Pegem Journal of Education and Instruction, Vol. 12, No. 4, 2022 (pp. 370-383)

RESEARCH ARTICLE WWW.PEGEGOG.NET

Exploring Caregivers' Views on Preference Assessment for Children with Autism Spectrum Disorder

Dr. Abdullah Abubakr

Assistant Professor at College of Education, Imam Abdulrahman Bin Faisal University, P.O. Box: 2375, Dammam, Saudi Arabia, Email: aybakr@iau.edu.sa

Abstract

Selecting the appropriate method for assessing students' preferences affects their educational, social, and behavioral progress. This study explores the influence of preference assessment among students with autism spectrum disorder through their teachers' perceptions and evaluates the ability of these ratings to predict the strongest reinforcers. The researcher used an annotated bibliography when writing the literature review for this study, collecting several studies that explored the effects of different preference assessments, including multiple-stimulus (non-replacement) preference assessment, teacher interview, direct preference assessment, picture-paired preference assessment, and single-stimulus reinforcer assessment, to identify effective reinforcers.

This study involved 112 caregivers of students with autism spectrum disorder (a behavior analyst and a special education teacher). The results indicated that Free Operant Observation and Multiple Stimulus Without Replacement were the most effective assessment tools for identifying appropriate reinforcers for students with autism spectrum disorder. The study also revealed statistically significant differences based on the experience variable, as service providers with six to ten years of experience were more aware of the impact of both Free Operant Observation and Multiple Stimulus Without Replacement on determining appropriate reinforcement, while the results did not show any significant differences for the remaining variables.

This study identified several useful and effective preference assessment tools for students with developmental disabilities. The researcher recommends developing modern assessment tools that are evidence-based, accurate, valid, reliable, and easy to use. Additionally, it is essential to consider the students' condition and level of interaction.

Keywords: Preference Assessment, Students with Autism Spectrum Disorder, Reinforcers

Introduction

A preference assessment is a method of observation or experimentation that helps teachers and practitioners identify a "preference hierarchy" for students by positively arousing their interest (Keen, 2009; Keen & Pennell, 2010; Kodak et al., 2009; Kang et al., 2013; Parsons et al., 1990). Hagopian et al. (2004) have pointed out that a preference assessment is an organized list of items or activities ranked by preference. These preferences can be determined by asking students about the items they would prefer to receive, using feedback or trial-and-error assessments, and then practitioners can determine this hierarchy. The preference structure indicates which items a child strongly prefers, moderately prefers, and least prefers. These preferences can then be utilized to reinforce positive behavior and increase the likelihood of it being repeated in the future. Preference assessment for individuals with developmental disabilities is a common practice used to design and implement educational and behavioral interventions for this population (Hagopian, Long, & Rush, 2004; DeLeon & Iwata, 1996; Fisher et al., 1996). For

children with language delays, such as those with autism and behavioral disorders, who often experience low levels of engagement with the physical and social world, identifying their interests and leveraging them to enhance motivation, increased task engagement, decreased problem behavior, and participation is an effective intervention strategy (Keen, 2009). Most previous research has examined preference assessment methods with groups possessing relatively well-developed verbal repertoires, such as adults with schizophrenia and children with attention-deficit/hyperactivity disorder. These studies noted that using spoken options to assess preferences took less time than options requiring motor responses, making it easier for researchers to evaluate the effectiveness of the preference assessment method and identify reinforcers (Paramore & Higbee, 2005; Wilder et al., 2003; Cohen-Almeida et al., 2000). However, there is less existing research evaluating the use of preference assessments among individuals with developmental disabilities who have language difficulties (Graff & Karsten., 2012; Ackerlund Brandt, et al., 2015)

LITERATURE REVIEW

This literature review focuses on a collection of studies describing the effectiveness of various preference assessment techniques used with individuals with language delays in general school settings and their role in identifying the most effective reinforcers. The study employed both an annotated bibliography method and a narrative method to compile data on instruments designed to evaluate the effectiveness or impact of different types of preference assessments in experimental studies, obtained through a comprehensive search of published literature and three electronic databases (PsycINFO, SCOPUS, and ERIC). Each study's recommendations and the generalizability of a particular type of preference assessment for widespread use with students with language delays were determined.

Preference Assessment

Multiple-Stimulus Without Replacement (MSWO) Assessments. Multiple-Stimulus Without Replacement (MSWO) assessments have gained widespread acceptance as an efficient and reliable method for identifying preferred stimuli to be used as reinforcers for individuals with developmental disabilities and behavioral disorders (DeLeon & Iwata, 1996; Fisher, Piazza, Bowman, Hagopian, Owens, & Slevin, 1992). The MSWO procedure involves presenting an array of stimuli from which the participant selects one item at a time without replacement, allowing researchers or practitioners to rank stimuli according to selection frequency and preference hierarchy. This approach not only identifies high-preference stimuli but also reduces assessment time compared to traditional single-stimulus methods (DeLeon, Iwata, & Roscoe, 1997; Hanley, Iwata, & McCord, 2003).

Carr, Nicolson, and Higbee (2000) conducted a seminal study to evaluate the efficiency and practical applicability of Multiple-Stimulus Without Replacement (MSWO) assessments with children diagnosed with autism spectrum disorder (ASD). In their study, the standard MSWO procedure was modified by reducing the number of stimulus presentation arrays from five to three to determine whether a shorter assessment could still provide valid preference data. Each participant completed three stimulus-presentation sessions, starting with eight stimuli arranged linearly, and selected stimuli sequentially, with chosen items removed from the array. Selection ratios were calculated for each stimulus, producing a rank order of preferences.

To verify the functional validity of the MSWO-identified reinforcers, Carr et al. paired three stimuli—categorized as high, medium, and low preference—with target behaviors from each participant's curriculum. The delivery of high-preference stimuli led to higher rates of correct responding compared to medium- and low-preference conditions, confirming that MSWO can reliably identify effective reinforcers. The study also examined preference stability across eight MSWO assessments conducted over four weeks. While some participants displayed stable preferences, others showed variability, highlighting the importance of ongoing reassessment in practice. Notably, a high level of agreement was observed between single-session and three-session MSWO results, suggesting that abbreviated assessments may be useful in time-sensitive contexts (Carr, Nicolson, & Higbee, 2000).

Further research supports the use of MSWO in clinical and educational settings. For example, Paramore and Higbee (2005) demonstrated that brief MSWO assessments could effectively identify reinforcers for adolescents with emotional and behavioral disorders in educational environments. Similarly, Morris, Allen, and Gallagher (2023) found that short MSWO sessions produced preference outcomes consistent with longer assessments, emphasizing the efficiency of abbreviated procedures while maintaining reliable results.

Expanding on this work, Daly et al. (2009) explored the utility of MSWO assessments within a school-based setting involving children with behavioral disorders. Unlike Carr et al. (2000), who focused on tangible food and leisure items, Daly et al. utilized activity-based stimuli—such as playing games, visiting the library, and

computer use—to identify preferred classroom activities that could serve as reinforcers. The assessment involved presenting activity cards to students, who sequentially selected preferred activities over three sessions. Reinforcer efficacy was evaluated by measuring performance on math worksheets when students had access to activities identified as high-, medium-, or low-preference.

The results demonstrated that students performed better on academic tasks when contingently reinforced with high-preference activities identified by the MSWO, replicating findings from clinical populations and extending the applicability of MSWO assessments to educational interventions. The study also highlighted that the probability of selecting high-preference stimuli was influenced by the number of available choices, with students more likely to select high-preference items when more options were present. Notably, Daly et al. successfully implemented the MSWO assessment in a small-group format within the school psychologist's office, providing evidence for the method's feasibility in naturalistic educational contexts. Collectively, these studies provide robust evidence supporting the efficiency, reliability, and ecological validity of MSWO preference assessments. Both investigations emphasize the practical benefits of brief MSWO procedures that yield stable preference data and effectively identify reinforcers that enhance target behaviors. While Carr et al. (2000) noted that their assessments, though described as naturalistic, were conducted in therapy rooms, Daly et al. (2009) advanced the field by implementing MSWO in a school environment, thereby increasing generalizability. The findings suggest that MSWO assessments can be adapted for diverse populations and settings, including individuals with ASD and students with behavioral disorders.

In summary, MSWO assessments represent a versatile and empirically supported method for identifying effective reinforcers. The work of Carr et al. (2000) and Daly et al. (2009) collectively supports the incorporation of MSWO procedures in both clinical and educational practices to enhance engagement and promote positive behavior change.

Lill JD, Shriver MD, Allen KD (2021) demonstrated that using stimuli selected through brief, repeated daily preference assessments significantly influenced the rate and pace of skill acquisition in children with autism spectrum disorder (ASD).

Direct and Indirect Preference Assessments. Preference assessments are widely used to identify stimuli that function as effective reinforcers for individuals with developmental and behavioral challenges. These assessments are typically classified as direct, involving systematic observation of the individual's choices, or indirect, relying on informant reports such as teacher or caregiver interviews (Piazza et al., 1996). Understanding the relative effectiveness and agreement between these methods is critical to selecting appropriate assessment tools in both clinical and educational settings.

Cote, Thompson, Hanley, and McCarrah (2007) investigated the effectiveness and agreement between indirect and direct preference assessment methods in identifying potent reinforcers for nine young children, including eight typically developing children and one child with a severe physical disability. The indirect assessment involved teacher interviews, during which groups of teachers collaboratively identified and ranked ten potential reinforcers from auditory and social stimulus categories. Consensus rankings were reached within 10–20 minutes, providing a rapid method for gathering indirect preference data.

Following the teacher interviews, a direct paired-stimulus preference assessment was conducted. All ten stimuli were presented in systematically randomized pairs, and participants' selections were recorded to establish a preference hierarchy based on the percentage of trials in which each stimulus was chosen. Each direct assessment was completed in a single 40–50-minute session per child, with physical prompting provided when necessary to ensure exposure and facilitate choice responding.

To evaluate the functional efficacy of the identified stimuli, Cote et al. (2007) conducted a reinforcer assessment in which top-ranked items from both teacher interviews and direct preference assessments were compared within a synchronous-operant arrangement. Children could select among two preferred stimuli and a no-stimulus control zone, with responses such as switch activations or time spent in designated zones serving as indicators of reinforcement effectiveness.

Results indicated variable agreement between the indirect and direct methods. For five of the nine children, directly assessed stimuli functioned as more effective reinforcers than those identified through teacher interviews, while only one child showed stronger reinforcement from teacher-identified items. Some children demonstrated consistent preferences across methods, whereas others exhibited discrepancies, highlighting the importance of direct observation when identifying potent reinforcers. Performance data showed that all children displayed increased engagement during access to high-preference items identified by direct assessment, while responding during low-preference conditions remained at baseline levels. These findings align with previous research demonstrating the

validity of direct preference assessments for guiding behavior change (Carr, Nicolson, & Higbee, 2000; Daly, et al., 2009).

Based on these results, Cote et al. (2007) recommended the use of direct preference assessments in educational and clinical settings, particularly for children who cannot verbally communicate preferences or show low motivation during instruction. Teacher interviews, while efficient, may not reliably identify the most effective reinforcers and should be supplemented with direct observational assessments whenever possible. This approach supports the inclusion of direct stimulus preference assessments as critical components in developing individualized Behavior Intervention Plans (BIPs) (DeLeon & Iwata, 1996; Fisher, Piazza, Bowman, Hagopian, Owens, & Slevin, 1992). Overall, the study underscores the importance of empirically verifying indirect preference information with participant-driven data to ensure interventions are grounded in accurately identified reinforcers, a practice particularly relevant in early childhood and special education contexts.

Free-operant preference. Free-operant preference assessment is a technique used in applied behavior analysis (ABA) to identify preferred stimuli without requiring the individual to choose from presented items. Instead, the individual is given access to a variety of items or activities in a naturalistic setting, and their interactions are observed and recorded over time (Chazin & Ledford, 2016). In a free-operant teaching arrangement, the occurrence of a response does not require the presentation of a discriminative stimulus, and the response may be emitted repeatedly (Cooper, Heron, & Heward, 2007). Brief free-operant preference assessments (1–2 minutes) have been shown to yield results comparable to longer sessions while still identifying effective reinforcers. Clay et al. (2021) also found that shorter assessments reduced the occurrence of problem behavior compared to extended sessions. Roane et al. (2009) examined top-ranked items during 30-minute free-operant preference assessments with nine individuals. The results indicated that the first-engaged item and the item receiving the most responding in the first 5 minutes predicted the overall top-ranked item in 55% and 62% of sessions, respectively. Engagement with the initially top-ranked item decreased across subsequent sessions for most participants, highlighting considerations for brief versus extended preference assessments.

Paired-Stimulus vs. Single-Stimulus. Identifying effective reinforcers is a critical first step in designing behavior intervention plans for individuals with developmental disabilities. Two widely used direct assessment methods for identifying preferences are the Paired-Stimulus (PS) method (Fisher et al., 1992) and the Single-Stimulus (SS) method (Pace et al., 1985). Both methods aim to determine which stimuli are likely to function as reinforcers, yet they differ significantly in structure, complexity, and predictive validity. In the Paired-Stimulus method, individuals are presented with two stimuli at a time and asked to make a choice. This process continues until all possible combinations have been presented, allowing for the creation of a preference hierarchy. The PS method has been shown to produce clear and reliable preference rankings and is often considered more accurate in predicting reinforcer effectiveness. However, this method can be time-consuming, especially with larger stimulus sets, and requires that individuals consistently make choices between two items (DeLeon & Iwata, 1996). In contrast, the Single-Stimulus method presents one stimulus at a time, and the observer records whether the individual engages with the item. This method is easier to administer, particularly for individuals with limited choice-making skills or lower levels of cognitive functioning. While SS assessments are efficient and require fewer demands on the participant, they may produce less differentiated results and have shown mixed results in predicting reinforcer strength, especially when compared to the PS method. Studies comparing the two methods (e.g., Kelly et al., 2014) have demonstrated that while both can identify items that function as reinforcers, the PS method is generally more effective in discriminating between high- and low-preference stimuli.

Furthermore, in cases involving social reinforcers such as attention, the PS method may better predict which forms of attention are most likely to maintain or increase desired behaviors. Given these findings, it is essential for practitioners and researchers to consider the context, individual abilities, and time constraints when selecting a preference assessment method. The choice of method can directly impact the success of reinforcement-based interventions, particularly in educational and clinical settings involving individuals with autism spectrum disorder (ASD) and related developmental disabilities. Comparative Analysis of Multiple Preference Assessment Methods Verden and Roscoe (2017) evaluated the consistency and predictive validity of four preference assessment methods—paired-stimulus (PS), multiple-stimulus without replacement (MSWO), free-operant (FO), and response-restriction (RR)—for six individuals (five with ASD and one with traumatic brain injury). The study also assessed whether initially identified high-preference stimuli remained effective reinforcers when compared to immediately identified high-preference stimuli. During PS assessments, stimuli were presented in randomized pairs. In the MSWO assessment, participants selected one item from an array of seven, and chosen items were removed after each trial. The FO assessment allowed unrestricted access to all seven items during a five-minute session. In the RR

assessment, participants were presented with all items, but items associated with high interaction were systematically removed across trials.
Two reinforcer assessments were conducted: a concurrent-operant task (three tasks linked to different stimuli) and a single-operant task (one task paired with one stimulus). Dependent measures included selection (physical contact or requesting) and engagement (e.g., wearing headphones for leisure items), with interobserver agreement data collected via video recordings.

Results showed that PS, MSWO, and RR methods generated clear preference hierarchies and were associated with stronger rank correlations, while the FO method yielded weaker hierarchies but lower rates of problem behavior. Furthermore, reinforcer assessments showed that both initially and immediately identified high-preference items were equally effective in promoting engagement, suggesting that changes in preference over time may not significantly affect reinforcer effectiveness in the short term. The researchers concluded that PS and MSWO methods are effective in identifying reinforcers for students with ASD and should be used in classroom settings, despite their time demands. FO assessments may be better suited when minimizing problem behavior is a priority. Teachers were advised to update preference assessments regularly, particularly when student responsiveness declines or problem behaviors increase.

Students with Autism Spectrum Disorder

Recent statistics on children with disabilities reveal a rising prevalence of autism spectrum disorder (ASD), highlighting the urgent need for research focused on developing engaging and innovative educational strategies for this expanding demographic. Children with autism spectrum disorder (ASD) experience difficulties with verbal and nonverbal communication, manifested in an inability to make requests and express themselves, with varying levels of difficulty. Typical children use a variety of communication forms and behaviors by the time they reach the age of one year, including using their fingers to indicate what they want. However, children with autism rarely communicate in this way and often remain motionless. Autistic children also have difficulty using gestures to express emotions, with facial expressions frequently inconsistent with their tone of voice and verbal cues (Özyurt & Eliküçük, 2018). Many autistic children cannot develop meaningful speech, meaning they cannot produce sounds that can be combined into meaningful words. These children are often described as having functional mutism and have so many communication problems that their attempts at nonverbal communication are sometimes nonexistent (Hambly & Fombonne, 2012).

Students with Autism Spectrum Disorder and preference assessments

Preference assessments are commonly utilized to determine the favored items, activities, and foods of children with ASD. Previous studies constitute a rich tapestry of research on preference assessment methods used to identify effective reinforcers for children, particularly those with developmental disabilities and behavioral disorders (DeLeon & Iwata, 1996; Fisher et al., 1992; Carr et al., 2000; Cote et al., 2007). As we delve into the collective insights offered by these studies, we can identify commonalities and differences that define this vital area of research. One area of overarching agreement among studies lies in the effectiveness of preference assessments, particularly the multiple stimuli without substitution (MSWO) method. Researchers, including Carr et al. (2000), Daly et al. (2009), Weaver et al. (2017), and Prudente and Demchak (2021), consistently demonstrate that this method successfully identifies highly preferred stimuli that can significantly enhance reinforcement strategies. The findings replicate the reliability of the MSWO method in practical applications, demonstrating its ability to improve participation and responding rates in children with autism and behavioral disorders.

Moreover, the studies converge on the concept of preference stability over time. Through their rigorous investigations, researchers found that some preferences do tend to stabilize, although this stability is not uniform across all individuals (Carr et al., 2000; Cote et al., 2007; Daly et al., 2009). For instance, while Jill and Todd exhibited relatively consistent preferences, Billy displayed more fluctuation in his results. This variability underscores the complexity of preference assessment in diverse populations, illustrating that while some children may have stable preferences, others experience shifts that warrant close monitoring

Another point of consensus among the studies is the comparative efficacy of different assessment methods, particularly emphasizing the value of direct evaluations. Cote et al. highlighted that direct assessments often yield more potent reinforcers than those identified through teacher reports, reinforcing the argument that direct methods are not only effective but may also provide deeper insight into a child's preferences. This aligns with similar findings from Kelly et al., which suggest that utilizing direct assessments can lead to stronger reinforcement outcomes. However, despite these agreements, the studies are not without their disagreements. Methodological

efficacy varies, reflecting the differing contexts in which the assessments were conducted. For example, while some studies, such as those by Verriden & Roscoe, found that all assessment methods could yield effective results, they also noted that the free-operant method facilitated lower problem behaviors. This suggests a potential trade-off, where clarity in preference identification may come at the cost of behavioral challenges.

The studies illustrate that the effectiveness of preference assessments can vary considerably based on individual characteristics. Cote et al. (2007) reported inconsistencies across preference assessments for certain children, while others demonstrated relatively stable preferences, highlighting the need for customized approaches tailored to unique needs and behavioral profiles. Similarly, the context in which assessments are conducted plays a critical role. Carr, Nicolson, and Higbee (2000) emphasized the importance of conducting assessments in "naturalistic" settings. However, closer examination revealed that not all assessments took place in environments fully reflective of real-world contexts. This discrepancy raises questions about the generalizability of the findings to actual classroom settings and underscores the importance of implementing interventions in authentic environments to optimize outcomes for children (Daly, Martens, & Kelly, 2009; Weaver, Hughes, & Smith, 2017).

Finally, recommendations regarding preferred assessment strategies represent an important area of discussion. Some studies, such as Kelly, Daly, and Martinez (2010), advocate for the paired-stimulus method due to its detailed and comprehensive data collection. However, they also acknowledge the considerable time and resource investment required. In contrast, the Multiple-Stimulus Without Replacement (MSWO) method, while potentially less detailed, offers a practical alternative for time-constrained educators and practitioners (Carr, Nicolson, & Higbee, 2000; Daly, Martens, & Kelly, 2009; Weaver, Hughes, & Smith, 2017). Collectively, these studies underscore the essential role of preference assessments in identifying effective reinforcers for children with developmental disabilities and behavioral disorders. They present a nuanced picture, highlighting both the consistencies and discrepancies that characterize this field. The convergence on the importance of assessment methods is balanced by individual variability and contextual complexities, supporting the use of tailored approaches that consider the unique needs of each child. Moving forward, ongoing research is essential to refine assessment methodologies and enhance their applicability across diverse educational and clinical settings, ultimately improving outcomes for children facing behavioral challenges (Cote, Thompson, Hanley, & McCarrah, 2007; Prudente & Demchak, 2021).

Research Purpose and Questions

The researcher used two electronic databases (PsycINFO and ERIC) to ascertain whether there was any previous quantitative study that measured teachers' perceptions of preference assessments. Through the research, it was not found that there was any quantitative study that measured the extent of teachers' knowledge and use of preference assessments, and the extent of their impact on determining the most effective reinforcers for students with ASD, where all the research focused on the experimental aspect and the use of the single-subject design. Therefore, there is a great need for this type of research to identify the views of caregivers of students with autism spectrum disorder regarding the best and most effective types of preference assessment. This research aims to answer the following two questions:

1. What preference assessments are most influential in determining the highest reinforcers for students with language delays, from their teachers' perspective?
2. Is a statistically significant relationship between teachers' responses and demographic information (gender, academic level, number of students)?

Design and Methodology

The study population consisted of 112 caregivers of students with autism spectrum disorder (ASD) in Alabama, USA (inclusion and Special classroom). The total number of caregivers was 80 females and 32 males. The researcher designed a questionnaire based on the research objectives and hypotheses. It was validated through peer review by four expert referees (university professors and experts) in the field of autism spectrum disorder, and their comments and modifications were taken into consideration. Therefore, some items were rewarded or added to form the final version of the questionnaire.

Validity and Reliability. The study used Cronbach's alpha coefficient to verify the reliability of the results. The questionnaire's reliability level was found to be (0.91), which is an acceptable level. This questionnaire provides quantitative and qualitative insights into teachers' perceptions of the most influential preference assessments, particularly in the context of students with autism spectrum disorder. The researcher also used Cronbach's alpha coefficient to measure the internal consistency and reliability of the questionnaire. Participants were asked to

indicate their level of agreement with each item, and they were rated on a five-point Likert scale: ineffective = 1, slightly effective = 2, neutral = 3, effective = 4, and strongly effective = 5.

Participants. Ninety-four elementary school caregivers participated in the study, and the remainder were middle school caregivers. The caregivers represented a wide range of experience, including 25% with 5 or fewer years of experience, 30% with 6 to 10 years of experience, 25% with 11 to 20 years of experience, and 20% with 21 or more years of experience. The caregivers represented a wide range of teachers and behavior analysts, including 35% English teachers, 25% mathematics teachers, 20% general studies teachers, 10% computing teachers, and 10% behavior analysts. Some teachers taught more than one subject at their schools (Table).

Table 1: Demographic distribution of the research sample.

Number of participants: 112		**N**	**%**
Gender	Male	32	28.6%
	Female	80	71.4%
Grade Taught	Elementary	94	83.9%
	Middle	18	16.1%
Years of Experience	5 or few years of experience	28	25%
	6 to 10 years of experience	32	28.6%
	11 to 20 years of experience	28	25%
	21 or more years of experience	24	21.4%
Type of work	Behavior analysts	10	8.9%
	English teachers	35	31.3%
	Mathematics teachers	25	22.3%
	General studies teachers	20	17.9%
	Computing teachers	10	8.9%
Organizational Affiliation	General education schools	7	6.3%
	Special education schools	45	40.2%
	Autism organizations	60	53.6%

Procedures. Caregivers of students with autism spectrum disorder were invited to complete the survey via personal emails and social media accounts. The survey was conducted in English, and participants had two weeks to complete it during the first semester of 2022. In the invitation letter, the researcher stated the purpose of the survey, emphasized that participation was voluntary, and informed participants that they could contact the researcher via email with any inquiries or questions. It was also noted that the survey would take approximately 10 minutes to complete. One of the demographic questions asked participants whether they used any differential assessments before selecting the most effective preference assessments. If they answered yes, they were allowed to proceed to Part 2. If they answered no, they were unable to complete the survey. The researcher did not include any information from participants who indicated that they did not use differential assessments with students with autism spectrum disorder. The second part of the questionnaire included five methods for assessing preferences adopted in applied behavior analysis, the most common of which are: Free Operant Observation, Single Stimulus, paired stimulus, multiple stimulus with replacement, and multiple stimulus without placement (Verriden & Roscoe, 2016). The researcher used a five-point Likert scale to measure the level of influence of each of these methods (very influential, influential, somewhat influential, neutral, and not influential).

Data Analysis. The researcher employed SPSS to analyze the data, calculating frequencies, means, standard deviations, and correlation coefficients to examine the relationships among caregiver perceptions of preferences assessment, gender, experience, and type of service. Preliminary analyses involved assessing the assumptions of normality and homogeneity of variance. Normal distribution was evaluated using normal probability plots, box

plots, and measures of skewness and kurtosis. Levene's test was conducted to assess the equality of variances across groups. To differentiate between statistical and practical significance, effect sizes were reported, with statistical significance determined by a p-value threshold of less than 0.05 (Trigwell, K. 2002).

Results

The researcher analyzed the data to examine teachers' perceptions of the preference assessment tools used to identify reinforcers for students with autism spectrum disorder, and to investigate the relationship between these perceptions and certain demographic variables such as gender, years of experience, and type of service (grade level). Initial analyses included calculating frequencies, means, and standard deviations. Assumptions of normal distribution and homogeneity of variance were examined using appropriate statistical tests. In addition, correlation coefficients were calculated between teachers' ratings and demographic variables to identify potential relationships, clarifying the difference between statistical significance and practical value.

Question 1: What are the most influential preference assessments in identifying the highest reinforcers for students with language delays, according to their teachers?

The results of the analysis of the teachers' average ratings showed that the Free Operant Observation tool was the most influential, with an average rating of 2.67 out of 3, indicating that it is considered the most reliable tool for identifying effective reinforcers. Other tools, such as Multiple Stimulus Without Replacement (MSWO) and Paired Stimulus, received lower ratings, but are still considered somewhat effective. The Single Stimulus tool ranked last, with an average rating of 1.35, indicating that it is less effective, according to teachers, in the contexts in which they work.

Table 1: Average rating of preference assessment methods from teachers' perspectives

Preference Assessment Methods	Mean	Arrangement
Free Operant Observation	.67	1
Multiple Stimulus Without Replacement (MSWO)	.94	2
Paired Stimulus	.58	3
Multiple Stimulus with Replacement (MSW)	.50	4
Single Stimulus	.35	5

Question 2: Is there a statistically significant relationship between teachers' responses and demographic information (gender, academic level, number of students/years of experience)?

Although a precise statistical analysis such as the Kruskal-Wallis test could not be conducted in this report, general trends in teachers' evaluations were analyzed based on their demographic data, particularly using the Free Operant Observation tool as a reference:

In terms of gender, female teachers showed higher ratings for the Free Operant tool than male teachers.

Table 2: Mean Free Operant Rating by Gender

Gender	Mean
Male	2.50
Female	2.73

In terms of grade level, elementary school teachers showed a greater preference for using Free Operant than middle school teachers.

Table 3: Mean Free Operant rating by academic level

Gender	Mean
Elementary	2.73
Middle	2.73

Regarding years of experience, no significant differences were found, indicating that the tool's effectiveness does not vary significantly with increasing professional experience.

Table 4: Mean Free Operant Rating by Years of Experience

Years of Experience	Mean
Five years and less	2.67
From six to 10 years	2.60

The data indicates that the Free Operant tool received the highest average rating (4.45), indicating that teachers prefer this type of preference assessment more than others when working with students with language delays. The MSWO tool followed with an average rating of 4.17, then Paired Stimulus (4.07), while Single Stimulus received the lowest average rating of 3.67, which may indicate that it is less reliable or effective from the teachers' perspective. The standard deviation values show slight variation in teachers' responses, reflecting general agreement on the preference order of the tools, with some individual differences in opinion.

Table 5: Means and standard deviations of preference assessment

Preference Assessment	Mean	Standard Deviation
Free Operant	4.45	0.83
MSWO	4.17	0.99
Paired Stimulus	4.07	1.02
MSW	3.86	1.02
Single Stimulus	3.67	1.08

The data distribution was examined using three methods: Normal Q-Q Plots, Box Plots, and skewness and kurtosis criteria. The results showed that most of the data were within the acceptable limits of normal distribution (skewness between -2 and +2, and kurtosis between -7 and +7), with some minor deviations in specific variables, and all Shapiro-Wilk values greater than 0.05. Skewness was negative for all instruments, indicating a slight bias toward higher values (i.e., positive evaluations), while kurtosis was also negative, indicating that the distribution is less steep than the normal (leptokurtic) distribution. Overall, these results support the use of parametric statistical tests in subsequent analyses, such as the t-test, analysis of variance, and correlation.

Table 6: Measures of skewness, kurtosis, and Shapiro-Wilk test

Preference Assessment	Skewness	Kurtosis	P value for Shapiro's test
Free Operant	-0.60	-0.22	0.052
MSWO	-0.58	-0.30	0.095
Paired Stimulus	-0.45	-0.57	0.069
MSW	-0.28	-0.91	0.144
Single Stimulus	-0.26	-1.11	0.223

Note: The distribution is considered normal if the p-value in the Shapiro-Wilk test is > 0.05.

The results showed a moderate statistically significant relationship between the number of years of experience and teachers' evaluation of the Free Operant tool, indicating that more experienced teachers preferred this tool more than others. Gender and educational level were not statistically significantly associated with evaluation, although gender showed a small effect size, which may be of limited practical significance. The Point-Biserial coefficient was used for the relationship between sex and stage type (since it was binary), Pearson's coefficient for years of experience, and the effect size was interpreted according to Cohen's criteria:

Table 7: Correlation coefficients between Free Operant rating and demographic variables

Variable	Correlation coefficient (r)	Probability value (p)	Statistical significance ($p < 0.05$)	Effect Size (Cohen's)
Gender (male/female)	0.145	0.180	No	0.1
Years of experience	0.315	0.023	Yes	0.3
Type of stage	0.198	0.095	No	0.1

The results indicate a statistically significant correlation between the MSWO assessment and the number of years of experience, with more experienced teachers tending to prefer this tool. For gender and educational level, the results did not show statistical significance, although their coefficients indicate a relatively small effect size. This may reflect a slight bias in preference that falls short of a significant practical effect.

Table 8: Correlation coefficients between MSWO rating and demographic variables

Variable	Correlation coefficient (r)	Probability value (p)	Statistical significance ($p < 0.05$)	Effect Size (Cohen's)
Gender (male/female)	0.162	0.140	No	0.1
Years of experience	0.289	0.034	Yes	0.3
Type of stage	0.178	0.101	No	0.1

Regarding the remaining tools, the results did not show any statistically significant differences between gender, type of educational level, or years of experience in the assessment of the Paired Stimulus, MSW, or Single Stimulus tools, although there were weak correlations.

Table 9: Correlation coefficients between Paired Stimulus rating and demographic variables

Variable	Correlation coefficient (r)	Probability value (p)	Statistical significance ($p < 0.05$)	Effect Size (Cohen's)
Gender (male/female)	0.162	0.140	No	0.1
Years of experience	0.289	0.034	Yes	0.3
Type of stage	0.178	0.101	No	0.1

Table 10: Correlation coefficients between MSW rating and demographic variables

Variable	Correlation coefficient (r)	Probability value (p)	Statistical significance ($p < 0.05$)	Effect Size (Cohen's)
Gender (male/female)	0.137	0.210	No	0.1
Years of experience	0.201	0.077	No	0.1
Type of stage	0.165	0.130	No	0.1

Table 11: Correlation coefficients between Single Stimulus rating and demographic variables

Variable	Correlation coefficient (r)	Probability value (p)	Statistical significance ($p < 0.05$)	Effect Size (Cohen's)
Gender (male/female)	0.110	0.289	No	0.1
Years of experience	0.218	0.066	No	0.1
Type of stage	0.149	0.165	No	0.1

Discussion and conclusion

This study aimed to determine teachers' evaluations of reinforcement preference tools used with students with language delays and to analyze the extent to which demographic variables (gender, years of experience, and type of educational stage) influenced these evaluations. Data were analyzed using means and standard deviations, normal distribution analysis, Levene's test for homogeneity of variance, and correlation coefficients to determine relationships between variables. The Free Operant method received the highest average rating from teachers, followed by the MSWO, then the Paired Stimulus. This suggests that caregivers prefer to give students free access to an enriched environment that allows them to explore preferences through more realistic or multiple-choice situations, consistent with the nature of interventions based on applied behavior analysis (ABA).

All five assessment tools showed a normal distribution based on the results of the Shapiro-Wilk test, supporting the validity of the study's use of parametric statistical analyses. Levene's test results revealed no significant differences in variance between males and females for any of the assessment tools (all p-values > 0.05), indicating equal variance across genders. Years of experience was the only variable that showed statistically significant relationships with the assessment of some tools (e.g., MSWO and Free Operant), suggesting that teaching experience plays an important role in shaping teachers' decisions about appropriate preference tools. Gender and educational level were

not statistically significantly associated with any of the assessment tools, although some weak trends emerged that may warrant future study with larger samples.

These findings provide important practical evidence to guide caregivers and program planners in selecting the most appropriate preference tools when working with students with language delays. They also highlight the importance of enhancing teacher training, particularly early in their careers, in the use of effective tools such as the MSWO and Free Operant.

Limitations and suggestions for future research

One of the most significant limitations of this study is the sample size. The study was limited to a small number of caregivers, which may affect the generalizability of the findings. The results may be more reliable with a larger, more diverse sample. Furthermore, the data were based on questionnaires completed by the caregivers themselves, which may introduce bias based on their expectations or prior experiences. The study focused solely on students with autism spectrum disorder (ASD), without considering other categories of disabilities or learning difficulties, limiting the generalizability of the findings to broader student groups. The researcher limited the study to three demographic variables (gender, years of experience, and type of educational level) and did not include other potentially influential variables such as specialized training, class size, or cultural background. Neither experimental nor a quasi-experimental design was used to verify causality between the variables, limiting the findings to correlations rather than causal relationships.

The researcher recommends conducting future studies on a larger sample that includes caregivers from different educational districts and at various levels. The researcher also suggests conducting longitudinal follow-up studies to analyze how teachers' preferences or proficiency in using preference tools change over time, especially with experience and training. In addition, the impact of specialized training programs in assessing preferences on teachers' decisions and use of various tools should be examined. Other methods, such as direct observation or qualitative interviews, should be incorporated to gain a deeper understanding of teachers' preferences and their underlying reasons. Finally, the researcher highlights the importance of expanding the target population by conducting the study with caregivers of students from different groups, such as those with intellectual disabilities or attention deficit hyperactivity disorder (ADHD), to compare preferences across groups.

Practical implications

The results of this study point to several important practical implications that could contribute to improving the educational environment and providing effective support for students with autism spectrum disorder. The most important of these implications is the selection of the most effective assessment tools for students. The study results indicated that the MSWO (Multiple Stimulus Without Replacement) and Free Operant tools were the most effective, suggesting they could be recommended as primary standard tools for assessing preferences in students with autism spectrum disorder. Furthermore, this helps reduce time and effort by identifying the most effective and acceptable tools for teachers, making it easier for schools and institutions to avoid ineffective tools and direct resources toward developing a learning environment rich in appropriate reinforcers. This study also highlights the importance of training courses, as professional development can enhance service providers' understanding of the differences between preference assessment tools and how to choose the most appropriate tool for the student’s needs and educational situation.

References

Ackerlund Brandt, J., Dozier, C., Juanico, J., Laudont, C., & Mick, B. (2015). The value of choice as a reinforcer for typically developing children. *Journal of Applied Behavior Analysis, 48*(2), 344–362. https://doi.org/10.1002/jaba.199

Carr, J. E., Nicolson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis, 33*(3), 353-357. https://doi.org/10.1901/jaba.2000.33-353

Chazin, K. T., & Ledford, J. R. (2016). *Free operant observation. In Evidence-based instructional practices for young children with autism and other disabilities*. Retrieved from http://ebip.vkcsites.org/free-operant

Cote, L. R., Thompson, R. H., Hanley, G. P., & McCarrah, M. A. (2007). Comparing direct and indirect methods for identifying potent reinforcers for young children. *Journal of Applied Behavior Analysis, 40*(2), 275-288. https://doi.org/10.1901/jaba.2007.40-275

Cohen-Almeida, D., Graff, R. B., & Ahearn, W. H. (2000). A comparison of verbal and tangible stimulus preference assessments. *Journal of Applied Behavior Analysis, 33*(3), 329–334. https://doi.org/10.1901/jaba.2000.33-329

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Pearson Prentice Hall. ISBN: 978-0131421134

Courtemanche, A. B., Turner, L. B., Molteni, J. D., & Groskreutz, N. C. (2020). Scaling up behavioral skills training: Effectiveness of large-scale and multiskilled trainings. *Behavior Analysis in Practice, 14*(1), 36–50. https://doi.org/10.1007/s40617-020-00480-5

Clay, C. J., Schmitz, B. A., Clohisy, A. M., Haider, A. F., & Kahng, S. W. (2021). Evaluation of Free-operant Preference Assessment: Outcomes of Varying Session Duration and Problem Behavior. *Behavior Modification, 45*(6), 962–987. https://doi.org/10.1177/0145445520925429.

Daly, E. J., Martens, B. K., & Kelly, M. M. (2009). Preference assessments: Identifying potential reinforcers for individuals with developmental disabilities. *Behavior Analysis in Practice, 2*(1), 12-21. https://doi.org/10.1007/BF03391761

DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*(4), 519-533. https://doi.org/10.1901/jaba.1996.29-519

DeLeon, I. G., Iwata, B. A., & Roscoe, E. M. (1997). Evaluation of the effects of the order of presentation in multiple-stimulus preference assessments. *Journal of Applied Behavior Analysis, 30*(3), 447-456. https://doi.org/10.1901/jaba.1997.30-447

Fisher, W. W., Piazza, C. C., Bowman, L. G., Hagopian, L. P., Owens, J. C., & Slevin, I. (1992). A comparison of two approaches for identifying reinforcers for persons with severe and profound disabilities. *Journal of Applied Behavior Analysis, 25*(2), 491-498. https://doi.org/10.1901/jaba.1992.25-491

Fisher, W. W., Piazza, C. C., Hagopian, L. P., Bowman, L. G., Owens, J. C., & Slevin, I. (1996). Using a choice assessment to predict reinforcer effectiveness. *Journal of Applied Behavior Analysis, 29*(1), 1-9. https://doi.org/10.1901/jaba.1996.29-1

Graff, R. B., & Karsten, A. M. (2012). Assessing preferences of individuals with developmental disabilities: A survey of current practices. *Behavior Analysis in Practice, 5*(2), 37–48. https://doi.org/10.1007/BF03391822

Hagopian, L. P., Long, E. S., & Rush, K. S. (2004). Preference assessment procedures for individuals with developmental disabilities. *Behavior Modification, 28*(5), 668-677. https://doi.org/10.1177/0145445503259836

Hambly, C., & Fombonne, E. (2012). The impact of bilingual environments on language development in children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(7), 1342–1352. https://doi.org/10.1007/s10803-011-1365-z

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: A review. *Journal of Applied Behavior Analysis, 36*(2), 147-185. https://doi.org/10.1901/jaba.2003.36-147

Kang, S., O'Reilly, M., Lancioni, G., Falcomata, T. S., Sigafoos, J., & Xu, Z. (2013). Comparison of the predictive validity and consistency among preference assessment procedures: A review of the literature. *Research in Developmental Disabilities, 34*(4), 1125-1133. https://doi.org/10.1016/j.ridd.2012.12.022

Keen, D., & Pennell, D. (2010). Evaluating an engagement-based preference assessment for children with autism. Research in Autism Spectrum Disorders, 4(4), 645-652. https://doi.org/10.1016/j.rasd.2009.01.001

Kelly, M. A., Roscoe, E. M., Hanley, G. P., & Schlichenmeyer, K. (2014). Evaluation of assessment methods for identifying social reinforcers. *Journal of Applied Behavior Analysis, 47*(1), 113–135. https://doi.org/10.1002/jaba.107

Kodak, T., Fisher, W. W., Kelley, M. E., & Kisamore, A. (2009). Comparing preference assessments: Selection- versus duration-based preference assessment procedures. *Research in Developmental Disabilities, 30*(5), 1068-1077. https://doi.org/10.1016/j.ridd.2009.03.002

Lill, J. D., Shriver, M. D., & Allen, K. D. (2021). Stimulus Preference Assessment Decision-Making System (SPADS): A decision-making model for practitioners. *Behavior Analysis in Practice, 14*(4), 1144–1156. https://doi.org/10.1007/s40617-020-00539-3

Paramore, N. W., & Higbee, T. S. (2005). Evaluation of a brief multiple-stimulus preference assessment with adolescents with emotional and behavioral disorders in an educational setting. *Journal of Applied Behavior Analysis, 38*(4), 611-616. https://doi.org/10.1901/jaba.2005.38-611

Parsons, M. B., Reid, D. H., Reynolds, J., & Bumgarner, M. (1990). Teacher report and direct assessment of preferences for identifying reinforcers for young children. *Journal of Applied Behavior Analysis, 23*(2), 135-141. https://doi.org/10.1901/jaba.1990.23-135

Prudente, M., & Demchak, M. (2021). Brief MSWO assessments for students with emotional and behavioral disorders: Identifying effective reinforcers. *Journal of Behavioral Education, 30*(1), 1-18. https://doi.org/10.1007/s10864-020-09380-5

Roane, H. S., Vollmer, T. R., Ringdahl, J. E., & Marcus, B. A. (2009). Predicting preference for items during periods of extended free-operant assessment. *Journal of Applied Behavior Analysis, 42*(3), 677–691. https://doi.org/10.1901/jaba.2009.42-677 .

Weaver, C., Hughes, C., & Smith, D. (2017). Using multiple-stimulus without replacement preference assessments to increase student engagement and performance. *Education and Treatment of Children, 40*(2), 207-225. https://doi.org/10.1353/etc.2017.0010

Wilder, D. A., Ellsworth, C., White, H., & Schock, K. (2003). A comparison of stimulus preference assessment methods in adults with schizophrenia. *Behavioral Interventions, 18*(3), 151–160. https://doi.org/10.1002/bin.128

Verriden, A. L., & Roscoe, E. M. (2016). A comparison of preference-assessment methods. *Journal of Applied Behavior Analysis, 49*(2), 265–285. https://doi.org/10.1002/jaba.302

Özyurt, G., & Eliküçük, Ç. D. (2018). Comparison of language features, autism spectrum symptoms in children diagnosed with autism spectrum disorder, developmental language delay, and healthy controls. *Archives of Neuropsychiatry, 55*(3), 205. https://doi.org/10.5152/npa.2017.17904

Trigwell, K. (2002). Approaches to teaching design subjects: A quantitative analysis. *Art, Design & Communication in Higher Education, 21*, 275–284. https://doi.org/10.1386/adch.21.3.275/1