



Use of an Online Training with Virtual Role Play to Teach Preference Assessment Implementation

Summer Bottini¹ · Jennifer Gillis¹

Accepted: 2 February 2021 / Published online: 10 February 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

Abstract

Identification of reinforcers is critical to the effectiveness of behavioral interventions. Stimulus preference assessments (SPA) are a frequently used method to identify putative reinforcers. Given the fluctuating nature of individual preferences, there is need for efficient training of providers that may regularly implement SPAs. The present study evaluated the utility of a web-delivered training with virtual role play to train SPA implementation. This study builds upon previous literature by utilizing a larger sample and incorporating role-play, a component often omitted from other efficient methods of training. Study 1 trained 40 undergraduate students to implement an SPA via web or in vivo. Results suggest both trainings were equivalently effective, and the web-delivered training reduced trainer time by approximately 25 min. Live role-play and feedback was still necessary with web-delivered training, consistent with suggestions that rehearsal and feedback is a vital component of training. Results also suggest web-delivered training may identify areas of weakness following training. A follow-up clinical pilot showed that the web-delivered training was also effective at training eight novice providers to competently implement the SPA with children with ASD in a special education school. This study demonstrates that web-delivered training with virtual role-play is likely another efficient training method for implementation of behavioral procedures.

Keywords MSWO · Training · Behavioral skills training · Preferences · Online training · Technology

Identification of effective reinforcers is the backbone of successful behavioral interventions. In the case of treating individuals with developmental delays, identifying reinforcers can be a complex task given associated impairments (e.g., difficulty communicating needs, difficulty making choices, etc.). Structured methods to identify stimuli that function as reinforcers have been developed for use with this

✉ Summer Bottini
sbottin1@binghamton.edu

¹ Psychology Department, Binghamton University, Binghamton, NY, USA

population, such as stimulus preference assessments (SPAs). An SPA is a procedure designed to objectively identify a hierarchy of preferred items for individuals (Virués-Ortega et al. 2014). Although SPAs do not guarantee that an item functions as a reinforcer, they are highly predictive of reinforcer effectiveness (Higbee et al. 2000; DeLeon and Iwata 1996). As such, the use of SPAs is advantageous in a variety of applied settings. Indeed, practitioners report frequently using SPAs with their clients with developmental disabilities, like Autism Spectrum Disorder (Graff and Karsten 2012). That being said, researchers have also found that preferences change over time and are unpredictable with respect to when those changes might occur (e.g., Carr et al. 2000; Hanley et al. 2006), making frequent, ongoing implementation of SPAs important. To increase the frequent use of SPAs across settings that serve individuals with developmental disabilities (e.g., schools, clinics, independent therapists, etc.), ongoing training of professionals and paraprofessionals in SPAs is needed (Pence et al. 2012).

Common SPA Training Methods

There has been a growing interest in identifying effective methods to train individuals how to implement SPAs. However, this remains a relatively new area of research, as demonstrated by Leaf and colleagues' (Leaf et al. 2019) systematic review that found only 19 studies that evaluated methods to train implementation of SPAs. Despite there being relatively few studies, all but one study showed that these training methods were moderately to highly effective. The most commonly examined training package was behavioral skills training (BST). A number of alternatives and adaptations have also emerged in recent years.

Behavioral Skills Training (BST) BST is a comprehensive training package that is effective in training numerous skills (e.g., functional analysis, discrete-trial training) to various populations (e.g., therapists, paraprofessionals, caregivers) (e.g., Conklin and Wallace 2019; Miles and Wilder 2009). BST comprises four primary components: instruction, modeling, rehearsal, and feedback (Parsons et al. 2012). Rehearsal and feedback are repeated until the trainee performs at a predetermined mastery criterion. Some studies have suggested that contingent feedback is the necessary and most potent component of this package (e.g., Roscoe et al. 2006; Roscoe and Fisher 2008); whereas, instructions alone is insufficient. There also exists some evidence that modeling may be a similarly important component related to the effectiveness of BST (e.g., Bearman et al. 2013). Researchers have demonstrated that BST is highly effective in training individuals to implement SPAs with high integrity (e.g., Bishop and Kenzer 2012; Lavie and Sturmey 2002; Roscoe and Fisher 2008; Roscoe et al. 2006). One limitation of BST, however, is that it requires many resources, including expert trainers and an extended amount of training time. This is problematic given limited resources available across intervention settings and may explain why BST is infrequently used in practice (DiGennaro Reed and Henley 2015).

Video Modeling Given time and resource constraints within applied settings, there have been efforts to identify other effective training methods that are more efficient and easier to implement. In the context of training SPA implementation, there has been a recent focus on the use of video modeling with voice-over or embedded instructions. Video modeling, in isolation, involves showing a trainee a video of a competent therapist implementing the desired skill (i.e., implementing the SPA). The trainee is then expected to imitate what he/she observed in the video (Catania et al. 2009). Instructions can be embedded within the video to provide additional context and information to the trainee. For example, voice-over instruction can emphasize what the trainee should be attending to in the video frame-by-frame. Numerous studies have demonstrated that video modeling with embedded instructions is an effective training method for training SPA implementation (Deliperi et al. 2015; Delli Bovi et al. 2017; Lipschultz et al. 2015; Rosales et al. 2015; Weldy et al. 2014). Across studies, it is common that trainees re-watch the videos if they do not meet mastery criteria, as assessed in a performance probe (e.g., implementation of the SPA with a confederate or trainer). Thus, live role-play is still a component of the video modeling training method.

Self-Instruction Researchers have also evaluated whether self-instructional materials can be effective in training individuals to implement SPAs with some positive findings (Hansard and Kazemi 2018; Ramon et al. 2015). For example, Ramon and colleagues (Ramon et al. 2015) demonstrated that use of a self-instructional manual was more effective at training undergraduates to implement an SPA than verbally describing how to implement the procedure. Of note, not all participants mastered SPA implementation after using the self-instructional manual. Other researchers have combined self-instructional materials with other common training components to increase its effectiveness. For example, Graff and Karsten (2012) trained individuals to implement SPAs using written step-by-step instructions followed by role-play. Similarly, Arnal Wishnowski and colleagues (Arnal Wishnowski et al. 2018) demonstrated that an online delivered self-instructional manual with video modeling was effective at training individuals to implement SPAs. Taken together, incorporation of modeling and/or role-play appears necessary to the reliable effectiveness of training SPAs.

Limitations of Available Training Methods

The existing literature has achieved a great deal: there now stands multiple training methods from which a supervisor/trainer may choose to use. That being said, there remains room for improvement. First and foremost, training efficiency is an ongoing concern. In their systematic review, Leaf and colleagues (Leaf et al. 2019) found that total training time could take up to six hours for teaching SPA implementation. As such, individual studies emphasize the importance of training efficiency. For example, Roscoe and Fisher (2008) highlighted that their feedback and role-play only took 15 to 20 min, which is notably shorter than other reported training durations. Researchers have also been inventive in developing training adaptations with

the intention to increase efficiency. In addition to antecedent based training methods, like self-instructional materials and video modeling, researchers have used group formats while training (Weldy et al. 2014; Bishop and Kenzer 2012), telehealth or web-based technology to remotely deliver training (Higgins et al. 2017), and pyramidal training to proliferate BST for SPA implementation (Pence et al. 2012). Despite these advancements, training can still require substantial resources with respect to time (e.g., remotely delivered BST can take upward of six hours; Higgins et al. 2017) and require availability of technology (e.g., recording technology or HIPAA-compliant web communication technology).

Additionally, the need for remote training capabilities has become abundantly clear since the onset of the COVID-19 pandemic and various states enacted shelter-in-place orders to decrease the spread of disease and medical burden (Desai and Patel 2020). Behavioral intervention is a high priority for many individuals with disabilities, and as such, continuity of care is critical. Consequently, telehealth delivery of applied behavior analytic services was approved for health insurance reimbursement (US Department of Health and Human Services 2020) and researchers have recommended providers engage in remote options when possible (Cox et al. 2020). Given these efforts to minimize in-person contact, development of virtual methods to effectively train providers is important. Moreover, many caregivers have been thrust into the role of service provider and/or educator during this era of telehealth, further substantiating the need for efficient and remote methods to train novice providers, including caregivers, in behavioral principles and procedures.

Besides issues of efficiency and the need for remote training, all existing literature on training SPA implementation utilizes single-case design (Leaf et al. 2019). Despite many strengths of single-case design, it is limited in generalizable conclusions. Across existing studies, the number of participants trained ranged from 2 to 18, with the majority only examining a handful of participants. Although these studies have contributed much to our understanding on this topic, a next logical step is to evaluate SPA training methods utilizing group design. Furthermore, the paucity of research specific to training SPA implementation requires additional replication across studies. For example, only a handful of studies have evaluated each training method (Leaf et al. 2019), few have specifically trained scoring and interpretation of SPAs (Lipschultz et al. 2015), and no studies have directly compared two or more effective SPA training procedures.

The Present Study

The present study compares a web-delivered training to a live training. We sought to extend the existing literature by directly comparing two full training packages using group design. We also intended to replicate previous findings demonstrating the effectiveness of BST and online-delivered instruction/modeling for teaching inexperienced individuals to implement an SPA. Further, our web-delivered training package was novel in that it included a series of questions that required participant responses, simulating a role-play. The purpose of the simulated question-based role-play was three-fold: First, it required trainees to pass ‘the dead man’s test’ to

demonstrate attending behavior in the absence of a trainer; second, we evaluated whether it would decrease the need for in-person role-play and feedback with a trainer; and, third, we evaluated whether it could identify areas for training improvement. Regarding this last point, Higgins et al. (2017) described the utility in collecting component skill data during SPA implementation to permit individualized follow-up training. Thus, we sought to evaluate whether online training could capture this type of information to inform more optimal training.

Method

Participants

Forty undergraduate students (23 female, 17 male) with no prior experience in preference assessments were recruited from a state university in the United States. Participants reported an average age of 19.39 years ($SD=1.22$, Range 18–24) and an average GPA of 3.27 ($SD=0.64$). Participants identified as Caucasian (61.0%), African American (14.6%), Asian/Pacific Islander (12.2%), Hispanic/Latinx (9.8%), and/or Multiracial (2.4%). We chose to recruit undergraduate students due to their minimal exposure to preference assessments and applied behavior analysis, which is consistent with trends in the currently available empirical literature (e.g., Ramon et al. 2015 stated the need for further research on training with less experienced participants, Leaf et al. 2019 showed approximately a third of recently published articles on this topic utilized undergraduate participants).

Setting and Materials

Training was conducted in one of two small rooms. One room was equipped with a computer on a desk and a chair for participants to complete the online training. The second room was equipped with a small desk, two chairs, and a TV on a desk. In the second room, the experimenter sat in one chair and projected the training PowerPoint onto the TV screen during training for participants who completed the live training condition. Role play and feedback following training was conducted in the second room (the TV was turned off and the experimenter sat opposite the participant at the table). During training, regardless of condition, participants were given guided notes, a blank datasheet,¹ and a pen with which they were allowed to take notes if they chose. The second room was also equipped with a video camera which was stationed on a small table in the corner of the room. This camera was turned on during role-play and feedback to code inter-observer agreement. For the role-play, participants were provided a new blank data sheet and different colored M&Ms to use during the SPA.

¹ Copy of guided notes and/or datasheet are available upon request.

Self-Competency Survey A seven item self-report questionnaire was administered to participants immediately after the training (i.e., before role-play and feedback). The purpose of this survey was to determine social validity of the training procedures, and it included questions like “To what extent did you think this training was useful in teaching the procedure? All questions were rated on a four-point Likert scale, such that lower scores indicated higher satisfaction and higher scores indicated higher dissatisfaction (lowest possible score = 7; highest possible score = 28). Responses were summed to derive a single total score. In addition to these items, the participants were asked whether they would have preferred an online training, in-person training, or had no preference.

General Procedure

After providing informed consent, participants were delivered a brief training in implementation and interpretation of the multiple stimulus without replacement (MSWO) preference assessment. Participants were randomly assigned, via randomizer.org, to receive training in an online format or in-person, each of which is described below. We chose to train participants in the MSWO as this is an efficient and accurate SPA, which has been recommended as a first approach to assessing preference (Karsten et al. 2011). Participants then conducted three MSWO administrations with a confederate. Both live- and online-training groups participated in this live role-play to maintain consistency for comparison of performance between groups following training. After each administration, participants were provided praise for components implemented correctly and corrective feedback on components omitted or implemented incorrectly. Role-play and feedback with each participant took approximately ten to 15 min. Participants implemented the MSWO with a confederate because (1) this is common practice within relevant literature (Leaf et al. 2019), (2) there exists evidence that performance with confederates generalizes to actual clients (e.g., Lipschultz et al. 2015; Roscoe et al. 2006; Ausenhus and Higgins 2019), and (3) use of a confederate permitted all component skills of interest to be probed (Higgins et al. 2017). After implementing an MSWO three times, participants were asked to analyze the data using the datasheet and identify the item(s) they would most likely use with the simulated client and least likely use with the simulated client. Finally, participants completed the self-competency survey and a brief demographic survey.

Training Conditions

Live Training Eighteen participants were assigned to the live training. For this training, participants were provided guided notes and a blank datasheet that corresponded with the presentation. The experimenter first presented a scripted PowerPoint on implementation of the MSWO. Next, the experimenter showed the participant two video models of MSWO implementation with a child, one that used toys and one that used food items. After the two videos, the experimenter presented a second scripted PowerPoint on data collection and interpretation of the MSWO. Any

Table 1 Approximate time each component of training took for individuals involved

	Live Training		Online Training	
	Trainee	Trainer	Trainee	Trainer
Set up	0	5	0	2
PowerPoint Presentations	25	25	25	0
Virtual Role Play	0	0	23	0
Live Role Play	15	15	15	15
Total	40	45	63	17

questions posed by the participant were answered throughout the training, as this is naturalistic for typical live training procedures.

Online Training Twenty-two participants were assigned to the online training. Similar to the live training condition, these participants were provided guided notes, a blank datasheet and were shown the same PowerPoint presentations and video models as the Live Training group. The only difference between the training conditions was that all training materials were delivered online via Qualtrics for this condition. For both PowerPoints, voice-over instruction was used in place of the live presenter and it followed the same script. After the two presentations and video models, participants were then directed through a simulated role-play via the online survey platform. Specifically, participants responded to a series of short video clips as if they were implementing an MSWO. They were directed to take data on the provided datasheet. After responding to each page, they were directed to a page that provided feedback regarding whether they were correct or not.

During the simulation, trials of the MSWO began such that participants watched short clips (e.g., of the confederate selecting an item) and then chose the most appropriate action they should do next in the assessment (e.g., remove other items from reach of the confederate). After each trial, they were also asked to take data and could click a button to check whether their datasheet looked as it should. Specifically, if they clicked the button, the screen displayed a picture of what the datasheet should have written on it at this point in the assessment so that the participant could compare their datasheet to this model. Upon completing the simulated MSWO, participants were then asked a series of questions regarding analysis and interpretation of the MSWO. In particular, participants were asked to calculate the proportion of times each item was selected and then were asked to identify which item should be used as a reward. As with all other components of the simulated session, participants were directed to a page that provided feedback regarding each of their answers. This concluded the online training. See Table 1 for an approximation of the time it took to implement each component for both trainings.

Table 2 Implementation components of MSWO and operational definitions for correct response

Component	Operational Definition
Pre-sample	Providers confederate each item, one at a time, prior to beginning the assessment
Spacing	Places array of items within 6 inches of the confederate. Items in the array are between 2 and 3 inches from one another
Instruction	Presents instruction, “pick one” or something similar, without gesturing to any specific item in the array
Within Trial Removal	Following confederate selection of an item, removes all other items in the array from reach of the confederate
Between Trial Removal	Once an item was selected by the confederate, it is not included in subsequent trial presentations
Item Rotation	Changes the order of items in the array between each trial
Multiple Item Response	If the confederate selects multiple items, he/she removes all items and re-presents the trial
No Item Response	If the confederate does not select an item within 10 s, he/she removes all items and re-presents the trial
Data Collection	Correctly writes down which item is selected on each trial
Identify Reward	Correctly states which item would be optimal to use as a potent reward

Dependent Variables

Implementation Performance For each of the three live role-plays of an MSWO, participant performance was coded by marking whether any error was made on any of the following components for each relevant trial: (1) whether the participant allowed the confederate to sample each item (pre-sample), (2) whether spacing between items was appropriate (spacing), (3) whether the participant delivered an instruction to the confederate to make a selection (instruction), (4) whether the participant removed all items from reach after the confederate made a selection (within-trial removal), (5) whether previously selected items were excluded from the subsequent trials (between-trial removal), (6) whether items were rotated between trials (item rotation), (7) whether the participant re-presented a trial if the confederate consumed multiple items (multiple item response), (8) whether the participant re-presented a trial if the confederate did not select an item within 10-s (no item response), (9) whether the participant took data appropriately (data collection), and (10) whether the participant correctly identified an appropriate reward. See Table 2 for a list of implementation components, including the abbreviated term used hereafter and operational definitions. The total number of errors was summed across trials for each participant. Additionally, participant calculations of proportion selected for each item was recorded as total number of analyses calculated correctly.

Virtual Performance During the simulated role-play, participant responses were recorded. These responses were retrospectively coded as to whether each participant made an error on relevant implementation components (pre-sample, instruction, within trial removal, between trial removal, no item response, and identify reward).

Table 3 Percentage of participants that made an error during each component of the initial live role play versus the virtual role play

Component	Live Role Play	Virtual Role Play
Pre-sample	100%	63.6% ($n = 14$)
Spacing	75.0% ($n = 30$)	–
Instruction	20.0% ($n = 8$)	0%
Within trial removal	37.5% ($n = 15$)	40.9% ($n = 9$)
Between trial removal	2.5% ($n = 1$)	0%
Item rotation	45.0% ($n = 18$)	–
Multiple item response	80.0% ($n = 32$)	–
No item response	60.0% ($n = 24$)	59.1% ($n = 13$)
Data collection	0%	–
Reward Identification	5.0% ($n = 2$)	45.5% ($n = 10$)
Correctly calculated all analyses	75.0% ($n = 30$)	45.5% ($n = 10$)

The number of errors in calculating proportion of times each item was selected was also recorded for each participant.

Interobserver Agreement

An independent rater coded performance during the role-play for 32 (78.05%) of participants. Total interobserver agreement was calculated as follows: (number of agreements / [number of agreements + number of disagreements]) \times 100. Average interobserver agreement was 98.21% (Range = 96.29%–100%).

Results

MSWO Implementation Performance

The total number of errors incurred during the first MSWO implementation with a confederate was compared between participants in the online and live training using an independent samples t-test. There was no difference between groups in the number of errors incurred, $t(38) = 0.826$, $p = 0.414$. Participants in the live training made an average of 9.35 errors during the first MSWO ($SD = 5.09$); whereas, participants in the online training made an average of 8.05 errors ($SD = 4.86$). See Table 3 for the number of participants that made each type of error after the first role-play.

When comparing performance between groups for the second MSWO implementation using an independent samples t-test, there was also no difference in the number of errors $t(38) = -0.16$, $p = 0.875$. Participants in the live training made an average of 2.50 errors during the first MSWO ($SD = 1.85$); whereas, participants in the online training made an average of 2.60 errors ($SD = 2.14$). Implementation was not compared for the third MSWO implementation due to the small number of errors incurred overall (Live: $M = 0.58$, $SD = 0.90$; Online: $M = 1.00$, $SD = 2.05$).

Additionally, an independent samples t-test was conducted to see whether training groups differed in the number of analyses (proportion of times item selected) they correctly calculated. There was no difference between groups, $t(31.92)=0.866$, $p=0.393$.

Virtual Performance

Virtual performance was able to be extracted for 22 of the 23 participants that were assigned to the online training (data for one participant was missing due to technology failure). The most frequently incurred error was related to pre-sample ($n=14$, 64%) and no item response ($n=13$, 59%). Ten participants made an error related to calculating analyses (45.5%) as did ten participants for identifying a reward (45.5%). Nine participants (40.9%) made an error related to within trial removal. No participants made errors regarding between trial removal or instruction. See Table 3 for the percentage of participants that made each type of error during the virtual role-play. As an exploratory analysis, a spearman rho correlation was conducted between percentage of participants that made errors during the virtual training versus the first initial role-play was conducted. This analysis revealed a high correlation coefficient, $r_s=0.76$, but yielded a p-value that would not be considered significant when taking into account the number of analyses conducted in this study ($p=0.04$; family-wise error would require a p-value below 0.005).

Self-Competency Survey

An independent samples t-test revealed no difference in self-competency scores between the online training group ($M=11.45$; $SD=2.63$) and the live training group ($M=11.16$, $SD=2.39$), $t(37)=0.363$, $p=0.719$. Eighteen participants indicated they would prefer receiving a live training in the future, 11 indicated they would prefer an online training in the future, eight indicated they had no preference, and four did not indicate a choice.

Discussion

The present study found that a web-delivered training was similarly effective as its live counterpart. Notably, both training packages resulted in a high number of errors during the initial in vivo role-play with a confederate. This suggests that role-play and feedback was a necessary component to both training types, as a low number of errors were incurred during the second and third role-plays across participants. This finding is consistent with literature demonstrating that role-play and feedback is a critical feature of BST (e.g., Roscoe et al. 2006; Roscoe and Fisher 2008). When considering high performance following role-play and feedback, the present findings support previous findings that BST is an effective training for SPA implementation (e.g., Bishop and Kenzer 2012; Lavie and Sturmey 2002; Roscoe and Fisher 2008).

The equivalent performance during the initial role-play between experimental groups suggests that our virtual role-play did not decrease the need for in-person role-play and feedback. That being said, this study supports other advantages of web-delivered training. Namely, it decreased the time a trainer was required to conduct BST. Specifically, the online training only required approximately 15 min of trainer time for each participant (for role-play and feedback), saving approximately 25-min of trainer time for each participant as instruction and video-modeling was presented online. Thus, online presentation of instruction and video-modeling may represent an alternative to group formats while training (e.g., Weldy et al. 2014; Bishop and Kenzer 2012) and save time for the critical role of individualized role-play and feedback (e.g., Roscoe et al. 2006). However, the online training took participants longer than the live training due to the virtual role-play portion (adding approximately the same amount of time saved for trainers). Since equivalent performance was observed between experimental groups, it is possible this additional component could be omitted. It is also possible that the virtual role-play was necessary to require an attending response while the trainee was not actively supervised by a trainer. Web-delivered training may be less effective than live training without a component that requires the trainee to pass the dead man's test. Determining whether virtual role-play is necessary for online trainings to be effective is outside the scope of the present study, but warrants future research, especially with respect to training implementation of technical behavioral procedures.

In addition to training efficiency, web-delivered training may be advantageous in identifying areas of trainee weakness following completion. When examining the types of errors incurred during the virtual role-play, there was some correspondence with those incurred during the initial live role-play across participants. For example, a high number of participants made errors on pre-sampling and the no item response for both the virtual role-play training and live training. Similarly, few participants made errors regarding delivering instruction and between trial removal for both the virtual role-play training and live training. Thus, an online virtual role-play may be able to capture performance for component skills of SPA implementation, which can be helpful in informing future training (Higgins et al. 2017). By reviewing online performance, a trainer may decide to alter training materials to better target the skills with low performance. Including this data in the present study is explorational in nature and should be interpreted with caution. That being said, it is likely worthwhile to more rigorously explore benefits of virtual role-plays (e.g., to refine training procedures) in future studies. Future research should also explore whether virtual role play may also identify areas of weakness for individual trainees following training completion. Such information may permit more targeted live role-play and feedback.

Regarding social validity, participants in each training group felt similarly regarding their implementation skills. It appears there is no detriment to trainee experience when utilizing an online-delivered training. This substantiates the utility of online-based training when it can be effectively used to increase training efficiency. Participants were relatively split regarding which type of training they would prefer in the future (online versus live); however, it is unknown whether a more consistent preference would have emerged if participants contacted both

forms of training. Further exploration of why individuals prefer online versus live trainings may be informative in regards to improving trainee engagement, but unfortunately was not assessed in the present study. Organizations or trainers may consider providing trainees a choice which type of training they would prefer. Availability of both live and online training materials may provide trainees and trainers more autonomy during training.

This study is not without limitations. First, we did not include a baseline measure of SPA implementation. There is little consistency in the literature regarding appropriate baseline measures for this topic (Deliperi et al. 2015). For example, previous studies use written descriptions (e.g., Graff and Karsten 2012), general vocal instructions (e.g., Lerman et al. 2004, 2008), or no training at all (e.g., Lavie and Sturmey 2002) for the baseline. We purposely selected a participant population with no experience in preference assessments, and all participants indicated on the demographic survey they had no prior learning history with SPAs. Additionally, our primary purpose was to compare two training methods using group design, decreasing the need for a true baseline. Second, the online training did not probe all target skills during the virtual role-play, which limits conclusions that can be made. Again we urge cautious interpretation of this data and present it only as an initial suggestion that this type of data may be useful to trainers/organizations and easily accessed via use of online mediums. Third, the virtual role-play increased training time for participants without increasing performance during the initial live role-play. Although this online training package decreased trainer time, it increased trainee time. Cost–benefit analysis is needed to analyze the relative advantages of saving trainer time versus trainee time. Since we did not evaluate the effectiveness of the online training without the virtual role-play, it remains unclear whether web-based instruction and video-modeling would be equivalently effective as instruction and video-modeling provided in a live context. Fourth, participants had no stake in implementation and generalization to actual children with developmental disabilities was not assessed, limiting external validity of findings.

A final limitation is that these findings continue to support the need for in-person role play, reducing applicability for fully-remote training options. This finding may be due to the fact that live practice and feedback was only available during in-person role-play. From an internal validity perspective, this was considered a performance assessment more so than a training component, despite having implications for necessary training components. In other words, using in-person role-play for both training groups was purposeful in order to have a consistent basis to compare performance. That being said, it is possible that live practice and feedback could be made virtual with a similarly effective impact on performance following an initial web-based training. For example, trainee and supervisor might video conference one another and act out their respective roles with separate sets of training materials. Alternatively, a provider may observe and provide feedback via video conferencing while a trainee implements the procedure with the actual client. This latter approach may be particularly useful when training caregivers via telehealth. Future research may focus on comparing remote methods of using in-the-moment practice and feedback as this continues to emerge as a critical component of training.

Conclusion

Although many effective training packages exist, there is an ongoing need for efficient training packages that would be used in applied settings and may be useful for telehealth services. The present study evaluated the potential utility of a web-delivered training in MSWO implementation for novice providers. We found both trainings were equivalent despite the web-training reducing approximately 25 min of trainer time. Web-delivered training is another way to reduce training time for trainers while maintaining effectiveness for training SPA implementation. Web-delivered training may also be particularly useful as the need for telehealth grows. Use of virtual role play may have additional benefits to increase active participation when training is remote and to identify ways to improve training in the future.

Data Availability Data is available upon request.

Code Availability N/A

Declarations

Ethics Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the Binghamton IRB and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Consent to Participate Informed consent was obtained from all individual participants included in the study.

Consent for Publication Consent was obtained from all authors to publish the present study.

Conflict of Interest The authors have no conflicts to declare.

References

- Arnal Wishnowski, L., Yu, C. T., Pear, J., Chand, C., & Saltel, . (2018). Effects of computer-aided instruction on the implementation of the MSWO stimulus preference assessment. *Behavioral Interventions*, 33, 56–68. <https://doi.org/10.1002/bin.1508>.
- Ausenhus, J. A., & Higgins, W. J. (2019). An evaluation of real-time feedback delivered via telehealth: Training staff to conduct preference assessments. *Behav Analysis Practice*, 12, 643–648. <https://doi.org/10.1007/s40617-018-00326-1>.
- Bearman, S. K., Weisz, J. R., Chorpita, B. F., Hoagwood, K., Ward, A., Ugueto, A. M., & Bernstein, A. (2013). More practice, less preach? The role of supervision processes and therapist characteristics in EBP implementation. *Administration and Policy in Mental Health and Mental Health Services Research*, 40, 518–529. <https://doi.org/10.1007/s10488-013-0485-5>.
- Bishop, M. R., & Kenzer, A. L. (2012). Teaching behavioral therapists to conduct brief preference assessments during therapy sessions. *Research in Autism Spectrum Disorders*, 6, 450–457. <https://doi.org/10.1016/j.rasd.2011.07.005>.
- Catania, C. N., Almeida, D., Liu-Constant, B., & DiGennaro Reed, F. D. (2009). Video modeling to train staff to implement discrete-trial instruction. *Journal of Applied Behavior Analysis*, 42, 387–392. <https://doi.org/10.1901/jaba.2009.42-387>.

- Carr, J. E., Nicholson, A. C., & Higbee, T. S. (2000). Evaluation of a brief multiple-stimulus preference assessment in a naturalistic context. *Journal of Applied Behavior Analysis, 33*, 353–357. <https://doi.org/10.1901/jaba.2000.33-353>.
- Conklin, S. M., & Wallace, M. D. (2019). Pyramidal parent training using behavioral skills training: Training caregivers in the use of a differential reinforcement procedure. *Behavioral Interventions, 34*(3), 377–387. <https://doi.org/10.1002/bin.1668>.
- Cox, D. J., Plavnick, J. B., & Brodhead, M. T. (2020). A proposed process for risk mitigation during the COVID-19 pandemic. *Behavior Analysis in Practice*. Advance online publication. <https://doi.org/10.1007/s40617-020-00439-1>
- DeLeon, I. G., & Iwata, B. A. (1996). Evaluation of a multiple stimulus presentation format for assessing reinforcer preferences. *Journal of Applied Behavior Analysis, 29*, 519–533. <https://doi.org/10.1901/jaba.1996.29-519>.
- Deliperi, P., & VLadescu, J. C., Reeve, K. F., Reeve, S. A., & DeBar, R. M. . (2015). Training staff to implement a paired-stimulus preference assessment using video modeling with voiceover instruction. *Behavioral Interventions, 30*, 314–332. <https://doi.org/10.1002/bin.1421>.
- Delli Bovi, G. M., & VLadescu, J. C., DeBar, R. M., Carroll, R. A., & Sarokoff, R. A. . (2017). Using video modeling with voice-over instruction to train public school staff to implement a preference assessment. *Behavioral Analysis in Practice, 10*, 72–76. <https://doi.org/10.1007/s40617-016-0135-y>.
- Desai, A. N., & Patel, P. (2020). Stopping the spread of COVID-19. *Journal of the American Medical Association*. Advance online publication. <https://doi.org/10.1001/jama.2020.4269>
- DiGennaro Reed, F. D., & Henley, A. J. (2015). A survey of staff training and performance management: The good, the bad, and the ugly. *Behavior Analysis in Practice, 8*(1), 16–26. <https://doi.org/10.1007/s40617-015-0044-5>.
- 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>.
- Hansard, C., & Kazemi, E. (2018). Evaluation of video self-instruction for implementing paired-stimulus preference assessments. *Journal of Applied Behavior Analysis, 51*, 675–680. <https://doi.org/10.1002/jaba.476>.
- Hanley, G. P., Iwata, B. A., & Roscoe, E. M. (2006). Some determinants of changes in preference over time. *Journal of Applied Behavior Analysis, 39*(2), 189–202. <https://doi.org/10.1901/jaba.2006.163-04>.
- Higbee, T. S., Carr, J. E., & Harrison, C. D. (2000). Further evaluation of the multiple-stimulus preference assessment. *Research in Developmental Disabilities, 21*(1), 61–73. [https://doi.org/10.1016/S0891-4222\(99\)00030-X](https://doi.org/10.1016/S0891-4222(99)00030-X).
- Higgins, W. J., Luczynski, K. C., Carroll, R. A., Fisher, W. W., & Mudford, O. C. (2017). Evaluation of a telehealth training package to remotely train staff to conduct a preference assessment. *Journal of Applied Behavior Analysis, 50*, 238–251. <https://doi.org/10.1002/jaba.370>.
- Karsten, A. M., Carr, J. E., & Lepper, T. L. (2011). Description of a practitioner model for identifying preferred stimuli with individuals with autism spectrum disorders. *Behavior Modification, 35*(4), 347–369. <https://doi.org/10.1177/0145445511405184>.
- Lavie, T., & Sturmey, P. (2002). Training staff to conduct a paired-stimulus preference assessment. *Journal of Applied Behavior Analysis, 35*(2), 209–211. <https://doi.org/10.1901/jaba.2002.35-209>.
- Leaf, J. B., Milne, C., Aljohani, W. A., Ferguson, J. L., Cihon, J. H., Oppenheim-Leaf, M. L., et al. (2019). Training change agents how to implement formal preference assessments: A review of the literature. *Journal of Developmental and Physical Disabilities, 32*, 41–56. <https://doi.org/10.1007/s10882-019-09668-2>.
- Lerman, D. C., Tetreault, A., Hovanetz, A., Strobel, M., & Garro, J. (2008). Further evaluation of a brief, intensive teacher-training model. *Journal of Applied Behavior Analysis, 41*, 243–248. <https://doi.org/10.1901/jaba.2008.41-243>.
- Lerman, D. C., Vorndran, C. M., Addison, L., & Kuhn, S. C. (2004). Preparing teachers in evidence-based practices for young children with autism. *School Psychology Review, 33*(4), 510–526. <https://doi.org/10.1080/02796015.2004.12086265>.
- Lipschultz, J. L., VLadescu, J. C., Reeve, K. F., Reeve, S. A., & Dipsey, C. R. (2015). Using video modeling with voiceover instruction to train staff to conduct stimulus preference assessments. *Journal of Developmental and Physical Disabilities, 27*, 505–532. <https://doi.org/10.1007/s10882-015-9434-4>.

- Miles, N. I., & Wilder, D. (2009). The effects of behavioral skills training on caregiver implementation of guided compliance. *Journal of Applied Behavior Analysis*, 42(2), 405–410. <https://doi.org/10.1901/jaba.2009.42-405>.
- Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-based staff training: A guide for practitioners. *Behavior Analysis in Practice*, 5(2), 2–11. <https://doi.org/10.1007/BF03391819>.
- Pence, S. T., St Peter, C. C., & Tetreault, A. S. (2012). Increasing accurate preference assessment implementation through pyramidal training. *Journal of Applied Behavior Analysis*, 45(2), 345–359. <https://doi.org/10.1901/jaba.2012.45-345>.
- Ramon, D., Yu, C. T., Martin, G. L., & Martin, T. (2015). Evaluation of a self-instructional manual to teach multiple-stimulus without replacement preference assessments. *Journal of Behavioral Education*, 24, 289–303. [1007/s10864-015-9222-3](https://doi.org/10.1007/s10864-015-9222-3)
- Rosales, R., Gongola, L., & Homlitas, C. (2015). An evaluation of video modeling with embedded instructions to teach implementation of stimulus preference assessments. *Journal of Applied Behavior Analysis*, 48, 209–214. <https://doi.org/10.1002/jaba.174>.
- Roscoe, E. M., & Fisher, W. W. (2008). Evaluation of an efficient method for training staff to implement stimulus preference assessments. *Journal of Applied Behavior Analysis*, 41, 249–254. <https://doi.org/10.1901/jaba.2008.41-249>.
- Roscoe, E. M., Fisher, W. W., Glover, A. C., & Volkert, V. M. (2006). Evaluating the relative effects of feedback and contingent money for staff training of stimulus preference assessments. *Journal of Applied Behavior Analysis*, 39, 63–77. <https://doi.org/10.1901/jaba.2006.7-05>.
- U.S. Department of Health and Human Services (2020). *Notification of enforcement discretion for telehealth remote communications during the COVID-19 nationwide public health emergency*. <https://www.hhs.gov/hipaa/for-professionals/special-topics/emergency-preparedness/notification-enforcement-discretion-telehealth/index.html>. Accessed 29 May 2020.
- Virués-Ortega, J., Pritchard, K., Grant, R. L., North, S., Hurtado-Parrado, C., Lee, M. S., et al. (2014). Clinical decision making and preference assessment for individuals with intellectual and developmental disabilities. *American Journal on Intellectual and Developmental Disabilities*, 119(2), 151–170. <https://doi.org/10.1352/1944-7558-119.2.151>.
- Weldy, C. R., Rapp, J. T., & Capocasa, K. (2014). Training staff to implement brief stimulus preference assessments. *Journal of Applied Behavior Analysis*, 47, 214–218. <https://doi.org/10.1002/jaba.98>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.