American Journal on Intellectual and Developmental Disabilities

Direct Assessments to Evaluate Psychotropic Medication Effects for Children with Disabilities

--Manuscript Draft--

Manuscript Number:	AJIDD-D-20-00104R2
Article Type:	Comprehensive Review
Keywords:	psychotropic medication; behavioral assessment; direct observation
Corresponding Author:	Blair Lloyd Vanderbilt University Nashville, TN UNITED STATES
First Author:	Jessica Torelli
Order of Authors:	Jessica Torelli
	Blair Lloyd
	Marney Pollack
Manuscript Region of Origin:	UNITED STATES
Abstract:	To evaluate effects of psychotropic medication for children with disabilities, direct assessments may offer a valuable supplement to caregiver reports. Relative to indirect assessment, direct measures of behavior can increase objectivity and sensitivity, and some have potential to isolate distinct behavioral and learning processes. We conducted a systematic, narrative literature review to identify and describe the types and qualities of direct assessment methods that have been used to evaluate effects of non-stimulant psychotropic medication for children with disabilities. We identified 50 studies and 78 direct assessments, which we organized and described using seven assessment categories. Only one study met all three direct assessment quality indicators. We use our descriptive results to highlight research trends and gaps that warrant further study.

Abstract

To evaluate effects of psychotropic medication for children with disabilities, direct assessments may offer a valuable supplement to caregiver reports. Relative to indirect assessment, direct measures of behavior can increase objectivity and sensitivity, and some have potential to isolate distinct behavioral and learning processes. We conducted a systematic, narrative literature review to identify and describe the types and qualities of direct assessment methods that have been used to evaluate effects of non-stimulant psychotropic medication for children with disabilities. We identified 50 studies and 78 direct assessments, which we organized and described using seven assessment categories. Only one study met all three direct assessment quality indicators. We use our descriptive results to highlight research trends and gaps that warrant further study.

Keywords: psychotropic medication, behavioral assessment, direct observation

A Systematic Review of Direct Assessments to Evaluate Psychotropic Medication Effects for Children with Disabilities

The prevalence of psychotropic medication to treat emotional or behavioral symptoms of psychiatric and behavioral disorders has increased for children and youth over the last several decades (Carlson, 2019; Olfson et al., 2015). Relative to typically developing children, children with disabilities are more likely to experience trauma, engage in challenging behavior, and be diagnosed with psychiatric conditions, each of which increases the likelihood of psychopharmacological treatment (McLaren et al., 2018). In fact, psychopharmacological treatment is particularly common among children with disabilities, with recent prevalence estimates exceeding 50% across multiple disability populations (e.g., Mattison et al., 2014; Scheifes et al., 2013; Spencer et al., 2013).

There are no doubt children who need and benefit from medication-based therapies. However, there are also a number of concerns associated with widespread use of psychotropic medication. First, levels of evidence supporting medication efficacy vary widely by medication class and population. For example, robust evidence supports the use of stimulant medications to treat core symptoms of Attention-Deficit/Hyperactivity Disorder (ADHD) for children and adolescents (Greenhill et al., 2002; Pliszka et al., 2007). However, other medication classes are commonly prescribed off-label (i.e., to treat symptoms or conditions other than those for which a medication was formally tested and approved). For example, antipsychotics and alpha-agonists are two medication classes commonly prescribed off-label to treat a range of behavioral problems for children with developmental disabilities (Vitiello, 2017; Zito et al., 2008). In addition to off-label prescriptions, polypharmacy (i.e., the simultaneous use of multiple medications) and long-term use of medications raise concerns around the safety and efficacy of

psychopharmacological treatment for children. This is because the vast majority of medication efficacy trials inform short-term effects of single medications (McVoy & Findling, 2017). All three concerning trends have been identified disproportionately among vulnerable child populations, including children with disabilities (McLaren & Lichtenstein, 2019).

For all children prescribed psychotropic medication, including the most vulnerable groups among them, it is critical to understand whether and how these medications produce therapeutic effects. In practice, prescribing physicians commonly rely on informal reports by children and their caregivers. Because many children with disabilities have limited communication repertoires, physicians must often rely solely on caregiver reports and their own global clinical impressions to make treatment decisions (Vitiello, 2017; Zarcone et al., 2008). Even in controlled clinical trials of medication efficacy, researchers almost exclusively rely on standardized rating scales completed by caregivers to evaluate therapeutic outcomes (Aman et al., 2004; Li et al., 2017; Parikh et al., 2008; Vitiello, 2017). While such indirect measures are time- and resource-efficient, they are also prone to bias, particularly when raters are not blind to treatment condition (Higgins et al., 2011; Waschbusch et al., 2009). Caregivers have also shown a tendency to attribute therapeutic gains to medication treatment even when other types of interventions (e.g., behavioral) are also shown to produce positive changes in behavior (Waschbusch et al., 2009). For these reasons, indirect measures by caregivers alone are unlikely to offer complete answers on whether and how medication produces a therapeutic effect.

Relative to indirect measures, direct measures of behavior require more time and resources, and are therefore less commonly applied for purposes of medication evaluation. But there are several reasons why direct assessments may offer a valuable supplement to caregiver ratings and reports. First, direct assessments often provide more objective accounts of behavior

than third-party or self-reports, particularly when those reports are influenced by reporter characteristics (e.g., parent expectations, child cognitive impairments; Zarcone et al., 2008; Yoder et al., 2018). While observer bias is still possible and important to guard against, it is less likely when planned coding protocols are followed and data from independent observers (ideally those blind to condition or study purpose) are regularly checked for reliability. Second, systematic direct observation using count-coding methods can provide more sensitive measures of behavior relative to indirect methods or rating scales (Yoder et al., 2018). Increased sensitivity is advantageous when the goal is to detect the presence of therapeutic effects.

Third, the types of information direct assessments offer are distinct from those of indirect methods. Indirect assessments are commonly used to measure global characteristics or general patterns of behavior that are stable across contexts. In contrast, systematic direct observation is uniquely suited to measure behavior as it occurs in specific social contexts or other defined environmental conditions (Cairns & Green, 1979). When observational contexts are highly structured or controlled, direct assessments can help isolate how one or more environmental variable influences behavior. As an example, a functional analysis involves a highly controlled set of conditions designed to identify what function problem behavior serves in a given context (Iwata et al., 1994; Hanley, 2012). As applied to medication evaluation, a functional analysis could inform not only whether levels of problem behavior changed after starting a new medication, but whether the environmental triggers or functions of those behaviors also changed (Cox & Virues-Ortega, 2016; Schaal & Hackenberg, 1994; Thompson et al., 2007). Thus, structured direct assessments could help identify whether and how medication interacts with features of a child's environment to produce a therapeutic outcome.

Fourth, direct measures of behavior are commonly used to monitor effects of behavioral

interventions for children with disabilities and challenging behavior. This is relevant given broad consensus among experts in child psychiatry that behavioral or other psychosocial interventions should be the first line of treatment to address challenging behavior before prescribing psychotropic medication (with stimulants as a notable exception; American Academy of Child and Adolescent Psychiatry [AACAP], 2012; American Academy of Pediatrics, 2011; Gleason et al., 2007; McLaren & Lichtenstein, 2019; Rush & Frances, 2000). Indeed, in a recent survey of clinicians who prescribed medication to children with autism, Rieken et al. (2019) found clinicians identified (a) whether the child was receiving psychosocial/behavioral treatment and (b) child progress in psychosocial/behavioral treatment among their most important data sources when monitoring effects of psychotropic medication. These findings also reflect current guidelines in psychiatry that support continuation of behavioral treatment after initiating a medication regimen, as psychotropic medications are expected to improve a child's response to other forms of psychosocial intervention (AACAP, 2012). Such guidance suggests the same direct observation methods commonly used to evaluate behavioral interventions could be leveraged to help determine whether the addition of medication produces the intended therapeutic effect (Lloyd et al., 2016).

A series of methodological recommendations have been made for studies evaluating effects of psychotropic medication (e.g., Courtemanche et al., 2011; Napolitano et al., 1999; Van Haaren & Weeden, 2013; Zarcone et al., 2008), but three features are especially relevant for those incorporating direct assessment methods. First, inter-observer agreement data should be collected and reported to inform the reliability of observational variables. Depending on research methodology, indices such as percentages of agreement, kappa, or intra-class correlation coefficients inform the consistency of scores across independent observers (Yoder et al., 2018).

Second, to minimize the likelihood of systematic bias, the observers who collect behavioral data should be blind to treatment condition (i.e., medication status; Higgins et al., 2011; Reichow et al., 2018). While direct observation data are less prone to bias than third-party or self-reports, studies in which observers are not blind to condition can lead to systematic over- or underestimation of behaviors measured (Ledford et al., 2018; Yoder et al., 2018). Third, for assessment conditions that are highly structured, procedural fidelity data should be collected to ensure assessment procedures were implemented as planned. While fidelity measures have historically focused on intervention implementation, fidelity data are just as critical in the context of assessment (Reed & Codding, 2014). Such measures help rule out alternative explanations for why behavior patterns might have changed from one observation or experimental condition to another. Meeting these standards is necessary to inform the potential added value of direct assessment approaches to medication evaluation.

We know of only two previous literature reviews that have focused on direct assessment of psychotropic medications. One focused on broad methodological features of studies published in a single journal, most of which evaluated stimulant medications for children and adults with ADHD (Van Haaren & Weeden, 2013). The other focused on synthesizing results of a single assessment method (i.e., functional analysis) across studies (Cox & Virues-Ortega, 2016). No review has yet summarized the types and qualities of direct assessment methods that have been used to evaluate psychotropic medication effects, particularly outside the stimulant literature. This was the focus of the current review. We aimed to address the following research questions:

1. For what child disability populations and treatment purposes have non-stimulant psychotropic medication been evaluated via direct assessments?
2. What types of direct assessments and behavioral measures have been used to evaluate effects

of non-stimulant psychotropic medication?

3. To what extent did studies meet each of three quality indicators specific to direct assessment (i.e., reliability, observer blinding, procedural fidelity)?

Method

Inclusion and Exclusion Criteria

To be included in this review, studies were required to (a) include a direct measure of behavior to evaluate effects of a psychotropic medication (other than stimulants) and (b) include at least one participant age 17 or younger who had a disability. We defined direct measures of behavior as those in which researchers used systematic direct observation or automated devices to record one or more behaviors as they occurred throughout an observation. Expert rating scales or checklists completed following direct observations were excluded. Psychotropic medications were defined as any medication prescribed to alter mood or behavior (National Institute of Mental Health, 2016). We excluded studies evaluating effects of stimulant medications only, given there already exists a substantial evidence base establishing effects of these medications, along with evidence-based practice parameters (Greenhill et al., 2002; Pliszka et al., 2007). We included studies that evaluated effects of non-stimulant psychotropic medications in addition to stimulants. Because the focus of this review was on the types and qualities of direct assessment methods, we did not exclude studies on the basis of experimental design rigor. That is, studies were included as long as they incorporated an active manipulation of psychotropic medication or comparison of conditions, one or more of which involved psychotropic medication treatment.

Search and Screening Procedures

To identify studies, we searched two electronic databases (i.e., PsycINFO and PubMed) and conducted ancestral searches for all studies meeting inclusion criteria and for several related

literature reviews. We used an iterative process to develop a list of search terms. We conducted initial searches using highly relevant terms (e.g., psychotropic medication, direct observation), then identified additional relevant search terms from key words of identified studies. As we conducted these searches, we also checked to ensure that results included studies we were already familiar with that met inclusion criteria. If they did not, we further expanded search terms using key words from those studies. Final search terms incorporated those related to medication (*medication*, *psychotrop**, *psychopharm**, *drug therapy*, *central nervous system agent*) and direct assessment (*direct observation**, *behavior**, *reinforce**, *preference assessment*, *functional analysis*, *symptom**, *behavior* mechanism**, *social behavior*, *behavior therapy*). Exact search terms varied by database to account for distinct indexing methods (i.e., PubMed's Medical Subject Headings [MeSH]). We limited results to peer-reviewed studies written in English and conducted with human participants. The search was last updated on May 29, 2020. Search procedures resulted in 3,509 studies excluding duplicates.

The first author and four research assistants screened studies for inclusion in two phases. First, we screened titles and abstracts only and excluded any studies that (a) were not empirical, (b) were systematic literature reviews or meta-analyses, (c) did not include human participants, (d) did not include participants 17 years or younger, (e) focused on drug abuse or addiction, (f) did not include a direct measure of behavior, or (g) did not include a psychotropic medication. If titles and abstracts did not provide sufficient information to make these determinations, we kept them for further screening. We screened out 3,273 articles in the initial screening.

Second, we screened full texts of articles to identify studies that included (a) a non-stimulant psychotropic medication; (b) an active manipulation or comparison of conditions related to a psychotropic medication; and (c) a direct measure of behavior. We screened in 26

studies during full-text screening. We then completed ancestral searches for all studies meeting inclusion criteria, during which we identified 15 additional studies. We also completed ancestral searches for five related reviews and conceptual papers (Cox & Virues-Ortega, 2016; Li et al., 2017; Napolitano et al., 1999; Schroeder et al., 1983; van Haaren & Weeden, 2013), from which we identified seven additional studies. Finally, in the process of conducting background research for this review, we identified two additional studies. In total, 50 studies met inclusion criteria. A PRISMA flow diagram (Stovold et al., 2014) of study inclusion is presented in Figure 1.

We assessed interrater agreement for both stages of study screening. Two graduate research assistants independently screened 15% of the initial study sample (518 studies) and agreed on 94.8% of studies for whether to exclude or keep for further screening. For full-text screening, two research assistants screened 18% of the study sample (43 studies) and agreed on 91.0% of studies on inclusion or exclusion. Throughout screening, raters met to discuss any articles with disagreements until reaching consensus on the screening decision.

Coding Procedures

We developed a coding manual to describe critical features of each study and address the research questions. Two doctoral students (first and third authors) and one faculty member in Special Education (second author) completed all coding procedures. To address the first research question, we coded participant age, gender, race or ethnicity, disability types, and psychiatric diagnoses. We coded these participant variables at the study level for studies using group designs and at the participant level for those using single case designs. Coding categories for psychiatric diagnoses were based on the DSM-5 manual (American Psychiatric Association, 2013); disability categories were coded based on author report. We also coded the purpose for which medication was prescribed or evaluated according to three coding categories: treatment of core

symptoms of a diagnosis or disorder; treatment of target behaviors (regardless of diagnosis); and not reported. For studies reporting this information, we noted the diagnoses and/or behavioral topographies indicated for treatment. Finally, we coded the medications that were evaluated (medication classes and generic names). If medication type varied by participant, we coded all medications evaluated.

To address the second research question, we classified direct assessments according to what procedures were in place when child behavior was being measured. We first distinguished between unstructured and structured direct assessments. We defined unstructured assessments as observations that took place during some typical activity or routine with no programmed tasks, antecedents, or consequences of target behaviors. Because we anticipated unstructured assessments would take place in naturalistic settings, we noted setting type for this assessment category. We defined structured assessments as conditions that did involve programmed antecedents or consequences of target behavior or presentation of some explicit task or protocol. We then used an inductive approach to create assessment categories within structured direct assessments. That is, we initially entered narrative descriptions of each assessment. We used these descriptions to generate a set of seven structured assessment categories to further organize and summarize the assessment types. More than one type of assessment could be coded per study. The resulting coding categories are listed and described in the results section.

For each direct assessment, we also coded the behaviors that were directly measured. We initially coded target behaviors based on author descriptions, then used these descriptions to generate a set of nested categories, including problem behavior (e.g., aggression, disruption, self-injury, rule violations), appropriate behavior (e.g., appropriate requests, on-task behavior, compliance, work completion), and motor-related behaviors (e.g., stereotypy, tics, movement).

Other categories of measured behaviors included correct or incorrect responding, reinforcers earned, arbitrary responses (e.g., switch-pressing), attentional behaviors, affect, other social behaviors, and response time. Multiple behavior types could be coded per assessment. To provide context, we also noted the broader study designs in which these direct assessments and behavioral measures were applied to evaluate effects of medication. We coded designs into three main categories according to the medication evaluation: between group comparisons, within subjects comparisons, and single case designs. For studies using single case designs, we noted whether the design allowed sufficient opportunities to demonstrate effects of medication, as many of these studies involved complex designs with more than one treatment component being evaluated. We considered designs with three opportunities to demonstrate the medication effect sufficient (Ganz & Ayres, 2018; Horner et al., 2005).

To address the third research question, we coded the presence or absence of each of three quality indicators related to direct assessment at the level of each study. First, we coded whether authors collected and reported reliability data on any of the behaviors measured via direct assessment. Second, we coded whether observers (i.e., data collectors during direct assessments) were blind to medication condition. Consistent with expert recommendations to assess risk of bias (Higgins et al., 2011; Reichow et al., 2018), we required study authors to report blinding status specific to observers to code this quality indicator as present. However, if authors reported all research staff were blind to condition, and the study was placebo controlled, we considered this sufficient evidence for observer blinding. If a study only included direct assessments that did not require human observers (i.e., used an automated recording device), reliability and blinding variables were coded as not applicable. Third, for structured direct assessments only, we coded whether authors collected and reported data on the fidelity of assessment implementation.

We assessed interrater agreement for all coded variables for 10 or more randomly-selected studies (at least 20% of study sample). We calculated mean percentages of agreement (i.e., number of agreements divided by sum of agreements and disagreements, multiplied by 100) by coded variable. Coders discussed all disagreements and referenced the study and coding manual to reach consensus. We changed primary data to the consensus code, but retained original disagreements in calculations of interrater agreement. Mean agreement was 94.4% across all coded variables, including participant characteristics and medication treatment codes (M = 93.9%; range, 86.7%–100%), assessment, target behavior, and design codes (M = 92.1%; range, 82.6%–100%); and assessment quality indicator codes (M = 98.5%; range, 92.3%–100%).

Results

Fifty published studies met inclusion criteria. These studies were published across 30 journals between 1963 and 2019. Most identified studies were published between 1970 and 1994; only 14 were published since 2000.

Child Populations, Medications, and Purpose of Medication Treatment

The mean participant age was 11.3 (SD = 5.4) among 22 studies using group designs and 10.4 (SD = 3.5) among 28 studies using single case designs. Race or ethnicity was reported in only seven studies. Among three group design studies reporting race or ethnicity, the majority of participants were White (76%–80%). Among four single case studies, race or ethnicities reported were White (n = 3), Black (n = 1), Asian (n = 2), and Somali (n = 1). For both studies using group designs and those using single case designs, most participants were male (80% and 70%, respectively) and had at least one psychiatric diagnosis (100% and 92%, respectively). Across both sets of studies, the most common disability categories represented were intellectual disability (59% and 82%, respectively) and autism spectrum disorder (41% and 23%,

respectively; see Table 1 for a summary of disabilities and disorders represented across studies).

The most common medication classes evaluated in the study sample were typical (first generation) antipsychotics (30 studies; 60%) and atypical antipsychotics (10 studies; 20%). Evaluations of typical antipsychotics were largely replaced by atypical antipsychotics starting around the year 2000. Other medication classes evaluated included anticonvulsants (5 studies; 10%), opiate antagonists (5 studies; 10%), alpha 2 agonists (3 studies; 6%), selective norepinephrine reuptake inhibitors (SNRIs; 2 studies), mood stabilizers (2 studies), and selective serotonin reuptake inhibitors (SSRIs; 1 study). Stimulants were evaluated in combination with non-stimulant medications in nine studies (18%). Medication classes are listed by study in Table 2; a complete list of medications evaluated is provided in Table 3. With respect to the purpose of medication treatment or evaluation, medications were described as treating targeted behaviors (e.g., self-injury, stereotypy, aggression, property destruction, tics) in 32 studies (64% of sample) and core symptoms of diagnoses (i.e., Autism, Psychosis, Tourette's syndrome, ADHD) in 10 studies (20% of sample). The purpose of medication treatment or evaluation was unclear or not reported in remaining studies.

Types of Direct Assessments and Target Behaviors Measured

Among the 50 studies, we identified a total of 78 direct assessments. A mean of 1.5 (range, 1–6) assessments were included per study. The majority of assessments (82.3%) were categorized as structured; the rest (17.7%) were unstructured. Research designs in which direct assessments were incorporated included between group designs (n = 8), within subjects designs (n = 13), single case designs (n = 28), or some combination (n = 1). Of 28 single case designs, only six provided sufficient opportunities to demonstrate medication effects (see Table 2).

Unstructured Assessments

Fourteen assessments (from each of 14 studies) were unstructured; that is, they did not involve any programmed tasks, nor did they include programmed antecedents or consequences of targeted behaviors. Broadly, these assessments were designed to directly evaluate changes in behaviors as they occurred in typical activities or routines following changes in medication conditions. Unstructured assessments were completed in residential treatment facilities or inpatient hospitals (e.g., on living units; *n* = 9), educational settings (*n* = 2), or a combination of school and home environments (*n* = 1). Assessment settings were not reported for two studies. During 10 of the 14 unstructured assessments, researchers collected data on some topography of problem behavior (e.g., self-injury, aggression, property destruction, rule violations). Data were also collected on stereotypy (*n* = 6), appropriate or on-task behavior (*n* = 6), movement or activity level (*n* = 5), and social behaviors (*n* = 4).

Functional Analyses

Functional analyses were used in eight assessments across eight studies (all of which were published since 2000). Functional analyses involved conditions designed to evoke and/or reinforce problem behavior (Hanley, 2012), and were therefore used to isolate the evocative or reinforcing properties of defined stimulus conditions. Six of these assessments systematically programmed both antecedents and consequences of target behavior, thus focusing on identifying contingencies that reinforced problem behavior. All six included standard play (i.e., control) conditions and test conditions designed to evaluate whether problem behavior was maintained by (a) negative reinforcement in the form of escaping non-preferred tasks, and/or (b) positive reinforcement in the form of accessing adult attention. Four studies also tested whether problem behavior was maintained by positive reinforcement in the form of accessing preferred items or activities. In one study, Zarcone et al. (2004) included additional individualized test conditions.

A social avoidance condition tested whether problem behavior was reinforced by avoiding social contact; a staff removal condition tested whether the presence or absence of certain individuals influenced levels of problem behavior.

The other two studies systematically programmed antecedents of target behavior only. These studies used a Structured Observational Analog Protocol (SOAP) as an analog measure of parent-child interactions (Handen et al., 2013; Grondhuis et al., 2019). The SOAP consisted of a series of one to three test conditions and one control condition. In the control condition, children had access to preferred toys and parents were asked to play with their child as they typically would at home. In each test condition, parents were coached to present a condition that would potentially evoke problem behavior (i.e., diverting attention, presenting task demands, restricting preferred items) and to respond to child behavior as they typically would. Handen et al. (2013) included three test conditions: presentation of task demands, restricted attention, and restricted tangibles. Grondhuis et al. (2019) included presentation of task demands only.

Among functional analyses, some measure of problem behavior was used in almost all assessments (n = 7), most of which used definitions that included multiple topographies of problem behavior. The remaining study collected data on the occurrence of tics (Anderson et al., 2002). The other type of behavior measured in these assessments was compliance or work completion (n = 3).

Behavioral Treatment Assessments

Nineteen structured assessments (from 17 studies) were defined by the presence or absence of a behavioral treatment. While authors of these studies did not frame their behavioral treatments as a type of assessment, they met our conceptualization for direct assessments because the behavioral treatments defined the procedures that were (or were not) in place when behavior

was assessed. Most behavioral treatments (n = 14) included reinforcement contingencies for alternative behaviors or behaviors incompatible with targeted problem behavior. Examples of reinforcement-based intervention components were differential reinforcement of alternative behavior, differential reinforcement of other behavior, and token economies. Punishment contingencies were just as common (n = 14). Examples of punishment procedures included timeout, response costs, overcorrection procedures, restraint or immobilization, and visual screening. Twelve assessments included both reinforcement and punishment components. Other behavioral treatment conditions included noncontingent reinforcement, response interruption and redirection, compliance training, and daily report cards. Among assessments defined by the presence or absence of behavioral treatment, problem behaviors (e.g., self-injury, aggression, disruption, rule violations) were the most common types of behavior measured (n = 14; 73.7%). Data were also collected on appropriate behaviors (e.g., on-task behavior, work completion; n = 5) and stereotypy (n = 4).

Discrimination Tasks

Discrimination tasks were used in 11 assessments across six studies; one study by Anderson et al. (1989) included five different discrimination tasks. Stimulus discrimination tasks involved the presentation or alternation of stimuli that signaled the availability or unavailability of reinforcement. Broadly, these assessments were designed to evaluate whether medication impacted stimulus control, or the extent to which a child's behavior would vary by stimulus presentation. This set of assessments included both simple and conditional discrimination tasks. Simple discrimination tasks typically involved reinforcement of an arbitrary response in the presence of some visual stimulus, but not in its absence. For example, one assessment by Anderson et al. (1989) involved reinforcing lever presses when a 10-dot visual pattern was

presented, but not when the screen was blank. Conditional stimulus discrimination tasks included matching-to-sample tasks (e.g., Aman et al., 1991) or others in which a 'correct' (reinforced) response was conditional upon another stimulus. The most common type of behavior measured in these assessments was correct or incorrect responses (n = 10; 90.9%), with accuracy defined according to whether responses occurred in the presence or absence of the correct (i.e., discriminative) stimulus. Other types of behaviors measured were movement-related behaviors (e.g., seat movement [wiggling], activity levels; n = 6), stereotypy (n = 5), response time (n = 3), arbitrary responses without regard to accuracy (e.g., bar pressing, moving a ball manipulandum; n = 2), and reinforcers earned (n = 1). No measures of problem behavior were included in any of the discrimination tasks.

Memory or Attention Tasks

Direct assessments evaluating performance on specific memory or attentional tasks were included in 14 assessments across seven studies. These assessments were used to evaluate a medication's effects on a child's short-term memory, learning retention, or selective attention. While these assessments would likely be considered cognitive-behavioral assessments, they all met our definition for including direct measures of behavior throughout the assessment or learning task. Five studies included Continuous Performance Tasks (CPT), which were designed to measure sustained and selective attention; e.g., Aman et al., 1991; Sallee et al., 1994; Snyder et al., 2002). These tasks involved rapid presentations of two or more different visual stimuli in random sequence. Children were instructed to press a lever or space bar only when a specific stimulus appeared. Other assessments included short term memory tasks (e.g., Aman et al., 1991; Sprague et al., 1970), serial learning tasks (e.g., Helper et al., 1963), and digit span tasks (e.g., Helper et al., 1963). Among these assessments, the most common behaviors measured were

correct or incorrect responses (*n* = 13; 92.9%). Other types of behaviors measured were response time (*n* = 6) and movement (i.e., seat movement [wiggling] while completing assessments; *n* = 5). No measures of problem behavior were included in any of the memory or attention assessments.

Other Instructional or Prosocial Tasks

Other instructional or prosocial task assessments were those in which some instructional or developmental task was presented that did not meet criteria for any of the other assessment categories (i.e., was not designed to evoke problem behavior; did not explicitly address stimulus control, attention, or memory). These conditions were used in five assessments across five studies. Other instructional or prosocial tasks included pre-academic tasks (e.g., printing letters, doing puzzles, stacking blocks), academic tasks (e.g., addition fact sheets), responding to experimenter questions, and following a set of simple instructions (e.g., walk to the door). Behaviors measured during these assessments were correct responding (*n* = 3), compliance or work completion (*n* = 3), on-task behavior (*n* = 2), attending behavior (*n* = 1), and affect (i.e., facial grimaces; *n* = 1).

Motor Tasks

Motor task assessments were those that included a task designed to evaluate performance of specific motor skills. Motor tasks were used in only three assessments across two studies. Aman et al. (1991) conducted a graduated holes task that was based on the Motor Steadiness Battery (Klove, 1963). This task involved holding a stylus steady inside a series of holes that varied in size, with the goal of not touching the sides. Helper et al. (1963) included two motor tasks. In one task, children were asked to tap a stylus as rapidly as possible on a metal plate for three 15-s periods. In the other task, children were asked to place a dot in the center of as many

circles as possible during a 1-min period. Helper et al. (1963) stated these assessments were designed to "detect drug effects upon such nonlearning factors as cooperation and coordination" (p. 3). Behaviors measured in motor assessments were incorrect responding (i.e., errors defined as contact between the stylus and metal plate; Aman et al., 1991) and other behaviors specific to each motor task (i.e., stylus tapping; dotting circles; Helper et al., 1963).

Other Structured Assessments

The remaining four assessments (across 3 studies) were structured but did not fit any of the above categories. Barrett et al. (1989) collected data on self-injurious behavior during observations that consisted of three activities: an outdoor walk, a snack period, and an opportunity to engage with play materials. Aman et al. (1984) collected data on stereotypy (body rocking) and lever pulls during a condition in which lever pulls were reinforced with edibles. They included both measures to evaluate whether chlorpromazine differentially impacted stereotypy and other learned responses (i.e., decreased stereotypy while increasing the learned behavior of level pulls). Finally, Sand and Carlson (1973) collected data on the occurrence of tics during a 5-min segment of an interview in which the child was asked to discuss his activities over the past few days.

Direct Assessment Quality Indicators

The extent to which studies met quality indicators related to direct assessments varied by indicator. Researchers collected and reported reliability data on at least one behavioral outcome measure in 32 of 45 studies using human observers (71.1%). Of the 45 studies using human observers, 24 (53.3%) reported blinding observers to study purpose or medication condition. And of the 44 studies that included one or more structured assessment, only two (4.5%) collected and reported procedural fidelity data on assessment procedures. Quality indicator status is presented

by study in Table 2. Notably, only one study met all three quality indicators (Singh et al., 1993); 19 studies met both reliability and blinding quality indicators.

Discussion

A number of concerning trends in pediatric pharmacological intervention (i.e., off-label prescribing, polypharmacy, long-term use) highlight a need to understand whether and how these medications produce a therapeutic effect. This need is especially urgent for children with disabilities who are disproportionately impacted and most vulnerable to a host of negative outcomes (McLaren et al., 2018). Direct assessments of behavior have potential to offer objective and sensitive measures of therapeutic effects to supplement commonly relied-upon caregiver ratings and reports. In fact, prescribing clinicians have rated direct assessment data as being important sources of information to draw from when monitoring medication effects for children with disabilities (Rieken et al., 2019). Direct assessments may be especially useful in isolating effects on behavioral and learning processes, or determining whether medication improves a child's responsiveness to other non-medical interventions. We conducted a literature review to summarize and describe the range of direct assessments used to evaluate effects of non-stimulant psychotropic medication among children with disabilities. Our goal was to highlight assessment methods that researchers and clinicians might consider incorporating in medication evaluation, and to identify what next steps are needed to better understand what direct assessments can contribute to medication monitoring.

Our first research question focused on child populations, medications, and purposes for which medications were evaluated via direct assessment, and our results revealed a few distinct trends. First, outside the stimulant literature, studies that have incorporated direct assessments of behavior largely included children and youth with intellectual and developmental disabilities and

most commonly evaluated effects of typical antipsychotics—a class of medication that has largely been replaced by atypical antipsychotics due to their associated extrapyramidal side effects (Crystal et al., 2009; Harrison et al., 2012; Meltzer, 2004). Second, for more than half (64%) of the studies, authors identified the purpose of medication treatment as addressing specific target behaviors, as opposed to core symptoms of diagnosed disabilities or psychiatric disorders. Coupled with the number of studies using single case designs (n = 28), this pattern may reflect researcher expectations that direct assessments are better suited to measure context-dependent behaviors than global characteristics (Yoder et al., 2018).

Our second research question, and the primary focus of this review, addressed the types of direct assessments and behavioral measures that have been used to better understand medication effects. We were particularly interested in structured assessments that might inform how medication impacts basic behavioral or learning processes, which could then inform medication's impact on the effectiveness of other non-medical interventions. While the majority of assessments met our definition of structured, we identified four assessment categories as having unique potential to isolate behavioral or learning processes. First, functional analyses were used to determine whether the motivating conditions or reinforcement contingencies maintaining problem behavior changed with medication. Second, a subset of behavioral treatment assessments provided opportunities to evaluate how the addition of a medication impacted the effectiveness of programmed reinforcement and/or punishment contingencies. Third, discrimination tasks were used to evaluate whether medication influenced stimulus control, or the degree to which a child would respond differently in the presence of different stimuli. And a fourth assessment category targeted specific learning processes related to memory and attention (e.g., short-term memory, selective attention, sustained attention). Notably, none of

the identified assessments directly informed changes in child preference. Behavioral researchers have speculated that psychotropic medications might impact the value of certain reinforcers—an effect that would have direct implications for behavioral treatment (Roane, 2008; Carlson et al., 2012). Indeed, some researchers have found evidence of such effects for stimulant medications (e.g., LaRue et al., 2008; Northup et al., 1997). Future studies might incorporate preference or reinforcer assessments to explore non-stimulant medications' impact on reinforcer value.

We were also interested in the extent to which assessments involved data collection on adaptive and prosocial behaviors in addition to focusing on reducing maladaptive behaviors (Valdovinos et al., 2009). Behaviors measured varied by assessment category. Measures of challenging behavior were most common among studies using functional analyses and behavioral treatment assessments, yet were largely absent from other assessment categories. Among assessments of stimulus discrimination, memory, or attention, measures of response accuracy (specific to the designated task) were most common. Measures of work completion, compliance, and on-task behavior were less common overall, and only appeared in a subset of studies using functional analysis, behavioral treatment assessments, and those assessing performance on other instructional or prosocial tasks. Other behaviors measured were neither maladaptive nor adaptive, and included stereotypy, activity levels (e.g., movement, wiggling), or task-specific fine motor behaviors. Measures of adaptive social behaviors, such as positive social interactions with adults or peers, were largely absent (see Grondhuis et al., 2019 and Handen et al., 2013 for exceptions). We also noted that while several studies using unstructured assessments and behavioral treatment assessments included direct measures of both maladaptive and adaptive behaviors, none included a battery of direct assessments meant to inform both reductive and skill-based effects of medication. These results suggest a need to directly assess

whether and how a medication impacts aspects of both child behavior and learning processes. Such comprehensive evaluations could help researchers and clinicians distinguish therapeutic effects of medication from common side effects (e.g., sedation).

Our third research question focused on the quality of direct assessments, as measured by the presence of three indicators: reliability of observational variables, observer blinding, and procedural fidelity of assessment procedures. Notably, only one of the 50 identified studies met all three indicators. Roughly half of the studies used blind observers for data collection, and almost none of those including structured assessments collected or reported fidelity data. While the lack of fidelity data might be expected given the age of the study sample, it nevertheless represents a threat to the internal validity of these studies (Reed & Codding, 2014). Taken together, these results highlight a need for additional research using direct assessments that meet these basic standards.

Another critical aspect of direct assessments is the validity of behavioral measures they produce. While we did not attempt to systematically evaluate assessment validity in this review, we did note general trends suggesting a need to validate observational measures for purposes of medication evaluation. Outside of the assessments designed based on individualized behavioral treatments, few studies reported validity data to support their assessment selection or cited standard assessment protocols on which they based their procedures. Assessments that were based on standard protocols included the functional analysis (e.g., Zarcone et al., 2004), the SOAP (Grondhuis et al., 2019; Handen et al., 2013), CPT (e.g., Snyder et al., 2002), the Student Behavior Teacher Response Observation Code (Waxmonsky et al., 2010), and the Motor Steadiness Battery (Aman et al., 1991). Yet studies using these assessments were outnumbered by those lacking explicit rationales for their assessment selections and procedures. This trend,

coupled with the variability in assessment procedures used to measure similar learning processes (e.g., stimulus discrimination), suggests that the variable outcomes common to this literature might be explained by variation in the scientific utility of assessments and measures used. One promising avenue for assessing construct validity of observational variables is to explore their associations with scores from other related assessments. Several studies in this review included a combination of indirect and direct observation measures, yet few directly evaluated correlations between them. In one noteworthy exception, Handen and colleagues (2013) evaluated effects of risperidone (with and without parent training) using a structured observational assessment (SOAP) and two standardized rating scales (Home Situations Questionnaire and Aberrant Behavior Checklist). They found that while therapeutic changes were detected across measures, changes in direct observation measures were not correlated with those from behavioral ratings. They interpreted this finding to suggest each type of measure assessed distinct aspects of therapeutic change. Incorporating multiple behavior change measures in medication evaluations will provide further opportunities to validate behavioral measures from direct assessments and determine the extent to which they offer complementary information to more traditional behavioral ratings.

Results of our review should be considered in light of several limitations. First, the 50 studies identified from our search procedures represent the published literature only. We did not search unpublished literatures and as a result, likely missed identifying other direct assessments that have been used for medication evaluation purposes but did not make it through the peer review process. Likewise, had we broadened our inclusion criteria related to age and medication class, we would have identified a broader range of direct assessments. Second, we did not attempt to summarize or quantify the effects of medication evaluations represented in this

sample. This decision was largely based on the heterogeneity of the study sample with respect to research methodology, independent variables (e.g., medication classes, types, and dosages; whether medication was being evaluated alone or in combination with other treatment components), treatment purposes, and dependent variables. Finally, because we did not limit our search by publication date, many studies included in this review evaluated effects of medications that are no longer commonly prescribed among children (e.g., typical antipsychotics). While prescription patterns have changed over time, the therapeutic goals of psychotropic medication treatment (e.g., addressing severe or chronic patterns of maladaptive behavior, ameliorating core symptoms of psychiatric disorders) have not. Thus, we consider the assessments designed to evaluate these outcomes relevant to the current psychopharmacological treatment landscape.

Despite these limitations, and the descriptive nature of this review, our results provide context and clarity around a call for future research. Publication patterns suggest direct assessments have, at least to some degree, fallen out of favor in medication evaluations. Yet our field still lacks an understanding of how non-stimulant medications impact learning and behavioral processes that lead to therapeutic outcomes. As others have argued before us (e.g., Thompson et al., 2007), using structured direct assessments designed to measure specific behavioral or learning processes has potential to meaningfully contribute to this understanding. But to realize their potential, we need to incorporate high quality direct assessments in rigorous medication evaluations. Results of such studies would inform conditions in which psychotropic medication increase responsiveness to other behavioral or psychosocial intervention.

References

*Aman, M. G., Marks, R. E., Turbott, S. H., Wilsher, C. P., & Merry, S. N. (1991). Methylphenidate and thioridazine in the treatment of intellectually subaverage children: Effects on cognitive-motor performance. *Journal of the American Academy of Child & Adolescent Psychiatry*, *30*(5), 816–824. https://doi.org/10.1016/S0890-8567(10)80022-X

Aman, M. G., Novotny, S., Samango-Sprouse, C., Lecavalier, E., Gadow, K., King, B. H., Pearson, D. A., Gernsbacher, M. A., & Chez, M. (2004). Outcome measures for clinical drug trials in autism. *CNS Spectrums, 9*(1)*,* 36–47. https://doi.org/d2fd

*Aman, M. G., Teehan, C. J., White, A. J., Turbott, S. H., & Vaithianathan, C. (1989). Haloperidol treatment with chronically medicated residents: Dose effects on clinical behavior and reinforcement contingencies. *American Journal on Mental Retardation*, *93*(4), 452–460.

*Aman, M. G., White, A. J., & Field, C. J. (1984). Chlorpromazine effects on stereotypic and conditioned behaviour of severely retarded patients: A pilot study. *Journal of Mental Deficiency Research*, *28*(4), 253–260. https://doi.org/fksztv

American Academy of Child and Adolescent Psychiatry. (2012, February). *A guide for community child serving agencies on psychotropic medications for children and adolescents*. Available from http://www.aacap.org/

American Academy of Pediatrics. (2011). ADHD: Clinical practice guideline for the diagnosis, evaluation, and treatment of attention-deficit/hyperactivity disorder in children and adolescents. *Pediatrics*, *128*(5), 1007–1022. https://doi.org/10.1542/peds.2011-2654

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

*Anderson, L. T., Campbell, M., Adams, P., Small, A. M., Perry, R., & Shell, J. (1989). The effects of haloperidol on discrimination learning and behavioral symptoms in autistic children. *Journal of Autism and Developmental Disorders*, *19*(2), 227–239. https://doi.org/10.1007/BF02211843

*Anderson, M. T., Vu, C., Derby, M., Goris, M., & McLaughlin, T. F. (2002). Using functional procedures to monitor medication effects in an outpatient and school setting. *Psychology in the Schools*, *39*(1), 73–76. https://doi.org/10.1002/pits.10006

*Barrett, R. P., Feinstein, C., & Hole, W. T. (1989). Effects of naloxone and naltrexone on self-injury: A double-blind, placebo-controlled analysis. *American Journal on Mental Retardation*, *93*(6), 644–651.

*Barrett, R. P., Payton, J. B., & Burkhart, J. E. (1988). Treatment of self-injury and disruptive behavior with carbamazepine (Tegretol) and behavior therapy. *Journal of the Multihandicapped Person*, *1*(2), 79–91. https://doi.org/10.1007/BF01098249

*Beale, I. L., Smith, O. A., & Webster, D. M. (1993). Effects of chlorpromazine and thioridazine on discrimination learning in children with mental retardation. *Journal of Developmental and Physical Disabilities*, *5*(1), 43–54. https://doi.org/10.1007/BF01046597

*Burgio, L. D., Page, T. J., & Capriotti, R. M. (1985). Clinical behavioral pharmacology: Methods for evaluating medications and contingency management. *Journal of Applied Behavior Analysis*, *18*(1), 45–59. https://doi.org/10.1901/jaba.1985.18-45

Cairns, R. B., & Green, J. A. (1979). How to assess personality and social patterns: Observations or ratings? In R. B. Cairns (Ed.), *Analysis of social interaction*. Erlbaum.

*Campbell, M., Anderson, L. T., Meier, M., Cohen, I. L., Small, A. M., Samit, C., & Sachar, E. J. (1978). A comparison of haloperidol and behavior therapy and their interaction in autistic

children. *Journal of the American Academy of Child Psychiatry*, *17*(4), 640–655. https://doi.org/10.1016/S0002-7138(09)61017-7

Carlson, J. S. (2019). Introduction to school psychopharmacology. In J. S. Carlson & J. A. Barterian (Eds.), *School psychopharmacology: Translating research into practice* (pp. 1–11). Springer.

Carlson, G., Pokryzywinski, J., Uran, K., & Valdovinos, M. (2012). The use of reinforcer assessments in evaluating psychotropic medication effects. *Journal of Developmental and Physical Disabilities, 24*, 515–528. https://doi.org/10.1007/s10882-012-9282-4

*Cinciripini, P. M., Epstein, L. H., & Kotanchik, N. L. (1980). Behavioral intervention for self-stimulatory, attending, and seizure behavior in a cerebral palsied child. *Journal of Behavior Therapy and Experimental Psychiatry*, *11*(4), 313–316. https://doi.org/cgkzdf

*Cohen, I. L., Campbell, M., Posner, D., Small, A. M., Triebel, D., & Anderson, L. T. (1980). Behavioral effects of haloperidol in young autistic children: An objective analysis using a within-subjects reversal design. *Journal of the American Academy of Child Psychiatry*, *19*(4), 665–677. https://doi.org/10.1016/S0002-7138(09)60969-9

Courtemanche, A. B., Schroeder, S. R., & Sheldon, J. B. (2011). Designs and analyses of psychotropic and behavioral interventions for the treatment of problem behavior among people with intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disabilities*, *116*(4), 315–328. https://doi.org/cztknb

Cox, A. D., & Virues-Ortega, J. (2016). Interactions between behavior function and psychotropic medication. *Journal of Applied Behavior Analysis*, *49*(1), 85–104. https://doi.org/f8dq7r

*Crosland, K. A., Zarcone, J. R., Lindauer, S. E., Valdovinos, M. G., Zarcone, T. J., Hellings, J. A., & Schroeder, S. R. (2003). Use of functional analysis methodology in the evaluation of

medication effects. *Journal of Autism and Developmental Disorders*, *33*(3), 271–279. https://doi.org/10.1023/A:1024402500425

Crystal, S., Olfson, M., Huang, C., Pincus, H., & Gerhard, T. (2009). Broadened use of atypical antipsychotics: Safety, effectiveness, and policy challenges. *Health Affairs*, *28*(5), 770–781. https://doi.org/10.1377/hlthaff.28.5.w770

*Danov, S. E., Tervo, R., Meyers, S., & Symons, F. J. (2012). Using functional analysis methodology to evaluate effects of an atypical antipsychotic on severe problem behavior. *Journal of Mental Health Research in Intellectual Disabilities*, *5*(3–4), 286–308. https://doi.org/10.1080/19315864.2011.594976

*Davidson, P. W., Kleene, B. M., Carroll, M., & Rockowitz, R. J. (1983). Effects of naloxone on self-injurious behavior: A case study. *Applied Research in Mental Retardation*, *4*(1), 1–4. https://doi.org/10.1016/S0270-3092(83)80013-7

*Davis, K. V. (1971). The effect of drugs on stereotyped and nonstereotyped operant behaviors in retardates. *Psychopharmacologia*, *22*(2), 195–213. https://doi.org/10.1007/BF00403628

*Davis, K. V., Sprague, R. L., & Werry, J. S. (1969). Stereotyped behavior and activity level in severe retardates: The effect of drugs. *American Journal of Mental Deficiency*, *73*(5), 721–727.

*Durand, V. M. (1982). A behavioral/pharmacological intervention for the treatment of severe self-injurious behavior. *Journal of Autism and Developmental Disorders*, *12*(3), 243–251. https://doi.org/10.1007/BF01531370

*Fisher, W., Piazza, C. C., & Page, T. J. (1989). Assessing independent and interactive effects of behavioral and pharmacologic interventions for a client with dual diagnoses. *Journal of Behavior Therapy and Experimental Psychiatry*, *20*(3), 241–250. https://doi.org/bf34qc

*Frazier, T. W., Youngstrom, E. A., Haycook, T., Sinoff, A., Dimitriou, F., Knapp, J., & Sinclair, L. (2010). Effectiveness of medication combined with intensive behavioral intervention for reducing aggression in youth with autism spectrum disorder. *Journal of Child and Adolescent Psychopharmacology*, *20*(3), 167–177. http://doi.org/btj9hp

Ganz, J. B., & Ayres, K. M. (2018). Methodological standards in single-case experimental design: Raising the bar. *Research in Developmental Disabilities*, *79*, 3–9. https://doi.org/10.1016/j.ridd.2018.03.003

Gleason, M. M., Egger, H. L., Emslie, G. J., Greenhill, L. L., Kowatch, R. A., Lieberman, A. F., Luby, J. L., Owens, J., Scahill, L. D., Scheeringa, M. S., Stafford, B., Wise, B., & Zeanah, C. H. (2007). Psychopharmacological treatment for very young children: Contexts and guidelines. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*(12), 1532–1572. https://doi.org/10.1097/chi.0b013e3181570d9e

Greenhill, L. L., Pliszka, S., & Dulcan, M. K. (2002). Practice parameter for the use of stimulant medications in the treatment of children, adolescents, and adults. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(2), 26–49. https://doi.org/d86rmk

*Grondhuis, S. N., Farmer, C. A., Arnold, L. E., Gadow, K. D., Findling, R. L., Molina, B. S. G., Kolko, D. J., Buchan-Page, K. A., Rice, R. R., Butter, E. M., & Aman, M. G. (2019). Standardized observation analogue procedure in the treatment of severe childhood aggression study. *Journal of Child and Adolescent Psychopharmacology, 30*(1), 48–54. https://doi.org/d7nx

*Handen, B. L., Johnson, C. R., Butter, E. M., Lecavalier, L., Scahill, L., Aman, M. G., McDougle, C. J., Arnold, L. E., Swiezy, N. B., Sukhodolsky, D. G., Mulick, J. A., White, S. W., Bearss, K., Hollway, J. A., Stigler, K. A., Dziura, J., Yu, S., Sacco, K., & Vitiello, B.

(2013). Use of a direct observational measure in a trial of risperidone and parent training in children with pervasive developmental disorders. *Journal of Developmental and Physical Disabilities*, *25*(3), 355–371. https://doi.org/10.1007/s10882-012-9316-y

Hanley, G. P. (2012). Functional assessment of problem behavior: Dispelling myths, overcoming implementation obstacles, and developing new lore. *Behavior Analysis in Practice, 5*(1), 54–72. https://doi.org/10.1007/BF03391818

Harrison, J. N., Cluxton-Keller, F., & Gross, D. (2012). Antipsychotic medication prescribing trends in children and adolescents. *Journal of Pediatric Health Care, 26*(2), 139–145. https://doi.org/10.1016/j.pedhc.2011.10.009

*Helper, M. M., Wilcott, R. C., & Garfield, S. L. (1963). Effects of chlorpromazine on learning and related processes in emotionally disturbed children. *Journal of Consulting Psychology*, *27*(1), 1–9. https://doi.org/10.1037/h0044323

*Herman, B. H., Hammock, M. K., Arthur-Smith, A., Egan, J., Chatoor, I., Werner, A., & Zelnik, N. (1987). Naltrexone decreases self-injurious behavior. *Annals of Neurology*, *22*(4), 550–552. https://doi.org/10.1002/ana.410220419

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal, 343,* 1–9. https://doi.org/10.1136/bmj.d5928

*Hollis, J. H., & St. Omer, V. V. (1972). Chlorpromazine: Direct measurement of differential behavioral effect. *Science*, *159*(3822), 1487–1489. https://doi.org/fqd2tk

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education.

Exceptional Children, 71(2), 165–179. https://doi.org/10.1177/001440290507100203

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1994). Toward a functional analysis of self-injury. *Journal of Applied Behavior Analysis, 27*(2), 197-209. https://doi.org/10.1901/jaba.1994.27-197

*Johnson, K., Johnson, C. R., & Sahl, R. A. (1994). Behavioral and naltrexone treatment of self-injurious behavior. *Journal of Developmental and Physical Disabilities*, *6*(2), 193–202. https://doi.org/10.1007/BF02579359

Klove, H. (1963). Clinical neuropsychology. In F. M. Forster (Ed.), *The medical clinics of North America* (Vol. 47, pp. 1647–1658). Saunders.

LaRue, R. H., Northup, J., Baumeister, A. A., Hawkins, M. F., Seale, L., Williams, T., & Ridgeway, A. (2008). An evaluation of stimulant medication on the reinforcing effects of play. *Journal of Applied Behavior Analysis, 41*(1), 143–147. https://doi.org/cswh6t

Ledford, J. R., Lane, J. D., & Gast, D. L. (2018). Dependent variables, measurement, and reliability. In J. R. Ledford & D. L. Gast (Eds.), *Single case research methodology* (3rd ed., pp. 97–131). Routledge.

Li, A., MacNeill, B., Curiel, H., & Poling, A. (2017). Risperidone in combination with other drugs: Experimental research in individuals with autism spectrum disorder. *Experimental and Clinical Psychopharmacology, 25*(5), 434–439. https://doi.org/10.1037/pha0000144

Lloyd, B. P., Torelli, J. N., & Symons, F. J. (2016). Issues in integrating psychotropic and intensive behavioral interventions for students with emotional and behavioral challenges in schools. *Journal of Emotional and Behavioral Disorders, 24*(3), 148–158. https://doi.org/f83jbb

*Luiselli, J. K. (1986). Behavior analysis of pharmacological and contingency management

interventions for self-injury. *Journal of Behavior Therapy and Experimental Psychiatry*, *17*(4), 275–284. https://doi.org/10.1016/0005-7916(86)90064-9

*Luiselli, J. K., Evans, T. P., & Boyce, D. A. (1986). Pharmacological assessment and comprehensive behavioral intervention in a case of pediatric self-injury. *Journal of Clinical Child Psychology*, *15*(4), 323–326. https://doi.org/10.1207/s15374424jccp1504_6

*Mace, F. C., Blum, N. J., Sierp, B. J., Delaney, B. A., & Mauk, J. E. (2001). Differential response of operant self-injury to pharmacologic versus behavioral treatment. *Journal of Developmental and Behavioral Pediatrics*, *22*(2), 85–91. http://doi.org/dfv4hh

Mattison, R. E., Rundberg-Rivera, V., & Michel, C. (2014). Psychotropic medication characteristics for special education students with emotional and/or behavioral disorders. *Journal of Child and Adolescent Psychopharmacology, 24*(6), 347–353. http://doi.org/f6d3t5

McLaren, J. L., Barnett, E. R., Concepcion Zayas, M. T., Lichtenstein, J., Acquilano, S. C., Schwartz, L. M., Woloshin, S., & Drake, R. E. (2018). Psychotropic medications for highly vulnerable children. *Expert Opinion on Pharmacotherapy*, *19*(6), 547–560. https://doi.org/ggk73n

McLaren, J. L., & Lichtenstein, J. D. (2019). The pursuit of the magic pill: The overuse of psychotropic medications in children with intellectual and developmental disabilities in the USA. *Epidemiology and Psychiatric Sciences, 28*(4), 365–368. https://doi.org/d9gv

McVoy, J., & Findling, R. L. (2017). *Clinical manual of child and adolescent psychopharmacology* (3rd ed.). American Psychiatric Association.

Meltzer, H. Y. (2004). What's atypical about atypical antipsychotic drugs? *Current Opinion in Pharmacology, 4*(1), 53–57. https://doi.org/10.1016/j.coph.2003.09.010

*Miguel, C. F., Clark, K., Tereshko, L., Ahearn, W. H., & Zarcone, J. (2009). The effects of response interruption and redirection and sertraline on vocal stereotypy. *Journal of Applied Behavior Analysis*, *42*(4), 883–888. https://doi.org/10.1901/jaba.2009.42-883

*Moore, T. R., Tervo, R. C., McComas, J. J., Rivard, P. F., & Symons, F. J. (2009). Longitudinal functional analysis of problem behavior during an atypical neuroleptic medication cross-over evaluation for an adolescent with developmental disabilities. *Education & Treatment of Children*, *32*(1), 105–119. https://doi.org/10.1353/etc.0.0042

Napolitano, D. A., Jack, S. L., Sheldon, J. B., Williams, D. C., McAdam, D. B., & Schroeder, S. R. (1999). Drug-behavior interactions in persons with mental retardation and developmental disabilities. *Mental Retardation and Developmental Disabilities Research Reviews*, *5*(4), 322–334. http://doi.org/fp2526

National Institute of Mental Health. (2016). *Mental health medications.* Available from https://www.nimh.nih.gov/

Northup, J., Fusilier, I., Swanson, V., Roane, H., & Borrero, J. (1997). An evaluation of methylphenidate as a potential establishing operation for some common classroom reinforcers. *Journal of Applied Behavior Analysis, 30*(4), 615–625. https://doi.org/fn89bf

Olfson, M., Druss, B. G., & Marcus, S. C. (2015). Trends in mental health care among children and adolescents. *New England Journal of Medicine*, *372*(21), 2029–2038. https://doi.org/ggj2s3

Parikh, M. S., Kolevzon, A., & Hollander, E. (2008). Psychopharmacology of aggression in children and adolescents with autism: A critical review of efficacy and tolerability. *Journal of Child and Adolescent Psychopharmacology, 18*(2), 157–178. http://doi.org/c7vgt2

*Piazza, C. C., Fisher, W., Hyman, S. L., Fleishell, J., Lou, K. K., & Cataldo, M. (1994).

Evaluation of pharmacologic treatment of destructive behaviors: Aggregated results from single-case experimental studies. *Journal of Developmental and Physical Disabilities*, *6*(2), 149–168. https://doi.org/10.1007/BF02579357

Pliszka, S., & AACAP Work Group on Quality Issues. (2007). Practice parameter for the assessment and treatment of children and adolescents with attention-deficit/hyperactivity disorder. *Journal of American Child Adolescent Psychiatry, 46(*7), 894–921. https://doi.org/dtdcks

Reed, F. D., & Codding, R. S. (2014). Advancements in procedural fidelity assessment and intervention: Introduction to the special issue. *Journal of Behavioral Education, 23*, 1–18. https://dx.doi.org/10.1007/s10864-013-9191-3

Reichow, B., Barton, E. E., & Maggin, D. M. (2018). Development and applications of the single-case design risk of bias tool for evaluating single-case design research study reports. *Research in Developmental Disabilities, 79*, 53–64. https://doi.org/gdq94g

Rieken, C. J., Griffith, A. K., D'Angelo, J. H., & Re, T. (2019). Psychotropic medication prescription for autism: Data sources for decision making. *Advances in Neurodevelopmental Disorders, 3*, 1–7. https://doi.org/10.1007/s41252-018-0078-0

Roane, H. S. (2008). On the applied use of progressive ratio schedules of reinforcement. *Journal of Applied Behavior Analysis, 41*(2), 155–161. https://doi.org/10.1901/jaba.2008.41-155

Rush, A. J., & Frances, A. (Eds.). (2000). Expert consensus guidelines series: Treatment of psychiatric and behavioral problems in mental retardation. *American Journal on Mental Retardation, 105*(3), 159–228.

*Ryan, E. P., Helsel, W. J., Lubetsky, M. J., Miewald, B. K., Hersen, M., & Bridge, J. (1989). Use of naltrexone in reducing self-injurious behavior: A single case analysis. *Journal of the*

Multihandicapped Person, *2*(4), 295–309. https://doi.org/10.1007/BF01098171

*Sallee, F. R., Sethuraman, G., & Rock, C. M. (1994). Effects of pimozide on cognition in children with Tourette syndrome: Interaction with comorbid attention deficit hyperactivity disorder. *Acta Psychiatrica Scandinavica*, *90*(1), 4–9. https://doi.org/b2sqr2

*Sand, P. L., & Carlson, C. (1973). Failure to establish control over tics in the Gilles de la Tourette syndrome with behaviour therapy techniques. *The British Journal of Psychiatry*, *122*(571), 665–670. https://doi.org/10.1192/bjp.122.6.665

*Sandford, D., & Nettelbeck, T. (1982). Medication and reinforcement within a token programme for disturbed mentally retarded residents. *Applied Research in Mental Retardation*, *3*(1), 21–36. https://doi.org/10.1016/0270-3092(82)90056-X

Schaal, D. W., & Hackenberg, T. (1994). Toward a functional analysis of drug treatment for behavior problems of people with developmental disabilities. *American Journal on Mental Retardation, 99*(2), 123–140.

Scheifes, A., de Jong, D., Stolker, J. J., Nijman, H. L. I., Egberts, T. C. G., & Heerdink, E. R. (2013). Prevalence and characteristics of psychotropic drug use in institutionalized children and adolescents with mild intellectual disability. *Research in Developmental Disabilities*, *34*(10), 3159–3167. https://doi.org/10.1016/j.ridd.2013.06.009

Schroeder, S. R., Lewis, M. H., & Lipton, M. A. (1983). Interactions of pharmacotherapy and behavior therapy among children with learning and behavioral disorders. *Advances in Learning & Behavioral Disabilities*, *2*, 179–225.

*Singh, N. N., & Aman, M. G. (1981). Effects of thioridazine dosage on the behavior of severely mentally retarded persons. *American Journal of Mental Deficiency*, *85*(6), 580–587.

*Singh, N. N., Landrum, T. J., Ellis, C. R., & Donatelli, L. S. (1993). Effects of thioridazine and

visual screening on stereotypy and social behavior in individuals with mental retardation. *Research in Developmental Disabilities*, *14*(3), 163–177. https://doi.org/cs9jmz

*Singh, N. N., & Winton, A. S. W. (1984). Behavioral monitoring of pharmacological interventions for self-injury. *Applied Research in Mental Retardation*, *5*(1), 161–170. https://doi.org/10.1016/S0270-3092(84)80028-4

*Slifer, K. J., Ivancic, M. T., Parrish, J. M., Page, T. J., & Burgio, L. D. (1986). Assessment and treatment of multiple behavior problems exhibited by a profoundly retarded adolescent. *Journal of Behavior Therapy and Experimental Psychiatry*, *17*(3), 203–213. https://doi.org/dbv8mw

*Snyder, R., Turgay, A., Aman, M., Binder, C., Fisman, S., & Carroll, A. (2002). Effects of risperidone on conduct and disruptive behavior disorders in children with subaverage IQs. *Journal of the American Academy of Child & Adolescent Psychiatry*, *41*(9), 1026–1036. https://doi.org/10.1097/00004583-200209000-00002

Spencer, D., Marshall, J., Post, B., Kulakodlu, M., Newschaffer, C., Dennen, T., Azocar, F., & Jain, A. (2013). Psychotropic medication use and polypharmacy in children with autism spectrum disorders. *Pediatrics, 132*(5), 833-840. https://doi.org/10.1542/peds.2012-3774

*Sprague, R. L., Barnes, K. R., & Werry, J. S. (1970). Methylphenidate and thioridazine: Learning, reaction time, activity, and classroom behavior in disturbed children. *American Journal of Orthopsychiatry*, *40*(4), 615–628. http://doi.org/bh2sqw

Stovold, E., Beecher, D., Foxlee, R., & Noel-Storr, A. (2014). Study flow diagrams in Cochrane systematic review updates: An adapted PRISMA flow diagram. *Systematic Reviews, 3*(1), 1–5. https://doi.org/10.1186/2046-4053-3-54

*Strong, C., Sulzbacher, S. I., & Kirkpatrick, M. A. (1974). Use of medication versus

reinforcement to modify a classroom behavior disorder. *Journal of Learning Disabilities*, *7*(4), 214–218. https://doi.org/10.1177/002221947400700406

*Sverd, J., Kupietz, S. S., Winsberg, B. G., Hurwic, M. J., & Becker, L. (1978). Effects of l-5-hydroxytryptophan in autistic children. *Journal of Autism & Childhood Schizophrenia*, *8*(2), 171–180. https://doi.org/10.1007/BF01537866

Thompson, T., Moore, T., & Symons, F. (2007). Psychotherapeutic medications and positive behavior support. In S. L. Odom, R. H. Horner, M. E. Snell, & J. Blacher (Eds.), *Handbook of developmental disabilities* (pp. 501–527). Guilford Press.

*Valdovinos, M. G., Nelson, S. M., Kuhle, J. L., & Dierks, A. M. (2009). Using analogue functional analysis to measure variations in problem behavior rate and function after psychotropic medication changes: A clinical demonstration. *Journal of Mental Health Research in Intellectual Disabilities*, *2*(4), 279–293. https://doi.org/fmppf2

Van Haaren, F., & Weeden, M. (2013). Some guidelines for conducting research in applied behavioral pharmacology. *Journal of Applied Behavior Analysis*, *46*(2), 498–506. https://doi.org/10.1002/jaba.47

Vitiello, B. (2017). Developmental aspects of pediatric psychopharmacology. In M. McVoy & R. L. Findling (Eds.), *Clinical manual of child and adolescent psychopharmacology* (3rd ed., pp. 1–28). American Psychiatric Association Publishing.

Waschbusch, D. A., Pelham, W. E., Waxmonsky, J., & Johnson, C. (2009). Are there placebo effects in the medication treatment of children with attention-deficit hyperactivity disorder? *Journal of Developmental & Behavioral Pediatrics, 30*(2), 158–168. http://doi.org/fttpkj

*Waxmonsky, J. G., Waschbusch, D. A., Pelham, W. E., Draganac-Cardona, L., Rotella, B., & Ryan, L. (2010). Effects of atomoxetine with and without behavior therapy on the school

and home functioning of children with attention-deficit/hyperactivity disorder. *The Journal of Clinical Psychiatry*, *71*(11), 1535–1551. https://doi.org/10.4088/JCP.09m05496pur

*Werry, J. S., & Aman, M. G. (1975). Methylphenidate and haloperidol in children: Effects on attention, memory, and activity. *Archives of General Psychiatry*, *32*(6), 790–795. https://doi.org/10.1001/archpsyc.1975.01760240118009

Yoder, P., Lloyd, B., & Symons, F. (2018). *Observational measurement of behavior*. Springer.

*Zarcone, J. R., Hellings, J. A., Crandall, K., Reese, R. M., Marquis, J., Fleming, K., Shores, R., Williams, D., & Schroeder, S. R. (2001). Effects of risperidone on aberrant behavior of persons with developmental disabilities: I. A double-blind crossover study using multiple measures. *American Journal on Mental Retardation*, *106*(6), 525–538. http://doi.org/ctkpc2

*Zarcone, J. R., Lindauer, S. E., Morse, P. S., Crosland, K. A., Valdovinos, M. G., McKerchar, T. L., Matthew Reese, R., Hellings, J. A., & Schroeder, S. R. (2004). Effects of risperidone on destructive behavior of persons with developmental disabilities: III. Functional analysis. *American Journal of Mental Retardation*, *109*(4), 310–321. http://doi.org/cps877

Zarcone, J., Napolitano, D., & Valdovinos, M. (2008). Measurement of problem behaviour during medication evaluations. *Journal of Intellectual Disability Research*, *52*(12), 1015–1028. https://doi.org/10.1111/j.1365-2788.2008.01109.x

Zito, J. M., Derivan, A. T., Kratochvil, C. J., Safer, D. J., Fegert, J. M., & Greenhill, L. L. (2008). Off-label psychopharmacologic prescribing for children: History supports close clinical monitoring. *Child and Adolescent Psychiatry and Mental Health, 2*(1), 24. https://doi.org/10.1186/1753-2000-2-24

Figure 1

PRISMA Flow Diagram Showing Screening and Article Identification Process

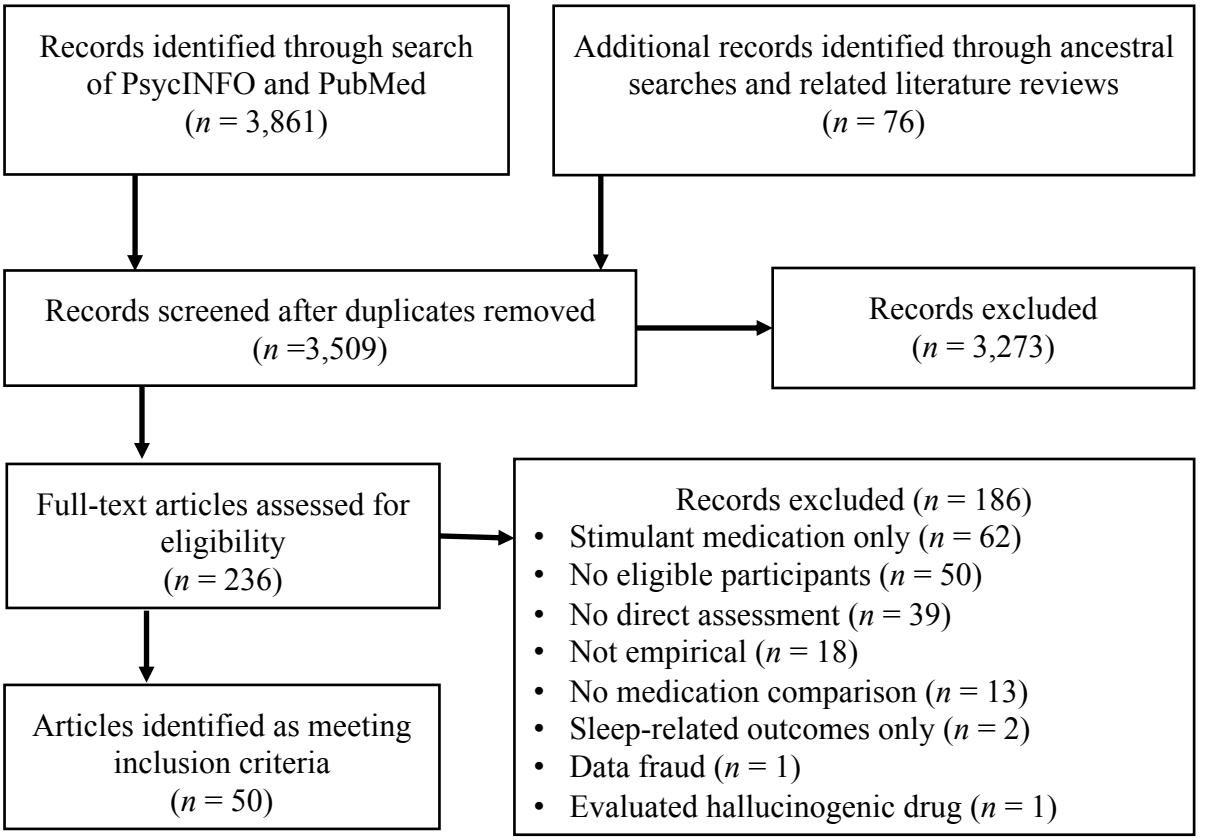

Table 1

Numbers and Percentages of Studies Reporting Participant Disabilities and Disorders

Disability/Disorder	*n* (%)
Intellectual disability	33 (66.0)
Autism spectrum disorder	18 (36.0)
Attention deficit hyperactivity disorder	9 (18.0)
Seizure disorder	7 (14.0)
Disruptive, impulse-control, or conduct disorder	6 (12.0)
Motor disorders	6 (12.0)
Cerebral palsy	5 (10.0)
Pervasive developmental disorder	5 (10.0)
Schizophrenia spectrum or other psychotic disorder	3 (6.0)
Developmental delay	2 (4.0)
Fragile X syndrome	2 (4.0)
Hearing impairment	2 (4.0)
Tuberous sclerosis	2 (4.0)
Visual impairment or blind	2 (4.0)
3q29 deletion	1 (2.0)
Agenesis of the corpus callosum	1 (2.0)
Coffin-Lowery syndrome	1 (2.0)
Cognitive delay	1 (2.0)
Cornelia de Lange syndrome	1 (2.0)
Cri-du-chat (5-p) syndrome	1 (2.0)
Cyclical behavior	1 (2.0)
Down syndrome	1 (2.0)
Emotional disturbance	1 (2.0)
L hemiplegia	1 (2.0)
Mood disorder	1 (2.0)
Personality disorder	1 (2.0)
Phenylketonuria	1 (2.0)
Obsessive-compulsive disorder	1 (2.0)
Organic brain syndrome	1 (2.0)
Rickets	1 (2.0)
Sensory integration disorder	1 (2.0)
Severe hyperactivity	1 (2.0)
Sleep-wake disorder	1 (2.0)
Subaverage IQ	1 (2.0)
Wardenberg's syndrome	1 (2.0)

Note. Disabilities and disorders are listed based on author report. IQ = intelligence quotient.

Table 2

Study Characteristics, Direct Assessment Types, and Quality Indicators Listed by Study

Study Characteristics				Direct Assessment Types								Direct Assessment Quality Indicators		
Citation	N	Design	Medication Class(es)	Unst	BTx	FA	Disc	M/A	I/P	Mot	Oth	Rel	Blind	Fid
Helper 1963	39	Between Group	Typical Antipsychotic	–	–	–	–	X	–	X	–	–	X	–
Campbell 1978	40	Between Group	Typical Antipsychotic	–	X	–	–	–	X	–	–	X	X	–
Sallee 1994	66	Between Group	Typical Antipsychotic	–	–	–	–	X	–	–	–	–	–	–
Snyder 2002	110	Between Group	Atypical Antipsychotic	–	–	–	–	X	–	–	–	n/a	n/a	–
Waxmonsky 2010	56	Between Group	SNRI	X	X	–	–	–	–	–	–	X	–	–
Handen 2013	124	Between Group	Atypical Antipsychotic	–	–	X	–	–	–	–	–	X	X	–
Grondhuis 2019	149	Between Group	Atypical Antipsychotic Stimulant	–	–	X	–	–	–	–	–	X	X	–
Frazier 2010	32	Between Group (Nonexperimental retrospective review)	Atypical Antipsychotic Typical Antipsychotic Anticonvulsant SNRI Alpha-2 Agonist Mood Stabilizers	–	X	–	–	–	–	–	–	–	–	–
Mace 2001	13	Between Group, Within Subjects	Typical Antipsychotic	–	X	–	–	–	–	–	–	X	X	–
Davis 1969	9*	Within Subjects	Typical Antipsychotic Stimulant	X	–	–	–	–	–	–	–	X	–	n/a
Sprague 1970	12	Within Subjects	Typical Antipsychotic Stimulant	X	–	–	–	X	–	–	–	X	X	–
Davis 1971	10*	Within Subjects	Typical Antipsychotic Stimulant	–	–	–	X	–	–	–	–	n/a	n/a	–
Werry 1975	24	Within Subjects	Typical Antipsychotic Stimulant	–	–	–	–	X	–	–	–	–	X	–
Cohen 1980	10	Within Subjects	Typical Antipsychotic	X	–	–	–	–	X	–	–	–	X	–
Singh 1981	19*	Within Subjects	Typical Antipsychotic Stimulant	X	–	–	–	X	X	–	–	X	X	–
Aman 1984	2	Within Subjects	Typical Antipsychotic	X	–	–	–	–	–	–	X	X	–	–
Anderson 1984	40	Within Subjects	Typical Antipsychotic	–	–	–	X	–	–	–	–	n/a	n/a	–
Herman 1987	3	Within Subjects	Opioid Antagonist	X	–	–	–	–	–	–	–	X	X	n/a
Aman 1989	20*	Within Subjects	Typical Antipsychotic Anticonvulsant	X	–	–	–	–	–	–	–	X	X	n/a
Anderson 1989	45	Within Subjects	Typical Antipsychotic	–	–	–	X	–	–	–	–	–	–	–
Aman 1991	27	Within Subjects	Typical Antipsychotic Stimulant	–	–	–	X	X	–	X	–	–	–	–
Zarcone 2001	3	Within Subjects	Atypical Antipsychotic	X	–	–	–	–	–	–	–	X	X	n/a
Hollis 1972	4	Single Case (Yes)	Typical Antipsychotic	–	–	–	X	–	–	–	–	n/a	n/a	–

Study Characteristics				Direct Assessment Types								Direct Assessment Quality Indicators		
Citation	N	Design	Medication Class(es)	Unst	BTx	FA	Disc	M/A	I/P	Mot	Oth	Rel	Blind	Fid
Sanford 1982	2	Single Case (Yes)	Typical Antipsychotic	–	X	–	–	–	–	–	–	X	–	–
Davidson 1983	1	Single Case (Yes)	Opioid Antagonist	X	–	–	–	–	–	–	–	X	X	n/a
Barrett 1989	1	Single Case (Yes)	Opioid Antagonist	–	–	–	–	–	–	–	X	X	X	–
Singh 1993	3	Single Case (Yes)	Typical Antipsychotic	–	X	–	–	–	–	–	–	X	X	X
Danov 2012	4	Single Case (Yes)	Atypical Antipsychotic	–	–	X	–	–	–	–	–	X	X	–
Sand 1973	1	Single Case (No)	Typical Antipsychotic	–	–	–	–	–	–	–	X	–	–	–
Strong 1974	1	Single Case (No)	Antihistamine	–	–	–	–	–	X	–	–	–	X	–
Sverd 1978	3	Single Case (No)	Decarboxylase Inhibitor	X	–	–	–	–	–	–	–	–	–	n/a
Cinciripini 1980	1	Single Case (No)	Anticonvulsant	–	X	–	–	–	–	–	–	X	–	–
Durand 1982	1	Single Case (No)	Typical Antipsychotic	X	X	–	–	–	–	–	–	X	–	–
Singh 1984	1	Single Case (No)	Typical Antipsychotic Anticonvulsant	X	–	–	–	–	–	–	–	X	X	–
Burgio 1985	1	Single Case (No)	Typical Antipsychotic Stimulant	X	–	–	–	–	X	–	–	–	X	–
Luiselli 1986a	1	Single Case (No)	Typical Antipsychotic	–	X	–	–	–	–	–	–	X	–	–
Luiselli 1986b	2	Single Case (No)	Typical Antipsychotic	–	X	–	–	–	–	–	–	X	–	–
Slifer 1986	1	Single Case (No)	Typical Antipsychotic	–	X	–	–	–	–	–	–	X	–	X
Barrett 1988	1	Single Case (No)	Anticonvulsant	–	X	–	–	–	–	–	–	X	X	–
Fisher 1989	1	Single Case (No)	Typical Antipsychotic	–	X	–	–	–	–	–	–	X	–	–
Ryan 1989	1	Single Case (No)	Opioid Antagonist	–	X	–	–	–	–	–	–	X	X	–
Beale 1993	4	Single Case (No)	Typical Antipsychotic	–	–	–	X	–	–	–	–	n/a	n/a	–
Johnson 1994	1	Single Case (No)	Opioid Antagonist	–	X	–	–	–	–	–	–	–	–	–
Piazza 1994	12	Single Case (No)	Typical Antipsychotic Anticonvulsant Alpha-2 Agonist Mood stabilizer Antidepressant Beta blocker	–	X	–	–	–	–	–	–	–	–	–
Anderson 2002	1	Single Case (No)	Typical Antipsychotic	–	–	X	–	–	–	–	–	X	–	–
Crosland 2003	1	Single Case (No)	Atypical Antipsychotic	–	–	X	–	–	–	–	–	X	X	–
Zarcone 2004	8	Single Case (No)	Atypical Antipsychotic	–	–	X	–	–	–	–	–	X	X	–
Miguel 2009	1	Single Case (No)	SSRI	–	X	–	–	–	–	–	–	X	–	–
Moore 2009	1	Single Case (No)	Atypical Antipsychotic	–	–	X	–	–	–	–	–	X	X	–
Valdovinos 2009	1	Single Case (No)	Atypical Antipsychotic Stimulant Alpha-2 Agonist	–	–	X	–	–	–	–	–	X	–	–
			Percentage of studies:	28.0	34.0	16.0	12.0	14.0	10.0	4.0	6.0	71.1	53.3	4.5

Note. X = present; – = absent; n/a = not applicable; Unst = Unstructured; BTx = Behavioral treatment; FA = Functional analysis; Disc = Discrimination tasks; M/A = Memory or attention tasks; I/P = Other instructional or prosocial tasks; Mot = Motor tasks; Oth = Other structured assessments; Rel = Reliability; Blind = Blinded observers; Fid = Fidelity; SNRI = selective norepinephrine reuptake inhibitor; SSRI = selective serotonin reuptake inhibitor. Asterisks indicate total *N* for group design studies in which only a subset of participants met inclusion criteria for this review.

Table 3

Numbers and Percentages of Studies Evaluating each Medication (Generic, Organized by Medication Class)

Medications	*n* (%)
Antipsychotic (Typical)	30 (60.0)
Haloperidol	14 (28.0)
Thioridazine	12 (24.0)
Chlorpromazine	5 (10.0)
Pimozide	3 (6.0)
Fluphenazine	1 (2.0)
Molindone	1 (2.0)
Antipsychotic (Atypical)	10 (20.0)
Risperidone	9 (18.0)
Aripiprazole	3 (6.0)
Clozapine	1 (2.0)
Olanzapine	1 (2.0)
Quetiapine	1 (2.0)
Ziprasidone	1 (2.0)
Stimulant (in combination with another medication class)	9 (18.0)
Methylphenidate	8 (16.0)
Dextroamphetamine	2 (4.0)
Anticonvulsant	5 (10.0)
Carbamazepine	5 (10.0)
Opioid Antagonist	5 (10.0)
Naltrexone	4 (8.0)
Naloxone	2 (4.0)
Alpha-2 Agonists	3 (6.0)
Clonidine	2 (4.0)
Guanfacine	1 (2.0)
Selective Norepinephrine Reuptake Inhibitor (SNRI)	2 (4.0)
Atomoxetine	2 (4.0)
Mood Stabilizers	2 (4.0)
Lithium	2 (4.0)
Divalproex sodium	1 (2.0)
Lamotrigine	1 (2.0)
Selective Serotonin Reuptake Inhibitor (SSRI)	1 (2.0)
Sertraline	1 (2.0)
Other	2 (4.0)
Carbidopa	1 (2.0)
Diphenhydramine hydrochloride	1 (2.0)
L-5-hydroxytryptophan	1 (2.0)
Nortriptyline	1 (2.0)
Propranolol	1 (2.0)