior analysis; DTT = discrete trial training; PND = percentage of non-overlapping data; MC = maintenance check; NR = not reported; G = generalization; MB = multiple baseline; GCARS = General Child Affect Rating Scales; PI = perseverative interest; MSRIQ = Mood Scale & Related Interview Questions; SD = discriminative stimulus; MTS = momentary time sampling; MP = multiple probe; FCT = functional communication training; PRT = pivotal response treatment; TD = typically developing; RIRD = response interruption & redirection; NP = non-perseverative interest