EFFECTS OF MASTERY CRITERIA ON SKILL MAINTENANCE OF TACTS WITH INDIVIDUALS WITH AUTISM SPECTRUM DISORDER: A REPLICATION

by

VIDISHA GANGIDI

(Under the Direction of Scott Ardoin and Rachel Cagliani)

ABSTRACT

A mastery criterion is selected in skill acquisition to demonstrate mastery of a taught skill. Recently, researchers have evaluated different mastery criterions and the effects it has on maintenance of the taught skill. Previous research suggest that skills taught with a stringent mastery criterion of 100% across three consecutive sessions achieved the best maintenance when assessed in four monthly probes. This study replicates and extends previous research from Longino et al., (2021) to evaluate the maintenance of tacting skills taught using a most-to-least prompting procedure combined with a progressive time delay with two+ children diagnosed with autism spectrum disorder.

INDEX WORDS: mastery criteria, autism spectrum disorder, skill maintenance

EFFECTS OF MASTERY CRITERIA ON SKILL MAINTENANCE OF TACTS WITH INDIVIDUALS WITH AUTISM SPECTRUM DISORDER: A REPLICATION

by

VIDISHA GANGIDI

B.S.F.C.S, University of Georgia, 2019

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2023

EFFECTS OF MASTERY CRITERIA ON SKILL MAINTENANCE OF TACTS WITH INDIVIDUALS WITH AUTISM SPECTRUM DISORDER: A REPLICATION

by

VIDISHA GANGIDI

Major Professor:	Scott Ardoin
	Rachel Cagliani
Committee:	Kristina Gerencser

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2023

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

INTRODUCTION

When programming for clients, practitioners determine a set criterion to be confident that the client has acquired the target behavior (Fienup and Brodsky, 2017). Practitioners test the maintenance of the behavior after a week or a month to ensure that the client has maintained the learned behavior (McDougale et. al., 2019). To determine what kind of mastery criteria would benefit children diagnosed with autism spectrum disorder (ASD), researchers began analyzing criteria to determine the efficacy of specific predetermined mastery criteria. Fuller and Fienup (2018) conducted a preliminary analysis involving three participants with ASD to test the response maintenance using three mastery criteria (50%, 80%, and 90%) on academic skills. The results of this study determined that the participants were more likely to maintain response requirements when taught at a higher performance level. These findings warranted further investigations to determine the most implemented mastery criteria among practitioners and evaluate the effects of different mastery criteria on the skill maintenance of individuals with ASD.

A survey by Richling et al. (2019) determined that 80% of mastery criteria across three sessions are behavior analysts' most reported criteria. To test the survey's results, researchers conducted a further examination using vocal tacting skills and receptive identification, which reassured that an 80% mastery criterion decreased response maintenance during follow-up maintenance compared to 100% mastery criterion. Researchers used least-to-most (LTM) prompting procedures, systematically increasing the intrusiveness of prompts, to teach both operants until the participants' responding reached mastery. In comparison, the 100% mastery criterion set led to higher response maintenance for vocal tacting and receptive identification

skills (Richling et al., 2019). To ensure the efficacy of the Richling study, Pitts and Hoerger (2020) conducted a systematic replication of the study to assess if the study would produce similar results. The partial replication resulted in higher skill maintenance using more stringent (i.e., 100%) mastery criteria (Pitts & Hoerger, 2020).

Similar to the LTM prompting procedures used in the previous articles (i.e., Richling et al., 2019; Pitts & Hoerger, 2020), the majority of mastery criterion studies utilize LTM prompt hierarchy as part of their teaching methods compared to most-to-least prompting (MTL). According to Cengher et al. (2016), LTM is more efficient to implement by following the least intrusive prompt and then moving to a more intrusive prompt if needed. MTL may be more effective because it leads to fewer errors while teaching skills. While previous research studies have implemented LTM prompting procedures, Longino et al. (2021) used MTL prompting procedures to test the effectiveness of three mastery criteria (80%, 90%, and 100%). Using the MTL, the target sets assigned to the higher accuracy criteria (100%) produced higher maintenance rates than the lower accuracy criteria (80%). In addition to varying prompting procedures, most studies include participants, who represent a small population on the spectrum. Research on the efficacy of mastery criteria requires participants with different (find a better word) skill levels across the spectrum.

The purpose of this study was to conduct a systematic replication of Longino et al. (2021) study aimed to evaluate the maintenance of vocal tacts over a course of a month with target sets assigned to three different mastery criteria (80%, 90%, and 100%) using MTL prompting procedure with participants with limited expressive skills levels across the autism spectrum.

METHOD

Participants and Setting

Two participants were recruited to participate in this study. All participants were receiving intensive behavioral analytic intervention in a clinic for 15 hours per week. To be included in the study, participants met the following criteria: (a) able to tact at least 6 non-reinforcing items (i.e., Verbal Behavior Milestones Assessment and Placement Program, VB-MAPP, Level 1 Tact Milestone 3), (b) able to sit and attend to therapist, while engaging in minimal challenging behavior (i.e., combined score of 4 or less on Behavior Problems and Instructional Control on the VB-MAPP Barriers Assessment), and (c) have little to no barriers to reinforcers (i.e., combined score of 4 or less on Weak Motivators and Reinforcer Dependent on the VB-MAPP Barriers Assessment; Sundberg, 2014).

All participants were vocal. Cameron was a 5-year-old boy diagnosed with ASD who displayed strong echoic skills, spontaneous tacting, manding (2–3-word utterances), and intraverbal skills. With a total of 70.5 across milestones, Cameron scored 8 points within Level 1 and Level 2 Tact category on the VB-MAPP (Sundberg, 2014). Scoring in Level 1 and into Level 2 categorized the learner as an early learner with some solid language and learning skills. Zuri was a 6-year-old girl diagnosed with ASD who displayed echoic, tacting, and manding (2-3-word utterances) skills. With a total score of 70.5 across milestones, Zuri scored 7 points within Level 1 and Level 2 (generalized tacting 50 items across three examples) in the Tact category on the VB-MAPP assessment.

Researchers conducted sessions in the client's designated area at the facility's clinic. The area included instructional materials, preferred items, a table, and at least three chairs for the participant, therapist, and secondary data collector. There were minimal items in the area to increase attending towards the therapist and the stimuli and limit distractions during instructional sessions. The therapist embedded the experiment sessions in the participant's general therapeutic program.

Target Behavior and Response Measurement

The target behavior for each participant was accurately tacting by vocally stating the name of the image depicted on the index card when the S^D "What is it?" was presented. Researchers assigned one set of three target stimuli to each mastery criterion examined (i.e., 80%, 90%, and 100%). Each set included three 2-dimensional picture targets with a white background. Researchers selected birds as the stimuli to be used in the study with the likelihood of participants had not received formal teaching to identify various bird names.

To reduce certain threats to internal validity, researchers selected targets from the pre-assessment and used equating procedures based on Cariveau et al., (2022) to create unique target sets based on five standards. These standards include: (a) the number of syllables in each target, (b) the number of syllables in each condition, (c) overlapping sounds, (d) first sounds, and (e) visual properties of the images (i.e., color). Target set 1 included an emu, a toucan, and a heron. Target set 2 included a quail, a seagull, and a blue jay. Target set 3 included a swan, a robin, and a falcon. Researchers then equated the target sets across both participants, see Table 1.

In a session, each target set included a total of 10 trials, which included the presentation of the target stimuli in a 3-3-4 format to counterbalance target presentations across participants.

Table 1: Equating targets across participants

Target Sets	Participant 1	Participant 2
80% Mastery	Emu Toucan Heron	Quail Seagull Blue Jay
90% Mastery	Quail Seagull Blue Jay	Swan Robin Falcon
100% Mastery	Swan Robin Falcon	Emu Toucan Heron

Researchers created three different versions of session data sheets to account for the target presented four times in a session. Based on the participant's response, therapists scored each trial as correct or incorrect during the baseline and maintenance sessions and independent or prompted during the teaching sessions. For baseline and maintenance sessions, a correct response indicated that the participant provided the correct bird name when the target was presented with the S^D ("What is it?"). An incorrect response indicated that the participant provided the incorrect bird name when the target was presented with the S^D. During teaching sessions, therapists marked independent when the participant provided the correct response while prompted indicated that the participant required an error correction for that trial. Researchers scored each session into percentages by dividing the number of independent responses by the total number of

trials and multiplying by 100. If the participant did not respond after the presentation of the S^D, the researchers re-presented the trial and followed procedure based on baseline, teaching, or maintenance conditions. This procedure ensured that nonresponses were not included in the scoring of the dependent variable.

Design and Procedures

Researchers evaluated the effects of three varying mastery criteria on maintaining skills using a multiple-baseline design with an adapted alternating-treatments design. Researchers compared the effects of 80%, 90%, and 100% accuracy across three consecutive sessions on skill maintenance during four weekly follow-up maintenance probes. The current study consisted of four phases including phase 1: pre-assessment and preference assessment, phase 2: baseline, phase 3: teaching, and phase 4: maintenance. Starting from phase 2, all three target sets included 10 discrete trials within a session, which were conducted 1-2 times per day across 1 to 5 days per week depending on participants' attendance.

Pre-Assessment

Researchers selected common targets that participants were unlikely to receive formal outside teaching for the experimental sessions. The case managers and therapists paired with the participants the number of syllabi that all the participants had previously demonstrated the ability to pronounce and respond to differentially. Then, researchers conducted a formal pre-assessment, consisting of 19 images, to gauge if the participants have prior knowledge of the targets. Therapists presented the targets in a VR-3 (variable ratio of 3 demands), meaning placing between 1 to 5 demands before providing reinforcement for both participants. Therapists did not

deliver any consequences for correct or incorrect responses but provided non-contingent reinforcement for 30 s (i.e., "Great job staying in your seat") for the stay in-seat and/or responding correctly to a mastered target across various operants.

Preference Assessment

Researchers conducted a multiple-stimulus without-replacement (MSWO) preference assessment prior to the onset of the experiment for each participant to identify highly preferred edibles and tangibles (DeLeon & Iwata, 1996). The preference assessment included up to seven edible or tangible items identified as preferred items based on caregiver or therapist reports. Before beginning each experimental session, therapists presented the three most highly preferred items and instructed the participant to choose a reinforcer to work for the session. If researcher noted concerns related to satiation or shift in participants preference, the satiated items identified in the MSWO were rotated out with other preferred items during to maintain compliance and responding.

Baseline

During baseline, the therapist presented one card with a target stimulus in front of the participant. The therapist provided the S^D "What is it?" while showing the participant the target stimulus. Therapists did not deliver any consequences for correct or incorrect responses in the baseline. Additionally, therapists did not deliver any prompts during this phase. Therapists implemented VR-3 (variable ratio of 3 responses) schedule of reinforcement, presenting between 1 to 5 trials before access to reinforcement. To reduce the likelihood of challenging behavior and increase attending, the therapist provided 30-s noncontingent access to a tangible from the

preference assessment and praise (i.e., “thank you for looking at the card”) based on the participant’s VR schedule.

Teaching

During this phase, the therapist utilized a MTL prompting procedure with a progressive delay, differential reinforcement for correct responses, and an error correction procedure. Therapists used an MTL prompting procedure combined with a progressive delay with the following steps: Phase 1: 0-s delay to a full vocal prompt, Phase 2: 2-s delay to a full vocal prompt, Phase 3: 4-s delay to a full vocal delay, and Phase 4: no prompt, just independent responses. Therapists provided error correction in Phase 4 for an incorrect response. Unlike the Longino study, the prompt hierarchy did not include a partial vocal prompt. This modification was selected because to promote errorless teaching and consider the participants’ extensive tact repertoire. Researchers increased prompt steps after participants met the mastery criteria assigned to the target set for two consecutive sessions. For instance, for the target set assigned to the 90% accuracy criterion, the therapist would decrease the intrusiveness of the prompt when the participant achieved two consecutive sessions at 90% correct (prompted or independent). If the participant engaged in three incorrect responses within a 10-trial session, the therapist was to complete the session at the current prompt step and increase the intrusiveness of the prompt for the following session.

During teaching sessions, therapists implemented a FR-1 (fixed ratio of 1 response) schedule of reinforcement for independent/correct prompted responses. Based on the prompting phase, therapist provided 30-s access to reinforcers identified in the MSWO for

independent/correct prompted responses. If an error occurred, the therapist re-presented the trial with an echoic prompt, "Say ….". Finally, the therapist terminated the trial and moved on to the next trial regardless of whether the participant engaged in the correct response following the error correction procedure.

Maintenance Probes

Researchers began the maintenance phase after the participant met the criteria assigned to the target set. Researchers conducted weekly probes under baseline conditions for four weeks following when the participant mastered the target set. Therapists ran target sets using a VR-3 schedule of reinforcement with no consequences provided for correct or incorrect responses. Therapists provided noncontingent reinforcement to ensure compliance by providing praise (i.e., "thank you for responding") and tangible reinforcers identified from the MSWO for 30-s. The participants mastered each target set for each condition at different points during the study.

Interobserver Agreement and Treatment Integrity

Researchers collected interobserver agreement (IOA) data in-person and video recording for 98% of all sessions throughout the study. Researchers calculated IOA for 100% of sessions for Cameron and 96% of sessions for Zuri. Researchers calculated the percentage by trial-by-trial equation, dividing total agreements over total agreements and disagreements and then multiplying by 100. Average agreement was 97% (range 70%-100%) for Cameron and 99.6% (range 90%-100%) for Zuri.

During baseline and treatment phases, a second observer collected data to measure treatment integrity. The observer used a checklist outlined with the procedures and scored each

component as Yes for correct and No for incorrect. The components included (a) conducting a free operant assessment with preferences identified from the MSWO; (b) ensuring the participant is attending; (c) delivering the correct SD; (d) following the correct procedure based on the independent and prompt responses during that particular phase of teaching; (e) implementing the consequence based on participant responding. Researchers calculated treatment integrity by dividing the number of correct components over the total number of components and then multiplying by 100. Data was collected for 82% (range 77%-88%) of all sessions throughout the study. Average treatment integrity scores were 98% (range 98%-100%) for Cameron and 96% (range 94%-100%) for Zuri.

Results

Figure 1 displays the results for three participants: Cameron (top panel), Zuri (middle panel), and Ken (bottom panel). During the baseline sessions, Cameron provided zero correct independent responses to all three target sets. During the teaching procedures, Cameron's responding fluctuated across all three experimental conditions. Cameron met criteria for the target set assigned 80% mastery across three sessions followed 100% mastery target set and then 90% mastery set. Cameron demonstrated mastery for the 80%, 90%, and 100% criterion target set after 32, 36, and 40 respectively. According to the Week 1 maintenance probes, Cameron demonstrated maintenance for 80% mastery set at the associated mastery level. At Week 2, Cameron demonstrated 70% correct responding for target set assigned to 100% and 100% correct responding for target set assigned to 90%. At Week 3, Cameron demonstrated

maintenance for 80% mastery target set. Average responding across the weekly was 87% for 80% mastery target set, 97% for 90% mastery target set, and 90% for 100% mastery target set.

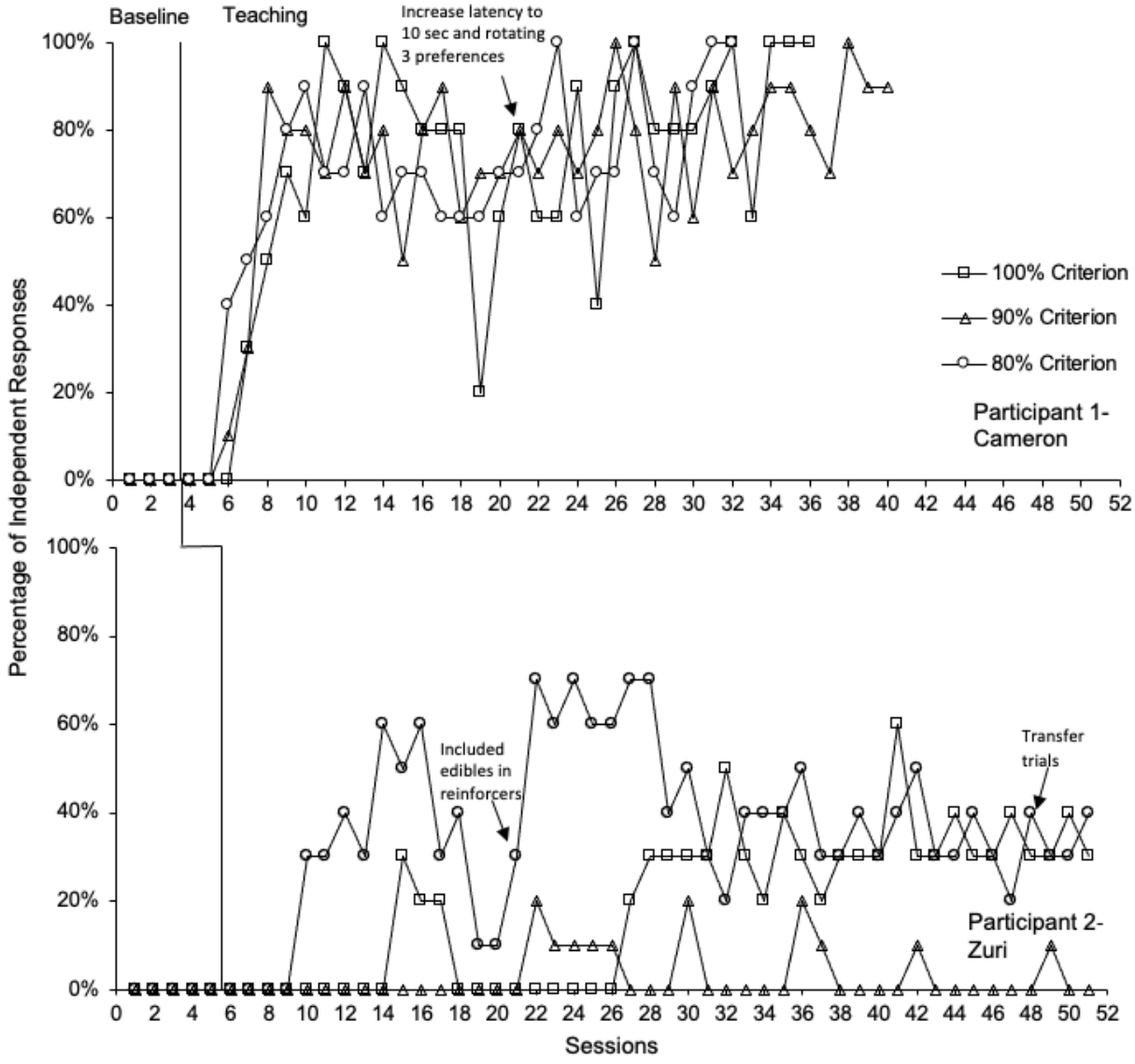

Figure 1: Percentages of independent responses across all target sets

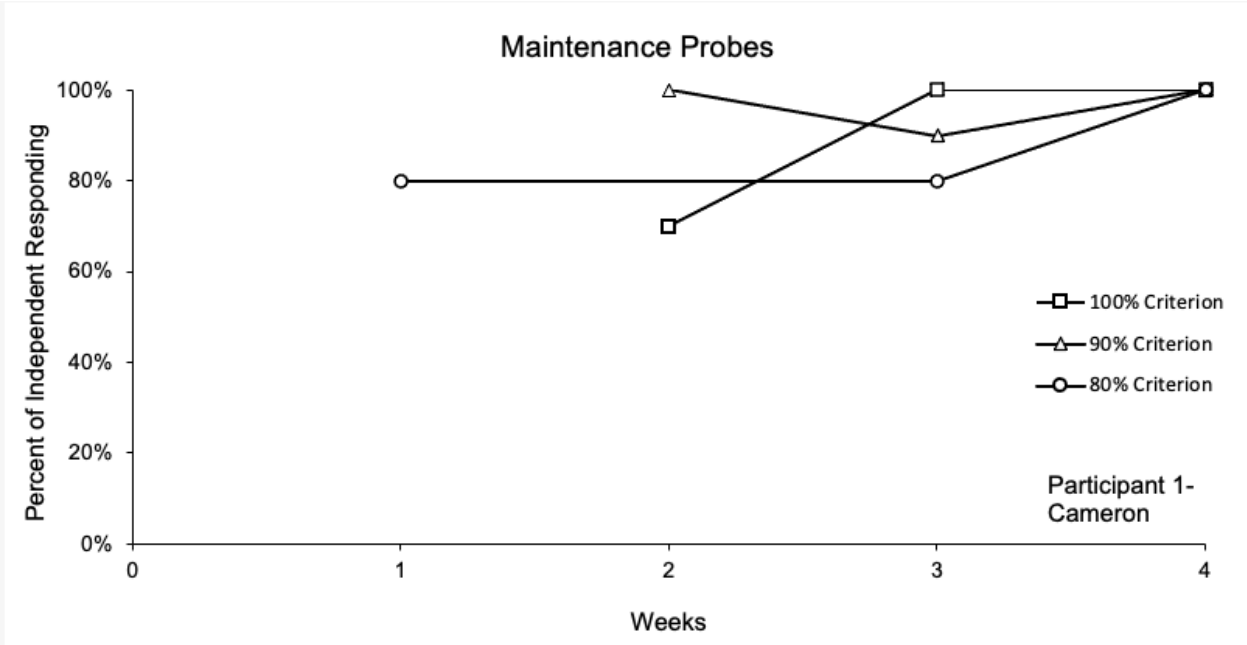

Figure 2: Percentages of independent responses across weekly maintenance probes

Maintenance probes for week 1 for 90% and 100% mastery sets and week 2 for 80% are missing data points due to attendance.

Zuri (bottom panel) responded with 0% correct responding across all three target sets during baseline. Zuri demonstrated low levels of independent responding in the beginning of the teaching sessions. Zuri's independent responding increased in the 80% mastery condition with the introduction of the 4-sec time delay before the presentation of the full vocal prompt. Zuri began responding to certain stimuli in the 100% mastery condition at session 15 but continued to have low levels of responding. Zuri periodically demonstrated independent responding for the 90% mastery conditions during sessions 22-26, 30, 36-37, 42, and 49. Responding remained

stagnant for the duration of the study. Researchers discontinued the study without assessing maintenance probes.

Discussion

This study was a replication that examined the effects on three mastery criteria (i.e., 80%, 90%, and 100%) on the maintenance of tact skills taught using a MTL prompting procedure with a progressive delay prompt hierarchy. Based on the procedures from the Longino et al. (2021), Cameron had higher levels of maintenance for 90% and 100% mastery target sets compared to 80% mastery. Fuller and Fienup, (2018) and Richling et al., (2019) indicated the efficacy of using 90% mastery criteria using LTM prompting hierarchy as a way of promoting independent responding and creating efficient programming unlike less stringent mastery criterion. Findings also follow results from a systematic review conducted by Wong et al. (2022) that suggested that higher levels of skill maintenance can be attributed to 90%-100% criterion levels. For Cameron, a less stringent mastery criteria produced high levels of skill maintenance due to using MTL prompting procedure to promote errorless teaching, implementing a time delay to increase independent responding, or the higher number of expressive skills Cameron displaying prior to the study.

Several limitations require further discussion when considering different variables that impacted the results. First, researchers cannot make a conclusive statement for this study as Zuri did not master the targets to assess response maintenance for all three target sets. While researchers considered modifications to support learning, researchers could not assess the

modifications' efficacy due to caregivers ending clinical services before the completion of the study.

Secondly, the rate of responding between both participants differed based on skill level, exposure to previous 2D tact programming, and success with other teaching styles. Zuri's regular clinic programming involved an adapted version of Project Impact, a naturalistic development behavioral intervention (NDBI). This programming supports teaching operants in the natural environment, while Cameron's programming included intensive discrete trial teaching (DTT) procedures to support skill acquisition. Implementing a DTT-style procedure with Zuri possibly led to a lower target acquisition rate during the teaching phase. Researchers also implemented transfer trials, providing an independent opportunity after the prompt to increase independent responding. Since transfer trials fall under DTT strategies, researchers were unsuccessful in increasing independent responses. An additional hypothesis to consider based on parental reports and clinical responding is defective conditional discrimination (difficulties discriminating between similar stimuli) based on the VB-MAPP (Sundberg, 2014).

Third, researchers changed the reinforcers available for each participant due to satiation. Therapists rotated 3 out of 5 reinforcers during teaching sessions to maintain motivation and consistent responding for Cameron. Researchers also included edibles (freeze-dried strawberries) to increase Zuri's motivation. The reinforcers from the MSWO were deemed not potent enough to capture an EO (establishing operation). Future studies may include a reinforcer testing procedure to assess if preferred items identified from the MSWO function as reinforcers.

Lastly, the differences in characteristics and skills between Cameron and Zuri led to variable responding to target sets. Cameron scored 8.5 points in the Tact category of the VB-MAPP (including generalized tacts of 50 items across three exemplars and tacts of 10 actions). In comparison, Zuri scored 7 points in the Tact category—the lower score in the tact category (generalized tacts of 50 items across three exemplars) but scored 1 point in the Defective Articulation under Barriers section of the VB-MAPP (some difficulty pronouncing certain words, but usually can be understood and articulation improves; Sundberg, 2014).

Although the study produced limited results on mastery criteria impacting skill maintenance, further research warrants investigating the impact of other variables, such as teaching procedures and barriers to learning, on children reaching mastery. Implementing these procedures with additional participants would provide more insight into the impact of mastery criteria on skill maintenance across the autism spectrum.

References

Cariveau, T., & Moseley, T. K. (2022). Equating and assigning targets in the adapted alternating treatment design: Review of special education journals. *Remedial and Special Education*, 43(1), 58-71.

Cengher, M., Budd, A., Farrell, N., & Fienup, D. M. (2017). A review of prompt-fading procedures: implications for effective and efficient skill acquistion. *Journal of Developmental Physical Disability*. 30, 155-173.

DeLeon, I., & Iwata, B. (n.d.). Evaluation of a multiple-stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis*, 519-533.

Fienup, D.M. & Brodsky, J. (2017). Effects of mastery criterion on the emergence of derived equivalence relations. *Journal of Applied Behavior Analysis*. 50, 843-848.

Fuller, J., & Fienup, D. (2018). A preliminary analysis of mastery criterion level: effects on response maintenance. *Behavior Analysis Practice*, 1-8.

Longino, E., Richling, S., McDougale, C. B., & Palmier, J. M. (2021). The effects of mastery criteria on maintenance: a replication with most-to-least prompting. *Behavior Analysis in Practice*, 15, 397-405.

McDougale, C., Richling, S., Longino, E., & O'Rourke, S. (2020). Mastery criteria and maintenance: a descriptive analysis of applied research procedures. *Behavior Analysis in Practice*, 13, 402-410.

Pitts, L., & Hoerger, M. (2021). Mastery criteria and the maintenance of skills in children with developmental disabilities. *Behavioral Interventions*, 36, 522-531.

Richling, S., Williams, W., & Carr, J. (2019). The effects of different mastery criteria on the skill maintenance of children with developmental disabilities. *Journal of Applied Behavior Analysis*, 52, 701-717.

Sundberg, M. (2014). *The verbal behavior milestones assessment and placement program. 2nd. Edition.* Concord, CA: AVB Press.

Wong, K., Fienup, D., Richling, S., Keen, A., & Mackay, K. (2022). Systematic review of acquisition mastery criteria and statistical analysis of associations with response maintenance and generalization. *Behavioral Interventions*, 37, 993-1012.