



Evaluating the Cool Versus Not Cool Procedure via Telehealth

Joseph H. Cihon^{1,2}  · Julia L. Ferguson^{1,2} · Matthew Lee¹ · Justin B. Leaf^{1,2} · Ron Leaf¹ · John McEachin¹

Accepted: 26 January 2021 / Published online: 12 February 2021
© Association for Behavior Analysis International 2021

Abstract

Autistics/individuals diagnosed with autism spectrum disorder (ASD) commonly display qualitative impairments in social behavior that commonly result in the use of interventions directly targeting the development of social skills. The COVID-19 pandemic has highlighted the need for effective social skills interventions that can be delivered directly via telehealth. The Cool Versus Not Cool procedure has continually been documented as effective within the literature. However, its reported use has been limited to in-person delivery. The purpose of this study was to evaluate the effectiveness of the Cool Versus Not Cool procedure conducted via telehealth to teach three children diagnosed with ASD to change the conversation when someone is bored. The results of a nonconcurrent multiple-baseline across-participants design demonstrated that all three participants reached the mastery criterion in four to eight sessions. Responding generalized to another adult for two of the three participants, and all three participants maintained correct responding. Social validity measures indicated the skill was important to teach, the intervention was acceptable and effective, and the telehealth format was an acceptable replacement for in-person intervention for these three participants.

Keywords Telehealth · Social skills · Cool Versus Not Cool · Autism · Social discrimination

Autistics/individuals diagnosed with autism spectrum disorder¹ (ASD) commonly display qualitative impairments in social behavior (American Psychiatric Association, 2013). The challenges related to social behavior for these individuals commonly result in the use of interventions directly targeting the development of social skills. To date, the effectiveness of several interventions for the development of social skills has been documented within the peer-reviewed literature. These include, but are not limited to, the teaching interaction procedure (e.g., Dotson et al., 2010; Leaf, Oppenheim-Leaf, et al., 2012a), video modeling (e.g., Rudy et al., 2014), discrete-trial teaching (e.g., Garcia-Albea et al., 2014), pivotal response treatment (e.g., Mohammadzaheri et al., 2014), behavioral

skills training (e.g., Stewart et al., 2007), and the Cool Versus Not Cool procedure (e.g., Milne et al., 2017).

Although it is most common for these interventions to occur in person, in a clinical or home setting, the COVID-19 pandemic has illustrated the need for effective interventions that can be delivered via telehealth (Cox et al., 2020; LeBlanc et al., 2020). Behavior-analytic research related to intervention for autistics/individuals diagnosed with ASD via telehealth is not new (Ferguson et al., 2019), but the sudden move to telehealth for many service providers has accelerated the need for research evaluating the effectiveness of behavior-analytic procedures delivered via telehealth (Cox et al., 2020; LeBlanc et al., 2020). To date, much of the behavior-analytic research related to telehealth-delivered behavior-analytic interventions has focused on training an individual to subsequently implement the intervention in person (see Ferguson et al., 2019, for a review). However, there are some recent, notable examples of the direct application of behavior-analytic interventions delivered via online tools.

For example, Pellegrino and DiGennaro Reed (2020) evaluated an intervention delivered via telehealth that used total task chaining and least-to-most prompting for two adults with intellectual and developmental disabilities. All sessions occurred via VSee, with the two adults participating from their apartments while the experimenter was located in a separate

¹ This terminology was selected to adhere to the 7th edition of the American Psychological Association's *Publication Manual* and to be inclusive of both those who prefer person-first language and those who prefer identity-first language.

✉ Joseph H. Cihon
jcihon@apfmail.org

¹ Autism Partnership Foundation, 200 Marina Dr., Seal Beach, CA 90740, USA

² Endicott College, Beverly, MA, USA

apartment within the same complex. Targeted skills were those that the participants expressed interest in learning (e.g., light cooking, money management). The results of a multiple-probe across-behaviors design indicated the intervention was effective for teaching both participants three self-selected skills. Furthermore, each of the skills maintained, and both participants indicated satisfaction with the procedures and outcomes. In another recent example, Ferguson et al. (2020) evaluated the effectiveness of discrete-trial teaching with instructive feedback delivered via telehealth to teach tact relations to six children diagnosed with ASD. The participants were divided into dyads, and each participant had their own primary (i.e., targeted tact) and secondary (i.e., instructive feedback) targets. All sessions occurred via Zoom, with the interventionist and participants located in different physical locations. The results indicated that all participants acquired primary and secondary responses, and five of six acquired primary and secondary observational responses (i.e., the targets for the other participant in the dyad).

Although the aforementioned (i.e., Ferguson et al., 2020; Pellegrino & DiGennaro Reed, 2020) and past research (e.g., Wacker et al., 2013) is promising for many who are shifting to service delivery via telehealth, the research is limited with respect to social skills interventions delivered directly via telehealth. One intervention approach that may transfer with little effort to a telehealth model is the Cool Versus Not Cool procedure. The Cool Versus Not Cool procedure is an approach to developing social discriminations through what is commonly referred to as discrimination training via instructor demonstration. Role-plays are then used to increase the likelihood the learner will engage in the desired social skill in the terminal environment. More specifically, the Cool Versus Not Cool procedure consists of five components: (a) labeling the targeted social skill (e.g., the “cool” skill), (b) the interventionist modeling the cool (i.e., the desired topography of the behavior) and not cool ways (i.e., the undesired topography of the behavior) to display the social skill, (c) providing the learner with the opportunity to label the model as cool or not cool and why the model was cool or not cool, (d) the learner role-playing the cool way, and (e) providing reinforcement or feedback based on learner responding throughout. It should also be noted the terms “cool” and “not cool” were selected based on the learners with whom the procedure was originally developed. Those learners were using those words already, and it was thought that using those same words to describe desired and undesired social behaviors would increase the likelihood of generalization and maintenance in the terminal environment (Leaf et al., 2020). As such, interventionists should determine the labels that will be most appropriate based on the learners with whom they provide intervention (e.g., “dope” and “weak,” “good choice” and “bad choice,” or “appropriate” and “inappropriate” may be more appropriate and effective than “cool” and “not cool” for some learners).

The Cool Versus Not Cool procedure has been demonstrated to be effective for teaching a variety of social skills to autistics/individuals diagnosed with ASD (e.g., interrupting, changing the game, greetings, joint attention, changing the conversation, abduction prevention, and eye contact; Leaf, Tsuji, et al., 2012b). The effectiveness of the Cool Versus Not Cool procedure has also been demonstrated in one-to-one settings (e.g., Leaf et al., 2015), small group settings (Au et al., 2016), and large group settings (Milne et al., 2017). In light of the recent increased need for effective telehealth-delivered interventions, the purpose of the present study was to evaluate the effectiveness of the Cool Versus Not Cool procedure conducted via telehealth for three children diagnosed with ASD.

Method

Participants

Three children independently diagnosed with ASD participated in the study. Participant demographic information is provided in Table 1. All participants had a previous history of receiving in-person social skills interventions including the Cool Versus Not Cool procedure. All participants had some experience with direct intervention delivered via telehealth due to the COVID-19 pandemic but had recently transitioned back to in-person intervention. All participants were also currently participating in a social skills group two times a week via telehealth. None of the participants had any previous experience with the use of the Cool Versus Not Cool procedure to target changing the conversation when someone is bored. Informed consent was obtained from each of the participants’ parents prior to participation in the study. The participants were free to leave at any point during the session; however, this never occurred, and participants assented to all sessions.

Interventionist

Julia L. Ferguson, the second author, served as the interventionist for all sessions, with the exception of generalization sessions. She was a 29-year-old White female with an undergraduate degree in applied behavior analysis and a master’s degree in behavior analysis and had begun her studies toward a doctoral degree in applied behavior analysis. She had over 8 years of experience implementing interventions based on the principles and procedures of applied behavior analysis for individuals diagnosed with ASD. This experience also included the use of the methods within this study to teach a variety of social skills to a variety of learners.

Table 1 Participant Demographic Information

Name	Age	Sex	Race/ethnicity	Diagnosis	Education	WPPSI-IV FSIQ	Vineland-3	SRS-2	SSiS
Winston	5 years 1 month	Male	Chinese American	ASD	Pre-K	97 (average)	79 (moderately low)	43 (within normal limits)	88 (low average)
Nick	4 years 0 months	Male	White & Chinese American	ASD	Pre-K	103 (average)	69 (low)	61 (mild range)	55 (very low)
Schmidt	5 years 1 month	Male	White & Chinese American	ASD	Pre-K	108 (average)	70 (low)	69 (moderate range)	70 (low)

Note. Standard scores and descriptions are provided for participants' full-scale IQ (FSIQ), Vineland-3 adaptive behavior composite (Sparrow, Cicchetti, & Saulnier, 2016), and Social Skills Improvement System (SSiS; Gresham & Elliott, 2008) results. A T score and descriptor are provided for participants' Social Rating Scale results (SRS-2; Constantino & Gruber, 2012). All FSIQ scores were obtained using the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV; Wechsler, 2012). ASD = autism spectrum disorder.

Setting

Throughout all conditions, probes, and intervention sessions, the participants and the interventionist were in different locations in Southern California. Winston was located at home for all of his sessions, whereas Nick and Schmidt were located at an isolated location within a private clinic. The interventionist conducted all sessions from her home. All sessions were conducted using the Zoom Video Communications, Inc. (Zoom; www.Zoom.us), platform using various devices with video and audio capabilities. Winston, Nick, and Schmidt used an iPad, and the interventionist used a laptop computer.

Dependent Measures

The participants' supervisors, who were responsible for training staff, developing curriculum and intervention strategies, and overseeing the participants' overall progress, were asked to provide a list of social skills that were likely to be included relatively soon within the participants' regularly scheduled clinical sessions. The lists provided by the supervisors were examined for areas of overlap across participants. Each of the supervisors noted the participants' challenges with changing the conversation when someone is bored (i.e., identifying possible boredom cues and changing their behavior as a result). Therefore, the main dependent variable for all three participants was changing the conversation when someone is bored. Changing the conversation when someone is bored was divided into seven component steps (see Table 2). Participant engagement in each of the steps was assessed during probe sessions (described later). The mastery criterion was defined as the participant engaging in each of the steps in the outlined order across three consecutive probe sessions. Generalization of changing the conversation when someone is bored was also measured once prior to intervention and once following a participant reaching the mastery criterion. Probe sessions to assess generalization after intervention occurred between 2

and 5 days following a participant reaching the mastery criterion.

Probe Sessions

Probe sessions occurred across each condition (i.e., baseline, intervention, generalization, and maintenance). Probe sessions consisted of one opportunity for the participant to demonstrate the targeted skill and lasted an average of 2 min (range 1–4 min). During the intervention condition, probe sessions always preceded intervention sessions with a 1 min break between sessions. Probe sessions were used to assess participant responding in the absence of direct instruction, prompting, or programmed reinforcement.

To begin each probe, the interventionist engaged the participant in a conversation about a preferred topic. These topics were determined by discussing with the participant's staff and supervisor the participant's preferred movies, video games, and activities. Following 2 min or a minimum of four exchanges, the interventionist began to engage in nonvocal boredom cues for up to 15 s. These cues consisted of looking away from the screen, looking at their phone or watch, or not responding to the participant. If the participant engaged in the steps for changing the conversation when someone is bored, the interventionist responded in accordance with the task analysis. That is, the interventionist engaged in a vocal/verbal response related to the statement made by the participant. If the participant did not engage in the steps for changing the conversation when someone is bored, the interventionist ended the probe by saying "thanks" and returning them back to the activity in which they were previously engaged.

General Procedure

Sessions occurred once a day, 2 to 5 days a week, depending on participant and interventionist availability. Intervention sessions lasted an average of 10 min (range 7–17 min). The

Table 2 Components of Changing the Conversation When Someone Is Bored

Step	Description
1	Faces the screen
2	Maintains a neutral or positive facial expression
3	Maintains a neutral or positive tone
4	Asks a question or makes a statement within 10 s of the interventionist engaging in a cue signaling boredom
5	Waits for the interventionist to respond
6	Asks a second question or makes a statement about a new topic within 10 s of the interventionist's response
7	Waits for the interventionist to respond

interventionist sent a link for the Zoom video conference to the participant's staff member. The staff member began the video conference for the participant and was available during the intervention sessions for role-plays. The staff member did not have any other interactions or functions during research sessions (i.e., they did not function as a shadow or provide any prompts or praise throughout the sessions).

Baseline

The purpose of baseline was to assess participant responding prior to any intervention or programmed reinforcement. Baseline sessions consisted of a probe session (previously described).

Generalization

Generalization of changing the conversation when someone is bored was assessed prior to intervention and after reaching the mastery criterion for all participants. To assess generalization, probe sessions occurred as described previously (i.e., via Zoom) with the exception that the participant's supervisor served as the conversation partner (i.e., someone different from the interventionist and staff member). Within the clinic, the participant's supervisor was responsible for training staff, developing curriculum and intervention strategies, and overseeing the participant's overall progress. Winston's supervisor was a 33-year-old White female with a bachelor's degree in psychology and 11 years of experience providing intervention for autistics/individuals diagnosed with ASD. Nick's supervisor was a 30-year-old White female with a bachelor's degree in psychology and 7.5 years of experience providing intervention for autistics/individuals diagnosed with ASD. Schmidt's supervisor was a 39-year-old Korean female with a master's degree in applied behavior analysis and 10 years of experience providing intervention for autistics/individuals diagnosed with ASD.

Intervention

Intervention consisted of the Cool Versus Not Cool procedure delivered via Zoom video conferences. To begin, the interventionist labeled the targeted skill (e.g., "Today we are going to work on changing the conversation when someone is bored."). The interventionist then provided a demonstration with the participant's staff member. The demonstration consisted of the interventionist and the participant's staff member engaging in a conversation via the Zoom connection. The staff member (who was in the same room as the participant) would then engage in nonvocal boredom cues. The interventionist followed the steps for changing the conversation when someone is bored (i.e., "cool" demonstration) or responded similarly to how the participant was responding to nonvocal boredom cues during probe sessions (i.e., "not cool" demonstration). The interventionist then ended the demonstration and asked the participant to label whether the interventionist responded in the cool or not cool way (e.g., "Was that cool or not cool?"). If the participant responded correctly, the interventionist provided praise and asked the participant to label why the demonstration was cool or not cool (e.g., "That's right! Why was it cool/not cool?"). If the participant responded incorrectly, the interventionist provided feedback and asked why the demonstration was cool or not cool (e.g., "No, that was actually not cool. Tell me why that was not cool."). There were a total of four demonstrations with two cool and two not cool demonstrations that were randomized prior to each session. Following the demonstrations, the interventionist then provided the participant with an opportunity to practice. During these role-plays, the interventionist engaged the participant in a conversation about a preferred topic and began to engage in nonvocal boredom cues. If the participant engaged in the steps for changing the conversation when someone is bored, the interventionist provided praise (e.g., "That's the way! Kiss your brain!"). If the participant did not engage in the steps for changing the conversation when someone is bored, the interventionist ended the role-play and provided corrective feedback (e.g., "You missed it. I was bored, and you didn't change the conversation."). This

continued until the participant engaged in all the steps for changing the conversation when someone is bored correctly.

Maintenance

Maintenance probes began 7 days following the participant reaching the mastery criterion for changing the conversation when someone is bored (as previously described). All maintenance sessions were conducted by the previously described interventionist. Three maintenance sessions occurred for each participant.

Experimental Design

A nonconcurrent multiple-baseline design (Watson & Workman, 1981) across participants, with a modification to improve experimental control, was used to evaluate the effectiveness of the Cool Versus Not Cool procedure delivered via telehealth on the participants' changing the conversation when someone is bored. This design was selected for its flexibility when conducting research in applied settings, which a concurrent multiple-baseline design does not allow. This flexibility was even more necessary for conducting research during an ongoing pandemic. Traditionally, within a nonconcurrent multiple-baseline design, baseline phases are predetermined, and participants are randomly assigned to each baseline length as they become available (Watson & Workman, 1981). Similar to previous studies (e.g., Cihon et al., 2019), an additional criterion common within multiple-baseline logic was used in this study in an attempt to improve the strength of the design. Participants progressed from baseline to intervention once a stable level of responding was observed during baseline. That is, we extended baseline sessions for the next participant until intervention effects were observed with the previous participant if necessary. Therefore, predetermined baseline phases and assignment were not used within this study. Experimental control was demonstrated when the intervention resulted in changes in a participant's behavior without changes in the remaining participants' behavior during baseline sessions (Baer et al., 1968; Carr, 2005). Furthermore, although the nonconcurrent multiple-baseline design permits participant removal if stable responding is not obtained, no participants were removed from this study for this or any other reason.

Interrater Agreement and Treatment Fidelity

Julia L. Ferguson, the second author (who was the interventionist in the study), also served as the primary rater for all sessions. Matthew Lee, the third author, served as the secondary rater. He held an undergraduate degree in psychology and had 6 months of experience with interventions based in applied behavior analysis for autistics/individuals diagnosed

with ASD. Interrater agreement was collected on the primary dependent variable on 37.5% of all sessions across participants and conditions. Agreements were defined as both raters scoring the same response on a step for changing the conversation when someone is bored. Disagreements were defined as one rater scoring one response and the other rater scoring a different response on a step for changing the conversation when someone is bored. Interrater agreement was calculated by dividing the number of agreements by the number of disagreements plus agreements and dividing by 100. Interrater agreement across all three participants was 100% during baseline, generalization, intervention, and maintenance conditions.

The fidelity of the interventionist's implementation of probes and the Cool Versus Not Cool procedure was also assessed. Matthew Lee (previously described) independently observed 35.7% of probe sessions and 37.8% of intervention sessions across all participants and conditions to score the interventionist's behavior. Correct interventionist behavior during probes consisted of (a) engaging the participant in a conversation about a preferred topic, (b) engaging in nonvocal boredom cues following 2 min or a minimum of four exchanges, (c) responding in accordance with the task analysis if the participant engaged in the steps of the targeted skill, and (d) ending the probe by saying "thanks" if the participant did not engage in the steps of the targeted skill. Treatment fidelity was calculated by dividing the number of steps the interventionist displayed correctly by the total number of steps and multiplying by 100. Treatment fidelity for probes averaged 100% across all participants and conditions.

Correct interventionist behavior during intervention sessions consisted of (a) labeling the targeted skill, (b) providing an instruction for the participant to watch the demonstration, (c) providing two cool and two not cool demonstrations, (d) providing the participant with an opportunity to rate whether the demonstration was cool or not cool after each demonstration, (e) providing the consequence that corresponded with the participant's response after each rating, (f) providing the participant with an opportunity to label why the demonstration was cool or not cool after each demonstration, (g) providing the consequence that corresponded with the participant's response after each answer, (h) providing the participant with an opportunity to role-play the targeted skill, (i) providing the consequence that corresponded with the participant's response, and (j) repeating role-plays until the participant engaged in all steps of the targeted skill correctly. Treatment fidelity for intervention sessions averaged 100% across all participants.

Social Validity

To assess social validity, a questionnaire with four questions was sent to the participants' supervisors and parents at the end of the study. The supervisors and parents were provided with a

video from the intervention condition (i.e., without the probe) selected at random for their child/client and were asked to fill out the questionnaire after viewing the video. The questions consisted of the following:

1. How important was it for the/you child to learn the skill the interventionist was teaching?
2. Please rate the degree to which you found the intervention to be acceptable.
3. Please rate the degree to which you found the intervention to be effective.
4. Do you feel that this method of social skills instruction is an acceptable replacement for in-person social skills instruction?

Responses to each of the questions were on a Likert scale from 1 (i.e., *not at all important, not at all acceptable, and not at all effective*) to 5 (i.e., *very important, very acceptable, and very effective*). The supervisors and parents then sent the questionnaires back to the researchers anonymously.

Results

Figure 1 displays the results for all three participants across all conditions. During baseline, all three participants engaged in the same number of steps for changing the conversation when someone is bored. All three participants were consistently engaging in the first three steps (i.e., facing the screen, maintaining a neutral or positive facial expression, and maintaining a neutral or positive tone) but not any of the remaining steps, which is not surprising given the possibility that those steps are likely necessary for any conversation-based skills. This responding continued during the assessment of generalization prior to the intervention condition. All three participants reached the mastery criterion (i.e., engaging in each of the steps in the outlined order across three consecutive probe sessions) during the intervention condition. The total intervention time required for Winston, Nick, and Schmidt to reach the mastery criterion was 70, 45, and 59 min, respectively. After reaching the mastery criterion, Winston's responding during the assessment of generalization returned to baseline levels, whereas Nick and Schmidt continued to engage in all steps correctly during the assessment of generalization. All three participants continued to engage in all steps correctly during maintenance sessions that occurred 7 days after they reached the mastery criterion.

All three supervisors and only one parent returned the social validity questionnaire. When asked how important it was for the child to learn the skill the interventionist was teaching, all three supervisors responded with a 4. When asked to rate the degree to which they found the intervention to be acceptable, all three supervisors responded with a 5. When asked to

rate the degree to which they found the intervention to be effective, the supervisors responded with a 4 on average (range 3–5). When asked if they felt that this method of social skills instruction is an acceptable replacement for in-person social skills instruction, all three supervisors responded with a 4. The only parent to return the social validity questionnaire responded with a 5 on all questions. Given the context in which this study occurred (i.e., COVID-19 pandemic) that may have resulted in increased stress and hardships for the parents and families, the researchers did not persist in requests to return the questionnaire if the parents did not do so after one follow-up.

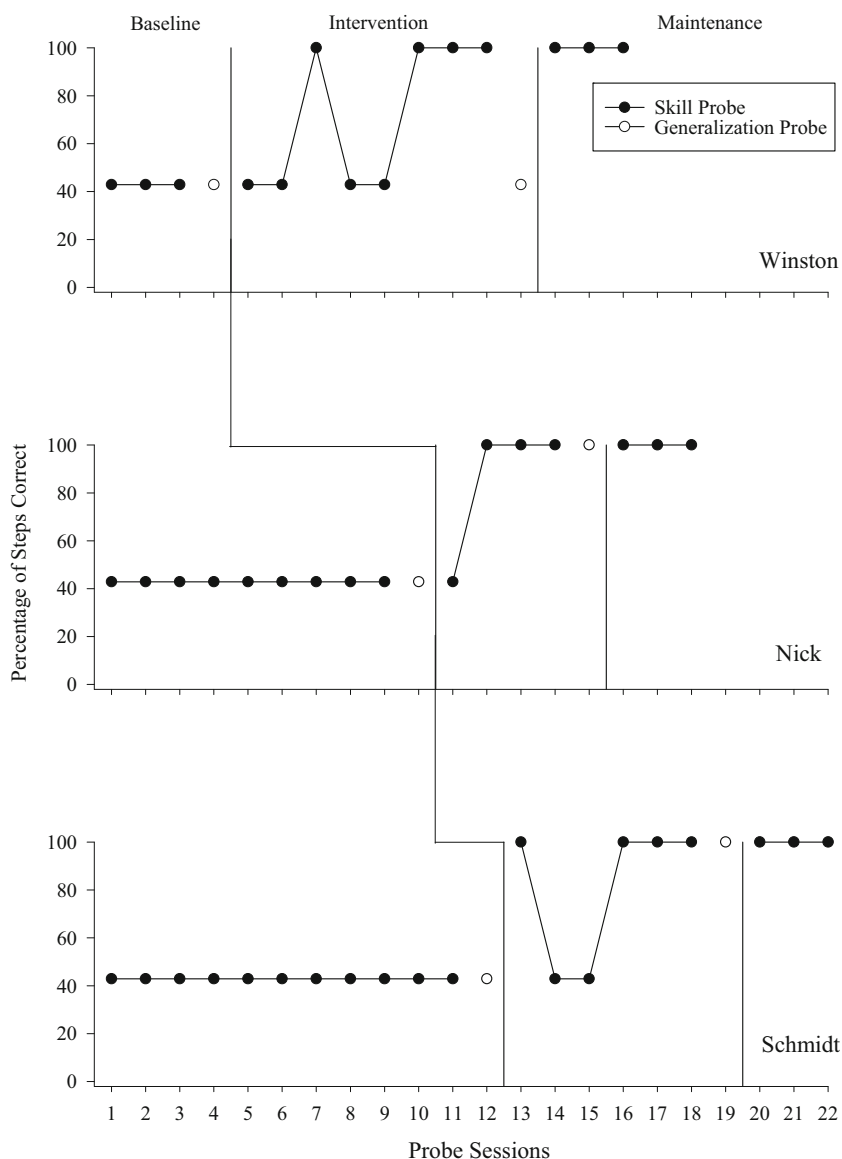
Discussion

The purpose of the present study was to evaluate the effectiveness of the Cool Versus Not Cool procedure conducted via telehealth to teach three children diagnosed with ASD to change the conversation when someone is bored. All three participants reached the mastery criterion in a relatively low number of sessions (range 4–8). Responding generalized to another adult for two of the three participants, and all three participants maintained correct responding on all steps of the targeted skill on all maintenance probes. Furthermore, responses to the social validity questionnaires indicated the skill was important to teach, the intervention was acceptable and effective, and the telehealth format was an acceptable replacement for in-person intervention for these three participants. These results have implications for clinicians providing intervention for autistics/individuals diagnosed with ASD.

The development of an effective social skills repertoire is a common focus of interventions for autistics/individuals diagnosed with ASD. Recent events have put many clinicians in the position of shifting this intervention to telehealth. Given the limited research on effective social skills interventions delivered via telehealth, this study provides clinicians with a viable option when making this shift. The results indicated that the Cool Versus Not Cool procedure can effectively be used when delivered via telehealth under conditions similar to this study (e.g., staff available, similar participant demographics, similar social skill). Furthermore, given previous comparisons with other interventions (e.g., Social Stories; Leaf et al., 2016) and a lack of data supporting the use of those other interventions via telehealth, clinicians may opt for the use of the Cool Versus Not Cool procedure over interventions such as Social Stories. However, future research will be necessary to compare these interventions within the same telehealth context to provide empirical evidence for the use of one procedure over the other.

Clinicians should also closely monitor the generalization of social skills that are targeted using the Cool Versus Not Cool procedure via telehealth. Winston's responding within the teaching context did not generalize to a very similar context.

Fig. 1 Participant Responding During Probe Sessions



It may be the case that a social skills intervention delivered via telehealth poses limitations that are absent in an in-person intervention. The social context involved in the delivery of interventions via telehealth differs greatly from the social context of in-person delivery. This difference could make developing the desired stimulus controls for social behavior difficult. For example, there may be subtleties within in-person interactions that become part of the contingencies that are not present or possible in a telehealth context. This difference may vary based on the social skill that is targeted, and future research may more efficiently evaluate this by targeting multiple skills within a study.

The results of the social validity questionnaire indicated the intervention was appropriate, effective, and an acceptable replacement for in-person social skills instruction for the respondents of the survey. These results provide more support for clinicians to use the Cool Versus Not Cool procedure when delivering social skills intervention via telehealth. This may be

a welcomed result with respect to the numerous considerations required when making the move to telehealth (Cox et al., 2020). It should be noted, however, that there was a low response rate to the survey from the parent respondents, which may affect the interpretation of the social validity data. Furthermore, clinicians should remain cautious and take into consideration individual differences, given the limited number of participants, skills, and social validity respondents within this study.

This study did not go without limitations that warrant discussion. First, although the interventionist conducted all sessions via Zoom, each of the participants had a staff member present at their location during all sessions who also assisted during the demonstration portion of the Cool Versus Not Cool procedure. This might not be possible in all situations. However, it is common for a parent or caregiver to be present during telehealth sessions. In these cases, a parent could fill the position of the staff member. Second, only one skill was

targeted with each of the participants. This limits the possible generality of the results to other social skills. Previous research has demonstrated the effectiveness of the Cool Versus Not Cool procedure to teach a variety of social skills, and it is possible those results would extend to a telehealth context. Future research will be necessary to evaluate whether that is the case. Third, all the participants in this study had a previous history of receiving in-person social skills interventions including the Cool Versus Not Cool procedure, had some experience with direct intervention delivered via telehealth, and had rather well-developed repertoires (see Table 1). As such, the results may not generalize to other autistics/individuals diagnosed with ASD with less experience with the Cool Versus Not Cool procedure and intervention delivered via telehealth or with less developed repertoires. Fourth, due to the setting in which the study occurred, a non-concurrent multiple-baseline design was used. Although non-concurrent multiple-baseline designs control for threats to internal validity (Harvey et al., 2004; Watson & Workman, 1981) and are common within applied research, a concurrent multiple-baseline design may be desired by future researchers. Finally, no measures of generalization were collected for in-person settings. These measures were not included for two primary reasons. First, due to the COVID-19 pandemic, increasing the number of people the participants contacted in person was deemed potentially harmful. Second, given the increase in the use of virtual environments, targeting social skills potentially useful in those environments may be useful in its own right given the increased use of communication technologies among young persons (Manago et al., 2020). Nonetheless, the results should be taken with caution with respect to the generalization of effects to in-person settings.

Despite the limitations, this study contributes to the limited literature on effective social skills interventions delivered via telehealth for autistics/individuals diagnosed with ASD. Given the uncertainty involved in what the intervention context will resemble in the future and the limited number of qualified professionals in some areas, effective, evidence-based interventions delivered via telehealth are more necessary than ever. We hope this study provides clinicians with an effective, evidence-based social skill intervention when shifting to telehealth-based instruction and targeting social skills for use in virtual environments and inspires future similar research to increase the number of treatment options available to interventionists.

Funding No grants were received to fund this project.

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the

institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

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

Conflict of interest The first, second, fourth, fifth, and six authors have written books and materials that describe the use of the procedure evaluated within this study. The royalties from the sales of those texts do not go to any of the authors.

References

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.).
- Au, A., Mountjoy, T., Leaf, J. B., Leaf, R., Taubman, M., McEachin, J., & Tsuji, K. (2016). Teaching social behaviour to individuals diagnosed with autism spectrum disorder using the Cool Versus Not Cool procedure in a small group instructional format. *Journal of Intellectual and Developmental Disability, 41*(2), 115–124. <https://doi.org/10.3109/13668250.2016.1149799>.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91–97. <https://doi.org/jaba.1968.1-91>.
- Carr, J. E. (2005). Recommendations for reporting multiple-baseline designs across participants. *Behavioral Interventions, 20*(3), 219–224. <https://doi.org/10.1002/bin.191>.
- Cihon, J. H., Ferguson, J. L., Milne, C. M., Leaf, J. B., McEachin, J., & Leaf, R. (2019). A preliminary evaluation of a token system with a flexible earning requirement. *Behavior Analysis in Practice, 12*(3), 548–556. <https://doi.org/10.1007/s40617-018-00316-3>.
- Constantino, J. N., & Gruber, C. (2012). *The Social Responsiveness Scale Manual, Second Edition (SRS-2)*. Los Angeles: Western Psychological Services.
- 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, 13*(2), 299–305. <https://doi.org/10.1007/s40617-020-00430-1>.
- Dotson, W. H., Leaf, J. B., Sheldon, J. B., & Sherman, J. A. (2010). Group teaching of conversational skills to adolescents on the autism spectrum. *Research in Autism Spectrum Disorders, 4*(2), 199–209. <https://doi.org/10.1016/j.rasd.2009.09.005>.
- Ferguson, J., Craig, E. A., & Dounavi, K. (2019). Telehealth as a model for providing behaviour analytic interventions to individuals with autism spectrum disorder: A systematic review. *Journal of Autism and Developmental Disorders, 49*(2), 582–616. <https://doi.org/10.1007/s10803-018-3724-5>.
- Ferguson, J. L., Majeski, M. J., McEachin, J., Leaf, R., Cihon, J. H., & Leaf, J. B. (2020). Evaluating discrete trial teaching with instructive feedback delivered in a dyad arrangement via telehealth. *Journal of Applied Behavior Analysis*. Advance online publication. <https://doi.org/10.1002/jaba.773>.
- Garcia-Albea, E., Reeve, S. A., Brothers, K. J., & Reeve, K. F. (2014). Using audio script fading and multiple-exemplar training to increase vocal interactions in children with autism. *Journal of Applied Behavior Analysis, 47*(2), 325–343. <https://doi.org/10.1002/jaba.125>.
- Gresham, F. M., & Elliott, S. N. (2008). *Social skills improvement system: Rating scales*. Bloomington, MN: Pearson Assessments.
- Harvey, M. T., May, M. E., & Kennedy, C. H. (2004). Nonconcurrent multiple baseline designs and the evaluation of educational systems.

- Journal of Behavioral Education*, 13(4), 267–276. <https://doi.org/10.1023/B:JOBE.0000044735.51022.5d>.
- Leaf, J. B., Milne, C. M., Leaf, J. A., Rafuse, J., Cihon, J. H., Ferguson, J. L., Oppenheim-Leaf, M. L., Leaf, R., McEachin, J., & Mountjoy, T. (2020). *The Autism Partnership method: Social skills groups*. Different Roads to Learning.
- Leaf, J. B., Mitchell, E., Townley-Cochran, D., McEachin, J., Taubman, M., & Leaf, R. (2016). Comparing Social Stories™ to Cool Versus Not Cool. *Education and Treatment of Children*, 39(2), 173–186.
- Leaf, J. B., Oppenheim-Leaf, M. L., Call, N. A., Sheldon, J. B., Sherman, J. A., Taubman, M., McEachin, J., Dayharsh, J., & Leaf, R. (2012a). Comparing the teaching interaction procedure to social stories for people with autism. *Journal of Applied Behavior Analysis*, 45(2), 281–298. <https://doi.org/10.1901/jaba.2012.45-281>.
- Leaf, J. B., Taubman, M., Leaf, J., Dale, S., Tsuji, K., Kassardjian, A., Alcalay, A., Milne, C., Mitchell, E., Townley-Cochran, T., Leaf, R., & McEachin, J. (2015). Teaching social interaction skills using Cool Versus Not Cool. *Child & Family Behavior Therapy*, 37(4), 321–334. <https://doi.org/10.1080/07317107.2015.1104778>.
- Leaf, J. B., Tsuji, K. H., Griggs, B., Edwards, A., Taubman, M., McEachin, J., Leaf, R., & Oppenheim-Leaf, M. L. (2012b). Teaching social skills to children with autism using the Cool Versus Not Cool procedure. *Education and Training in Autism and Developmental Disabilities*, 47(2), 165–175.
- LeBlanc, L. A., Lerman, D. C., & Normand, M. P. (2020). Behavior analytic contributions to public health and telehealth. *Journal of Applied Behavior Analysis*, 53(3), 1208–1218. <https://doi.org/10.1002/jaba.749>.
- Manago, A. M., Brown, G., Lawley, K. A., & Anderson, G. (2020). Adolescents' daily face-to-face and computer-mediated communication: Associations with autonomy and closeness to parents and friends. *Developmental Psychology*, 56(1), 153–164. <https://doi.org/10.1037/dev0000851>.
- Milne, C., Leaf, J. A., Leaf, J. B., Cihon, J. H., Torres, N., Townley-Cochran, D., Taubman, M., Leaf, R., McEachin, J., & Oppenheim-Leaf, M. (2017). Teaching joint attention and peer to peer communication using the Cool Versus Not Cool procedure in a large group setting. *Journal of Developmental and Physical Disabilities*, 41(2), 115–135. <https://doi.org/10.1007/s10882-017-9556-y>.
- Mohammadzahari, F., Koegel, L. K., Rezaee, M., & Rafiee, S. M. (2014). A randomized clinical trial comparison between pivotal response treatment (PRT) and structured applied behavior analysis (ABA) intervention for children with autism. *Journal of Autism and Developmental Disorders*, 44(11), 2769–2777. <https://doi.org/10.1007/s10803-014-2137-3>.
- Pellegrino, A. J., & DiGennaro Reed, F. D. (2020). Using telehealth to teach valued skills to adults with intellectual and developmental disabilities. *Journal of Applied Behavior Analysis*, 53(3), 1276–1289. <https://doi.org/10.1002/jaba.734>.
- Rudy, N. A., Betz, A. M., Malone, E., Henry, J. E., & Chong, I. M. (2014). Effects of video modeling on teaching bids for joint attention to children with autism. *Behavioral Interventions*, 29(4), 269–285. <https://doi.org/10.1002/bin.1398>.
- Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland-3: Vineland adaptive behavior scales* (3rd ed.). NCS Pearson, Inc.
- Stewart, K. K., Carr, J. E., & LeBlanc, L. A. (2007). Evaluation of family-implemented behavioral skills training for teaching social skills to a child with Asperger's disorder. *Clinical Case Studies*, 6(3), 252–262. <https://doi.org/10.1177/1534650106286940>.
- Wacker, D. P., Lee, J. F., Dalmau, Y. C. P., Kopelman, T. G., Lindgren, S. D., Kuhle, J., Pelzel, K. E., & Waldron, D. B. (2013). Conducting functional analyses of problem behavior via telehealth. *Journal of Applied Behavior Analysis*, 46(1), 31–46. <https://doi.org/10.1002/jaba.29>.
- Watson, P. J., & Workman, E. A. (1981). The non-concurrent multiple baseline across-individuals design: An extension of the traditional multiple baseline design. *Journal of Behavior Therapy and Experimental Psychiatry*, 12(3), 257–259. [https://doi.org/10.1016/0005-7916\(81\)90055-0](https://doi.org/10.1016/0005-7916(81)90055-0).
- Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence-fourth edition technical manual and interpretive manual* (4th ed.). Pearson Assessments.

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