Examination of the Effects of Auditory and Textual Stimuli on Response Accuracy and Latency during a Math Task and Tangram Puzzle

Tiffany Kodak¹ · Samantha Bergmann² · Brittany LeBlanc³ · Michael J. Harman⁴ · Maryam Ayazi⁵

Published online: 24 September 2018

© Association for Behavior Analysis International 2018

Abstract

Although Skinner (1957) provided a behavioral account of verbal thinking, additional research is needed to evaluate stimuli that may influence covert verbal behavior that occurs between the onset of a verbal stimulus and the emission of a response during an episode of verbal thinking. The present investigation examined the effects of auditory distractors and/or textual stimuli during arithmetic problems and tangram puzzles on the participants' response latency and accuracy. In addition, we measured and categorized occurrences of vocal verbal behavior during the response interval. In Experiments 1 and 2, the experimenter played auditory distractors during a proportion of arithmetic problems. In Experiment 2, the experimenter also presented a textual stimulus of the arithmetic problem. In Experiment 3, the experimenter played auditory distractors during a proportion of tangram puzzles. Results showed that auditory distractors led to longer response latencies and reduced accuracy in Experiment 1. The addition of the textual stimulus during trials in Experiment 2 improved accuracy and reduced differences in response latency when the auditory distractors were and were not present during the response interval. The auditory distractors during tangram puzzles in Experiment 3 produced no differential effects on accuracy or latency to respond.

Keywords Echoic behavior · Response latency · Verbal behavior · Verbal thinking

✉ Tiffany Kodak
tiffany.kodak@marquette.edu

¹ Department of Psychology, Marquette University, 525 N. 6th St, Milwaukee, WI 53203, USA

² Department of Behavior Analysis, University of North Texas, Denton, TX, USA

³ May Institute, Randolph, MA, USA

⁴ Department of Psychology, Briar Cliff University, Sioux City, IA, USA

⁵ Department of Psychology, University of Wisconsin-Milwaukee, Milwaukee, WI, USA

Skinner (1957) provided a behavioral account of thinking in which he described that a person may engage in various types of covert verbal behavior while completing an activity that requires verbal thinking. For example, Skinner described that to solve a novel mental arithmetic problem, we must infer that covert verbal behavior occurred between the initial statement of the problem (e.g., “What is 3 plus 10 plus 10 plus 5?”) and the individual’s response to the problem (e.g., “28”). Following the initial statement of the problem, a person might engage in covert verbal behavior such as, “Okay, 3 plus 10 plus 10 plus 5. Well, 3 plus 10 is 13, and 10 plus 5 is 15. So 13 plus 15 is 28.” This sequence of covert verbal behavior may account for the time that elapses (i.e., the latency) between the verbal stimulus of the problem and the person’s vocal response (“28”; Palmer, 2009).

Behavior analysts have rarely investigated covert behavior, focusing on observable behavior instead. Although observable behavior is easier to measure, can be subjected to comparisons of reliability to determine whether multiple observers agree on the occurrence of behavior, and may be a preferred subject of measurement for many of the research questions investigated by behavior analysts, the field of behavior analysis can benefit from evaluations of covert behavior in order to better predict and control overt behavior (Skinner, 1953).

When completing a task that likely requires covert verbal behavior, such as mental arithmetic, Skinner (1957) described several types of verbal operants that may occur. For example, a person may engage in covert echoic behavior by repeating the arithmetic problem (the person repeats “3 plus 10 plus 10 plus 5” to herself). Echoic behavior is defined as vocal behavior under the control of an immediately preceding auditory stimulus (Skinner, 1957). The consequence for echoic behavior is some form of generalized conditioned reinforcement (e.g., praise). A specific type of covert echoic behavior, called self-echoic behavior, also may occur during mental arithmetic. For example, a person might echo her own covert verbal behavior (e.g., she continually repeats “10 plus 5” while adding the first two numbers in the arithmetic problem). This behavior is described as self-echoic because the preceding auditory stimulus is the person’s own covert verbal behavior. Self-echoic behavior produces automatic reinforcement, such as assisting the person in engaging in behavior that produces its own reinforcement (e.g., echoing one’s own covert verbal behavior of a string of numbers that are subsequently entered into a phone to successfully call another person).

Although it is possible to identify numerous instances in which self-echoic behavior may be likely to occur during verbal thinking, there is a paucity of research on self-echoic behavior. Two studies examined the occurrence of self-echoic behavior in children (Esch, Esch, McCart, & Petursdottir, 2010; Esch, Mahoney, Kestner, LaLonde, & Esch, 2013). Esch et al. (2013) found that typically developing children engaged in more accurate self-echoic behavior than children diagnosed with autism spectrum disorder. Results of Esch et al. (2013) showed correct self-echoic responses varied from 20% to 100% of trials for typically developing children. Reductions in correct self-echoic behavior were noted following a 5-s delay. Nevertheless, additional research is needed to evaluate the conditions under which self-echoic behavior is likely to occur and to examine variables that might disrupt the occurrence of accurate self-echoic behavior.

Another verbal operant that may occur during verbal thinking is self-listener behavior. Skinner (1957) described that a person serves as his or her own listener when engaging in covert verbal behavior because this listener behavior leads to other behavior that produces its

own source of reinforcement. For example, when driving to a location that one doesn't visit often, a person might engage in covert verbal behavior by saying "turn left at the bank" to oneself, and then respond as a listener to this covert behavior by turning left when the bank is on the corner of the cross street. This self-listener behavior produces its own reinforcement by assisting the person in performing the necessary turns on surface streets, which is automatically reinforced by arriving at the correct location.

Although verbal behavior may occur at the covert level, "the line between public and private is not fixed" (Skinner, 1953, p. 282). Skinner (1957) described that covert verbal behavior may become overt under specific conditions, such as when engaging in overt behavior offers some advantage to the person. For example, if a person is in a noisy room and is trying to repeat a telephone number until the call can be placed, covert self-echoic behavior of the telephone number may become overt. Skinner (1957) states,

In the solution of a difficult problem, mathematical or otherwise, we resort to overt responses, vocal or written. For the same reason such covert behavior as counting money or adding figures is likely to become overt in the presence of distracting stimuli. (p. 436)

Thus, environments in which distracting stimuli occur may occasion overt verbal behavior during verbal thinking. Nevertheless, distracting stimuli should not occasion overt verbal behavior during tasks that do not require verbal thinking. For example, if a person is running on a treadmill, distracting stimuli (e.g., a loud vacuum, people talking) should not have an effect on running nor occasion overt verbal behavior. However, an examination of the conditions under which distracting stimuli occasion overt verbal behavior and alter performance on tasks that do and do not require verbal thinking is warranted.

The purpose of the present investigation was to evaluate Skinner's (1957) assertions regarding covert verbal thinking. Auditory distractors were presented during (a) a proportion of arithmetic problems that either did or did not include accompanying textual stimuli and (b) a tangram puzzle (i.e., a putative task that does not require verbal thinking). We evaluated whether covert verbal behavior (e.g., echoic and self-echoic behavior) would become overt more often in the presence of auditory distractors and whether this effect varied depending on the type of task. We also evaluated whether playing auditory distractors during arithmetic problems or tangram puzzles would increase the latency to respond to problems or puzzles and decrease accuracy in comparison to solving problems or puzzles in the absence of auditory distractors. Finally, we evaluated whether the presence of textual stimuli (i.e., the written math problem) during arithmetic problems would alter response accuracy, latency, and the occurrence of overt verbal behavior.

Method

Subjects, Setting, and Materials

The experiments included 115 undergraduate student participants; 45 participated in Experiment 1, 48 participated in Experiment 2, and 22 participated in Experiment 3. An

a priori power analysis was performed to estimate the necessary sample size. Due to a paucity of similar literature, we estimated that the experimental manipulations would result in a moderate effect. With an alpha of 0.05 and power of 0.80, the estimated sample size was approximately $N = 40$ for Experiments 1 and 2. All participants received one extra credit point for participation in the study, which was applied to the psychology course that the participant selected in an online research database. Refer to Table 1 for participant demographics. Sessions were conducted in a private room in a university-based laboratory. The room contained a table, chairs, and relevant study materials (e.g., cardboard divider, speaker, audio recording device).

The experimenter vocally presented 20 math problems (Experiments 1 and 2) or presented 20 tangram puzzles (Experiment 3). Each math problem was composed of one number that was between 20 and 29, one number between 10 and 19, and one number between 1 and 9. For example, the experimenter presented the problem “25 plus 13 plus 7.” Numbers were randomly selected by a computer program, and math problems were randomly assigned to trials. Textual stimuli presented on flash cards in Experiment 2 contained math problems printed in 72-point Calibri font on 10 cm × 15 cm index cards. Each tangram puzzle was an arbitrary shape composed of five colored pieces. Colored tangram puzzles were printed on 21.6 cm × 27.9 cm pieces of paper.

During 10 trials, the experimenter played a computer-generated auditory distractor file. The file contained numbers that ranged from 1 to 54, which corresponded to the

Table 1 Individual demographic data as a percentage of the sample

Characteristic	Experiment 1 (<i>n</i> = 45)	Experiment 2 (<i>n</i> = 48)	Experiment 3 (<i>n</i> = 22)
Sex			
Female	69	69	21
Male	29	31	1
Decline to answer	2	0	0
Age range			
18–24	87	79	21
25–34	11	19	1
35–44	2	0	0
45–54	0	0	0
55–64	0	2	0
Race/ethnicity			
Caucasian/White	58	54	11
African American/Black	9	23	3
Asian	22	10	5
Hispanic/Latino	7	10	3
American Indian/Alaska Native	2	0	0
Native Hawaiian/ other Pacific Islander	2	0	1
Decline to answer	0	2	0

All data are self-identified and self-reported

largest sum possible across all math problems included in the experiments. The auditory distractor file was a MP3 file played on a small speaker located on the tabletop that consisted of a female's voice that presented two numbers per second, played at 80 dB to approximate the volume of a loud conversation occurring in a public setting. Prior to the participant's arrival for a session, the experimenter confirmed the decibel reading by playing the auditory distractor file and measuring the decibel reading with a decibel reader placed at the approximate location of the participant's head while seated in the chair at the designated location at the table.

On the table in the session room, the experimenter placed tape to indicate the location at which the participant should sit in the chair in order to maintain a distance from the speaker that allowed the auditory distractor file to be played at approximately 80 dB. The participant was instructed to sit in the chair so that his or her body was within the taped lines. In addition, the participant was instructed to remove any fabric or materials that covered any portion of his or her ears. The experimenter placed a cardboard divider on the table at the start of the experiment so that the participant could not see the experimenter's data sheet, iPad® (for playing the sound file), session timers, or any other materials.

If the participant provided consent for audio recording, the experimenter placed a small audio recorder next to the participant on the table to capture the participant's vocal responses during the session. Each audio recording file was denoted by a randomly assigned participant number.

Response Measurement, Interobserver Agreement, and Treatment Integrity

Observers collected paper-and-pencil data on the dependent variables during each trial of a 20-trial session. The experimenters scored *latency to respond*, which was the amount of time from the presentation of the math problem or tangram puzzle to the participant's response. A *response* was scored when the participant stated his or her answer to the math problem or completed the tangram puzzle. If the participant provided more than one response to a math problem (e.g., "32, or maybe 33"), the experimenter recorded the first response. The responses to math problems were compared to an answer sheet following completion of the session to determine the *accuracy* of the participant's responses in each trial. The tangram puzzle was determined to be accurate when all colored pieces were placed in the exact location in the puzzle. *Accuracy* was calculated by dividing the total number of trials with a correct response by the total number of trials in the session. Two types of trials were presented (experimental and control). Thus, accuracy scores were calculated for trials in each trial type separately. We also calculated the mean percentage of correct responses (Experiment 3 only) by dividing the number of trials with a correct response in each condition (experimental or control) by the total number of trials, multiplied by 100. Data collectors also scored *vocalizations* if the participant engaged in any audible vocal verbal behavior that occurred between the experimenter stating the math problem or presenting the tangram puzzle and the participant providing his or her response. Data on vocalizations during the response interval were subsequently categorized according to predefined categories. The categories for math problems included (a) echoing numbers from the math problem stated by the experimenter, (b) saying numbers not from the math problem (e.g., engaging in addition of numerals), and (c) other vocal

verbal behavior. No categories were created for tangram puzzles due to the absence of vocal verbal behavior during trials.

The participants responded to an open-ended question following the completion of trials in Experiments 1 and 2 (described next). Participant responses to the open-ended question were recorded and subsequently placed into the following categories: (a) covert behavior performed by the participant related to numbers (e.g., “I tried to picture the numbers.”), (b) drawing or writing on the table (e.g., “I wrote the numbers.”), (c) using fingers to count (e.g., “I used my fingers to count.”), (d) attempting to ignore the auditory distractors (e.g., “I tried to block out the noise.”), (e) being distracted by auditory distractors (“I got distracted listening to the numbers.”), (f) auditory distractors interfering with the math problem (e.g., “Sometimes I confused your [experimenter] numbers for the recording.”), (g) reacting to sound (e.g., “The recording was stress provoking.”), (h) loss of stimulus control by the math problem (e.g., “It was really hard to remember all the numbers.”), (i) self-reported self-listener behavior (e.g., “I quietly talked to myself a little bit.”), and (j) engaging in covert self-echoic behavior for numbers (e.g., “I repeated the numbers under my breath.”).

A second observer independently collected data on all dependent variables during 98% of sessions in Experiment 1, 95% of sessions in Experiment 2, and 77% of sessions in Experiment 3. We calculated trial-by-trial interobserver agreement (IOA) by dividing the number of trials in which the observers recorded the same behavior by the total number of trials in a session and converted the ratio to a percentage. Agreement was scored for *latency* as long as the duration of time differed by no more than 2 s. Mean IOA for all dependent variables across conditions was 98% (range 65%–100%) in Experiment 1, 99% (range 85%–100%) in Experiment 2, and 99% (range 95%–100%) in Experiment 3.

We also calculated IOA for (a) 93% and 100% of the categorization of vocalizations of participants for whom overt vocal responses were recorded and (b) 100% and 98% of participants for whom responses to the open-ended question were reported in Experiments 1 and 2, respectively. An agreement was scored if two independent data collectors placed the participant’s response in the same category. Mean IOA in Experiment 1 for categorization of vocalizations was 85% (range 50%–100%). Mean IOA for categorization of responses to the open-ended question was 96% (range 85%–100%) in Experiment 1. Mean IOA for categorization of vocalizations in Experiment 2 was 86% (range 60%–100%). Mean IOA for categorization of responses to the open-ended question was 98% (range 85%–100%) in Experiment 2.

Observers also collected treatment integrity data. An instance of treatment integrity was scored in Experiments 1 and 2 if the experimenter (a) presented the math problem assigned to the trial; (b) presented the entire math problem within 2 s; (c) did not repeat any of the numbers in the math problem, once presented; (d) waited up to 1 min for the participant to respond; (e) initiated the auditory stimulus within 1 s of presenting the math problem, if relevant; (f) withheld any feedback following the participant’s response; (g) instructed the participant not to cover his or her ears, if relevant; and (h) presented the flash card within 0.05 s of the vocal presentation of the math problem (Experiment 2 only). An instance of treatment integrity was scored in Experiment 3 if the experimenter (a) presented the tangram puzzle assigned to the trial; (b) presented the instruction to complete the puzzle and removed the cover from the puzzle; (c) provided the correct puzzle pieces needed to complete the tangram puzzle; (d) waited

up to 1 min for the participant to respond; (e) initiated the auditory stimulus within 1 s of presenting the tangram puzzle, if relevant; (f) withheld any feedback following the participant's response; and (g) instructed the participant not to cover his or her ears, if relevant. The experimenter was required to perform all relevant trial behavior correctly to score an instance of treatment integrity during each trial.

To calculate treatment integrity, we divided the number of trials implemented with integrity by the total number of trials in a session and converted the ratio to a percentage. Mean treatment integrity data were 99% (range 95%–100%) for Experiment 1, 98% (range 85%–100%) for Experiment 2, and 99% (range 90%–100%) for Experiment 3.

Experimental Design

We evaluated the effects of the presentation of the auditory distractors on latency to respond and the accuracy of the participant's response in an alternating treatments design. The order of trials (experimental or control) was random. The same math problems were included in Experiments 1 and 2. The specific math problem assigned to a trial, trial type, and experiment was equated across participants. For example, a math problem included in a control trial for some participants was included in an experimental trial for other participants. The specific tangram puzzle assigned to a trial and trial type in Experiment 3 also was equated across participants. In addition, math problems and tangram puzzles were randomly ordered such that a math problem or puzzle presented in the first trial of an experiment for one participant was presented in different trials for other participants.

Experiment 1

Prior to initiating the session and after obtaining informed consent, the experimenter read instructions to the participant. The instructions were as follows:

I am going to read a math problem. I can only read the problem to you one time. Please do your best to solve the problem. Do not cover or plug your ears while you are trying to solve the problem. When you think you have the answer, please say the answer out loud. You have one minute to solve each problem. If you haven't solved the problem by the end of the minute, I will ask you to stop and I will present the next math problem. Do you have any questions before we begin?

The experimenter answered questions about the absence of a paper and pencil, the total number of math problems, and the option to take a bathroom break. If the participant asked questions about how to solve the problem (e.g., "Can I use my fingers to solve problems?"), the experimenter said that the participant could use whatever strategy he or she would like to solve the problem. The experimenter responded, "I can't answer that question," for all other questions.

A session consisted of 20 trials with 10 trials of each type presented in a session. Each participant completed one session. During each trial, the experimenter presented the math problem, started the session timer, waited up to 1 min for the participant to respond, and stopped the session timer once the participant responded or after 1 min

elapsed. If the participant did not respond within 1 min, the experimenter said “stop,” and initiated the next trial. Participants did not have a paper, pen, or any other materials to assist them in solving the math problems.

Following completion of the session, the experimenter asked the participant one open-ended question. The experimenter asked, “Did you do anything to make it easier for you to solve any of the problems?” and recorded the participant’s response to the question. After all responses were recorded, the experimenter thanked the participant for his or her participation and read the debriefing statement. Following debriefing, the experimenter told the participant that the extra credit point would be posted on the online database by the end of the day and walked the participant out of the session room.

Experimental Trials

The experimenter played the auditory distractor file at 80 dB while the participant solved the math problem. The experimenter turned off the auditory distractor file between trials.

Control Trials

The experimenter waited quietly for the participant to provide a response (i.e., the auditory distractor file was not played).

Experiment 2

The session procedures were identical to those described in Experiment 1 with one exception. The experimenter presented a textual stimulus (i.e., a math problem) on a flash card during each trial. That is, while the experimenter vocally presented the math problem, she presented a textual stimulus on an index card at the top of the cardboard divider with the math problem typed horizontally on the card. The textual stimulus remained present during the trial until the participant provided a response or 1 min elapsed. Thereafter, the experimenter flipped over the next card in the deck, a blank white index card, to cover the next addition problem during the intertrial interval. Textual stimuli were presented during all experimental and control trials.

Experiment 3

The session procedures were identical to those described in Experiment 1 with a few exceptions. The instructions read to participants were as follows:

I am going to give you a puzzle. Please do your best to complete the puzzle. Do not cover or plug your ears while you are trying to complete the puzzle. When you think you are done, please tell me that you are done. You have one minute to complete the puzzle. If you haven’t completed the puzzle by the end of the minute, I will ask you to stop and I will present the next puzzle. Do you have any questions before we begin?

During each trial, the experimenter placed a pile of puzzle pieces on the table along with a covered tangram puzzle. The experimenter removed the cover from the puzzle at the start of the trial and started the timer. Following completion of all trials, the experimenter did not ask the open-ended question.

Results

The results of all three experiments were analyzed in two ways. First, for Experiments 1 and 2, we conducted statistical analyses to measure differences between and within experiments in (a) the latency to respond, (b) the average response accuracy, and (c) the percentage of trials with vocalizations. The initial set of statistical analyses sought to determine any significant main effects and interactions for all three dependent variables in a separate 2 (Condition: control trials and experimental trials) $\times$ 2 (Experiment: 1 and 2) mixed two-factor ANOVA. Box's test for homogeneity was used to test for sphericity in variance between experiments. Four post hoc comparisons were conducted using independent-samples *t*-tests (across experiments) and repeated-measures *t*-tests (across conditions) to test for significant differences between conditions and experiments. For the independent-samples comparisons, Levene's test was used to measure any significant departures from equal variances. The Bonferroni alpha correction procedure was used to limit the probability of a Type I error during post hoc comparisons, $\alpha_{\text{pairwise}} = .0125$.

Statistical analyses of participants' latency to respond in Experiments 1 and 2 showed significant main effects of condition, $F(1, 86) = 10.67, p = .003, \eta_p^2 = .01$, and of experiment, $F(1, 878) = 129.79, p < .001, \eta_p^2 = .13$. Further, a significant interaction was measured for condition level by experiment, $F(1, 878) = 3.84, p = .050, \eta_p^2 = .004$. Bonferroni-corrected post hoc tests yielded significant changes in latency to respond within Experiment 1 from the control ($M = 11.85$) to the experimental trials ($M = 13.20$), $t(439) = -2.75, p = .006$. Thus, participants' latency to respond was longer during experimental trials in which the auditory distractors occurred than during control trials without the presence of auditory distractors. The same comparison was not significant ($p = .218$) in Experiment 2 ($M_{\text{Ctrl}} = 7.52, M_{\text{Exp}} = 7.80$); thus, latency to respond was consistent across experimental and control trials in which an index card that displayed the written math problem was present. Furthermore, latency to respond in the control trials in Experiment 1 ($M = 11.85$) was significantly longer than the response latency in the control trials in Experiment 2 ($M = 7.5$), $t(878) = 9.81, p < .001$. Similarly, the experimental trials in Experiment 1 produced significantly longer response times ($M = 13.20$) than the experimental trials in Experiment 2 ($M = 7.80$), $t(878) = 9.57, p < .001$. These results suggest that the presence of the textual stimulus in all trials of Experiment 2 reduced participants' latency to respond regardless of the presence of auditory distractors during the response interval.

Statistical analyses of the accuracy of responses showed significant main effects of condition, $F(1, 86) = 18.61, p < .001, \eta_p^2 = .18$, and of experiment, $F(1, 86) = 26.81, p < .001, \eta_p^2 = .24$. Bonferroni-corrected post hoc tests measured significant decrements in the mean accuracy of participants in Experiment 1 from the control trials ($M = .63$) to the experimental trials ($M = .50$), $t(43) = 3.92, p < .001$. The same comparison was made for changes in accuracy between trial types for Experiment 2 ($M_{\text{Ctrl}} = .81, M_{\text{Exp}} =$

.76), but changes were not significant ($p = .052$). Participants had significantly more correct responses during the experimental trials in Experiment 2 ($M = .76$) compared to Experiment 1 ($M = .50$), $t(86) = -5.45$, $p < .001$. Therefore, adding textual stimuli in Experiment 2 improved accuracy despite the presence of auditory distractors during response intervals.

The analyses for vocalizations in Experiments 1 and 2 showed that only the comparison by condition level yielded a significant main effect for differences in marginal means for the percentage of trials in which vocalizations occurred, $F(1, 86) = 4.49$, $p = .037$, $\eta_p^2 = .05$. The post hoc comparisons yielded no significant differences in the percentage of trials with vocalizations for Experiment 1 ($M_{\text{Ctrl}} = .14$, $M_{\text{Exp}} = .17$) and Experiment 2 ($M_{\text{Ctrl}} = .14$, $M_{\text{Exp}} = .19$). Nevertheless, our categorization of vocalizations, based on transcribed participant responses, yielded some noteworthy outcomes (see Table 2). In Experiment 1, 15 participants vocalized during 43% (64 out of 150 trials) of experimental trials and/or 36% (54 out of 150 trials) of control trials. Of the 64 experimental trials in which vocalizations occurred, 41% were echoic behavior (i.e., repeating numbers in the problem), 63% were saying numbers not in the problem (e.g., overt problem solving such as grouping numbers), and 27% were some other form of vocal verbal behavior. For the 54 control trials in which vocalizations occurred, 70% were echoic behavior, 44% involved saying numbers not in the problem, and 28% were some other form of vocal verbal behavior. In Experiment 2, 17 participants vocalized during 36% (61 out of 170 trials) of experimental trials and/or 31% (52 out of 170 trials) of control trials. Of the 61 experimental trials in which vocalizations occurred, 2% of vocalizations were echoic behavior, 82% were numbers not in the problem, and 26% were some other form of vocal verbal behavior. Of the 52 control trials in which vocalizations occurred, 6% were echoic behavior, 90% were saying numbers not in the problem, and 19% were some other form of vocal verbal behavior. Thus, less overt echoic behavior occurred during Experiment 2 in which textual stimuli remained present during the response interval, and a higher proportion of vocalizations were allocated to solving the problem (e.g., grouping numbers).

Statistical analyses for Experiment 3 included two repeated-measures t -tests that were conducted to determine the significance of participants' mean differences between control and experimental conditions for (a) the latency to complete tangram puzzles and (b) the accuracy of puzzle completion. For both measures (latency and accuracy), the

Table 2 Percentage of verbal behavior across trial type and experiment

Trial type	Verbal operant	Experiment 1 ($n = 15$)	Experiment 2 ($n = 17$)
Experimental		(42% of trials)	(36% of trials)
	Echoic	41%	2%
	Numbers not in problem	63%	82%
	Other vocal verbal behavior	27%	2%
Control		(36% of trials)	(31% of trials)
	Echoic	70%	6%
	Numbers not in problem	44%	90%
	Other vocal verbal behavior	28%	19%

mean differences between control and experimental conditions were nonsignificant ($p > 0.05$). Thus, presenting auditory distractors during a proportion of trials did not have an effect on participants' latency nor on the accuracy of completion of tangram puzzles. In addition, participants did not engage in any vocal verbal behavior during trials in Experiment 3.

The second type of analysis of our data involved visual inspection of scatterplots of data from Experiments 1 and 2. Figure 1 displays the participants' latency to respond in Experiments 1 and 2. Each data point shows a participant's latency to respond (s) during the experimental trials (y -axis) and control trials (x -axis). Data points above the diagonal line show participants who had shorter latencies to respond during the experimental trials, and data points below the line show participants for whom latency to respond was shorter during the control trials. In Experiment 1, participants had longer latencies to respond during the experimental trials than during the control trials. In contrast, participants did not appear to have differences in their latencies to respond across trial type in Experiment 2. Thus, the inclusion of the textual stimulus in Experiment 2 reduced participants' latency to respond regardless of the type of trial.

Figure 2 displays the accuracy of responding for each participant in Experiment 1 (top) and Experiment 2 (bottom). Each data point shows a participant's percentage of correct responses during the experimental trials (y -axis) and control trials (x -axis). Data points that fall on the diagonal line indicate that the participant's percentages of correct responses were identical across trial type. Data points above the diagonal line indicate a higher percentage of correct responses during the experimental trials, and data points below the diagonal line show a higher percentage of correct responses during the control trials. In addition, data points in the upper right portion of the graph (regardless

Fig. 1 The mean latency to respond (s) in experimental and control trials during Experiments 1 and 2

of whether the point is above or below the diagonal line) indicate a higher level of accuracy in responding than data points in the lower left portion of the graph.

The accuracy data for Experiment 1 (Fig. 2, top) show that participants engaged in a higher percentage of correct responses during control trials (i.e., trials conducted without the presentation of auditory distractors). However, approximately 50% of the participants had relatively low levels of accuracy because their mean percentage of correct responses was at or below 50% for one or both trial types. The accuracy data for

Fig. 2 The percentage of correct responses (accuracy) in experimental and control trials during Experiment 1 (top) and Experiment 2 (bottom). Numbers next to data points represent the number of participants whose data are shown in each point

Experiment 2 (Fig. 2, bottom) show an increase in the percentage of correct responding because all but three of the participants had percentages of correct responding that exceeded 50% in at least one type of trial. In addition, differences in accuracy across trial type appeared minimal because a similar number of data points were above and below the diagonal line. These results indicate that adding the textual stimulus in all trials in Experiment 2 substantially reduced differences in accuracy across trial type and increased participants' mean percentage of correct responding across both types of trials.

Figure 3 displays the participants' latency to respond to tangram puzzles in Experiment 3. Participants' latency to respond was similar in experimental and control trials, as shown by responding that fell on the diagonal line. Thus, unlike the results obtained when auditory distractors were presented during math problems in Experiment 1, the inclusion of auditory distractors during completion of tangram puzzles did not alter participants' latency to complete the puzzle. Similarly, accuracy data for Experiment 3 show that participants completed all of the tangram puzzles correctly in experimental trials and all but one puzzle correctly during control trials (Fig. 4). Therefore, presenting auditory distractors during experimental trials did not reduce accurate completion of tangram puzzles.

Data on the percentage of vocalizations during experimental and control trials (Fig. 5) show that participants had similar levels of vocalizations across trial types in Experiments 1 and 2. Thus, participants were as likely to vocalize during the response interval in the experimental trials as they were during control trials. The presence of the auditory distractors did not appear to influence the occurrence of vocalizations during the response interval, although the topography and function of vocalizations differed across experiments. Participants engaged in echoic behavior more often in Experiment 1, in which no textual stimuli were present during trials, in comparison to Experiment 2, in which textual stimuli of math problems were present. These data suggest that engaging in overt echoic behavior conferred an advantage to the participant in the absence of textual stimuli during completion of math problems. In contrast, participants did not engage in any vocal verbal behavior during Experiment 3 while completing tangram puzzles. The absence of any vocal verbal behavior during Experiment 3 suggests that completion of tangram puzzles may not require verbal behavior.

Fig. 3 The mean latency to respond (s) in experimental and control trials during Experiment 3 with tangram puzzles

Fig. 4 The mean percentage of trials correct during control and experimental trials for tangram puzzles in Experiment 3

Participants' responses to an open-ended question about whether they engaged in any behavior (covert or overt) to help them solve the math problems in Experiments 1 and 2 showed that participants' responses fell into certain categories that relate to the experimental procedures and outcomes. For example, 34% of participants reported that they experienced a loss of stimulus control by the math problems (e.g., they could not remember the numbers in the problem), they had a reaction to auditory distractors, they tried to ignore or block out auditory distractors, solving the problem was easier without auditory distractors present, or auditory distractors interfered with their completion of math problems. Participants also reported covert behavior that may have influenced their responding, such as self-echoic behavior (12% of participants), visualizing the numbers (15% of participants), and trying to engage in self-listener behavior (13% of participants). Finally, participants reported overt behavior on which we did not collect data. For example, 28% of participants said that they used their fingers to count, and 6% of participants said that they wrote on the table with their fingers to assist in completing math problems. The experimenters' anecdotal observations confirmed that

Fig. 5 The percentage of trials with vocalizations in experimental and control trials during Experiments 1 and 2

some participants did count with their fingers or use their fingers to write on the table during the response interval, which is consistent with the participants' self-reports of behavior.

Discussion

The current investigation found that participants had lower levels of accuracy and longer latencies to respond during arithmetic problems when auditory distractors containing numbers were played during the response interval during experimental trials in Experiment 1. However, response accuracy was also relatively low for the control trials in Experiment 1. When a textual stimulus of the math problem was presented throughout the response interval in Experiment 2, accuracy increased and latency to respond decreased for both trial types. In addition, differences in accuracy and latency across trial types were no longer significant when the textual stimulus was added to trials in Experiment 2. Auditory distractors played during completion of tangram puzzles in Experiment 3 did not have an effect on latency to complete puzzles nor accuracy of puzzle completion. The function of vocalizations varied across experiments; more echoic behavior occurred when no textual stimuli were present in Experiment 1. Thus, the absence of the textual stimulus created a condition under which putative covert echoic behavior became overt.

The present study sought to arrange conditions in the environment to test Skinner's (1957) assertions regarding verbal thinking, self-listener, and overt behavior during verbal thinking. For example, Skinner described that the covert behavior that occurs during verbal thinking may become overt if there is an advantage to be gained by the speaker. Skinner described that an individual who is reading a challenging text in a noisy environment may covertly read a lengthy sentence several times. If the individual is not able to respond to his or her own self as a listener because of the noise level in the environment, the individual may engage in overt reading. In Experiment 1, we played auditory distractors during half of the math problems to simulate conditions under which there may be an advantage to engage in overt behavior that was previously covert, and we measured the occurrence of vocalizations across conditions with and without auditory distractors. The results of Experiment 1 showed that participants were not more likely to vocalize in the presence of the auditory distractors, but vocalizations occurred across trials. Nevertheless, echoic behavior occurred more frequently during Experiment 1 in which no textual stimuli were present. That is, the participant was more likely to overtly echo the math problem presented by the experimenter and engage in self-echoic behavior to maintain the numbers in the math problem when no other supplemental stimuli (e.g., the written math problem) were available. Therefore, we identified a condition under which putative covert echoic behavior becomes overt. Our data collection did not permit an analysis of whether engagement in overt echoic and self-echoic behavior increased accuracy, and this is an analysis that could be included in future investigations.

We also evaluated Skinner's (1957) statements regarding conditions that might disrupt self-echoic behavior or responding to oneself as a listener during verbal thinking by measuring response accuracy and latency to respond. These measures allowed us to evaluate whether the auditory distractors presented in experimental trials

might produce a sufficient disruption to responding. Our results showed that the auditory distractors did increase participants' latency to respond in comparison to control trials in Experiment 1. Although we cannot identify the specific variables that may have influenced latency to respond, we systematically manipulated several variables to determine the impact on response latency in the present investigation. In Experiment 1, only the presence of the auditory distractors was manipulated across trials. Results showed that this stimulus led to longer latencies to respond. These results replicate other studies that examined the effect of auditory distractors on responding (Slowiaczek & Clifton, 1980). Slowiaczek and Clifton (1980) had participants engage in irrelevant vocal verbal behavior (e.g., counting) while reading a passage. The authors evaluated whether blocking what they described as subvocalizations with the repetitive vocalizations during passage reading would influence measures of reading comprehension. The results showed that the subvocalization-blocking procedure led to significant reductions in scores on measures of comprehension. Although the authors did not provide a behavior-analytic conceptualization of their data, the outcomes of their study align with those of the current investigation. Both studies showed that auditory stimuli generated by the individual or occurring in the environment have an effect on related behavior, and together, these findings support Skinner's assertions regarding the benefit of responding to oneself as a listener on verbal behavior.

Our measure of participant vocalizations during the math task permitted an examination of overt verbal thinking. That is, a proportion of participants did engage in overt echoic and self-echoic behavior, said numbers not directly in the problem while solving it (e.g., grouping numbers, partial answers when adding portions of the math problem), and emitted other vocal verbal behavior during the response interval. Although participants were not instructed to speak aloud while solving the problem, nor did they vocally solve the entire problem aloud, aspects of their behavior are consistent with Skinner's (1957) description of verbal thinking. Future studies could instruct participants to solve the problem aloud to measure and more fully evaluate the vocal verbal behavior that occurs between the presentation of the math problem and the participant's emission of the correct answer.

The auditory distractors in the present investigation included numbers that were different from the numbers in the math problem but played at a volume (80 dB) that may disrupt self-listener behavior. It is possible that the numbers presented as auditory distractors led to echoic responding that interfered with the emission of the correct response. For example, on several occasions, participants emitted a response to the math problem within 2 s of the same number that was presented in the auditory distractor (e.g., participant said "24" within 2 s of "24" being presented within the auditory distractor). Echoic behavior under faulty sources of control (i.e., under the control of extraneous auditory stimuli rather than the individual's own covert verbal behavior) reduced accuracy. Our results are similar to those of Nosik, Williams, Binder, and Carr (2017), who found letters presented as visual or auditory distractors during a paired-letter task had the greatest effects on responding in comparison to unrelated stimuli (e.g., unrelated shapes, muffled talking). The authors noted that the presentation of letters as distractors may have disrupted covert verbal behavior that occurred during the paired-letter task. Nevertheless, in the present investigation, accuracy was significantly higher during experimental trials in Experiment 2 in which multiple sources of control were available (i.e., textual stimulus, experimenter's vocalization of the math

problem, participant's self-echoic and self-listener behavior), even though the same auditory distractor file played during response intervals. The presence of the textual math problem during trials may have reduced the likelihood of the emission of behavior under faulty and irrelevant sources of stimulus control (e.g., echoing auditory distractors) and improved accuracy in responding. However, echoing auditory distractors was not a dependent variable in the present study, and the influence of this behavior on responding across trials and experiments remains unknown. Additional research is needed to evaluate sources of faulty stimulus control on responding and the effects of distractors that are similar or dissimilar to the experimental task.

The tangram puzzles in Experiment 3 served as a control condition for examination of the effects of auditory distractors on the performance of different tasks. To study verbal thinking as described by Skinner (1957), we arranged one task that may occasion verbal thinking and another that may not. We hypothesized that novel math problems occasioned covert verbal behavior to respond accurately. For example, answering "23 plus 15 plus 7" correctly, in the absence of previous practice, may evoke covert verbal behavior (e.g., echoing parts of the problem, adding numbers together). In contrast, tangram puzzles (as arranged in our study) may not require verbal thinking. For example, placing each puzzle piece involves a visual match-to-sample task, and nonhuman animals (who may not display verbal behavior; refer to Leigland, 1997, for a discussion of this topic) can accurately complete visual match-to-sample tasks (e.g., Frank & Wasserman, 2005; Iversen, Sidman, & Carrigan, 1986). The results of Experiments 1 and 3 show that auditory distractors had differential effects on behavior, depending on the task. That is, auditory distractors were not simply disruptive to the completion of all types of tasks. Rather, only during tasks in which putative covert verbal behavior occurs during the interval between the discriminative stimulus (i.e., a math problem) and the response (i.e., the answer to the math problem) did the presentation of auditory distractors increase the latency to respond, decrease the accuracy of a response, and/or evoke overt verbal behavior. Thus, Experiment 3 lends support to the differential effects of distracting stimuli on specific tasks for which verbal thinking may or may not occur.

An alternative account of our results is that auditory distractors in Experiments 1 and 2 may have weakened stimulus control in addition to or rather than aspects of verbal thinking (e.g., self-echoic and self-listener behavior). The presentation of auditory distractors may have weakened the stimulus control of the vocally presented math problem in Experiment 1 by presenting alternative numbers that could have acquired faulty stimulus control over responding. In other words, this weakened or loss of stimulus control might be described as not "remembering" the problem, which was reported by 34% of participants in response to the open-ended question at the conclusion of the experiment. This weakening of stimulus control could be resolved by presenting a textual stimulus of the math problem, which occurred in Experiment 2. The results of Experiment 2 show that participants' accuracy increased, latency to respond decreased, and the occurrence of echoic/textual behavior decreased in the presence of a textual stimulus. Thus, the manipulation of textual stimuli, rather than auditory distractors, may have had an effect on verbal behavior. Nevertheless, one purpose of our study was to identify the condition(s) under which covert behavior may become overt during verbal thinking. Our measure of vocalizations in the presence and absence of the textual stimuli in Experiments 1 and 2 shows that some participants

engaged in overt vocal verbal behavior, and the type of vocalizations changed across experiments. Thus, although the presence of textual stimuli may have altered performance more so than the auditory distractors, participants' responding in the presence and absence of textual stimuli suggests that we were able to examine conditions that led to overt vocal verbal behavior.

One limitation of the current investigation is that we are unable to verify that verbal reports from the participants matched their performance. That is, if a participant reported that auditory distractors hindered his or her performance, we did not compare the latency or accuracy of responses across experimental and control trials for that individual. Verbal reports may be influenced by rule-governed or contingency-shaped behavior and can be inaccurate (Catania, Matthews, & Shimoff, 1982; Hayes, White, & Bissett, 1998). Although there are methods for verifying the accuracy of verbal reports of behavior (e.g., the "silent dog" method; Hayes et al., 1998), we did not use these methods in the present study. Future studies could examine whether participants' verbal self-reports of their verbal thinking during math problems with and without auditory distractors align with their performance to further examine the accuracy of these reports during math problem completion. For example, participants could be instructed to think aloud during their completion of math problems to measure the occurrence of echoic behavior, self-echoic behavior, and other relevant behaviors during math problem completion.

Another limitation of the present investigation is that we did not directly manipulate the presence or absence of echoic, self-echoic, self-editing, or self-listener behavior during the experiments. Rather, we manipulated variables that may influence the occurrence of these behaviors. It might be possible to block the occurrence of these behaviors during completion of math problems by having participants sing a song or engage in other competing, overt verbal behavior. Research examining the effects of blocking these behaviors during completion of math problems could produce further evidence for the occurrence of these covert verbal operants during verbal thinking.

Measures of echoic behavior during Experiment 2 cannot be distinguished from textual responses. That is, because a textual stimulus of the math problem was present during all trials in Experiment 2, participants' responses that included repeating the math problem may have been under textual and/or echoic control. Nevertheless, higher levels of echoic behavior were observed in Experiment 1 in which no textual stimuli were present. Thus, echoing the math problem was more likely when no textual stimuli were present and only a transitory auditory stimulus of the math problem occurred at the onset of trials.

We hypothesized that the manner in which we presented the tangram puzzle in Experiment 3 allowed for a comparison of the effects of auditory distractors on tasks that vary in relation to verbal thinking. The tangram puzzle in Experiment 3 included an outline of an arbitrary shape composed of five colored pieces. Thus, participants matched each colored puzzle piece to the corresponding location in the puzzle outline. We hypothesized that this visual-visual match-to-sample task would not require verbal thinking, although the determination of verbal thinking tasks is an empirical question. It is possible that an alternative arrangement for the tangram puzzle could have altered whether this task would be characterized as a verbal thinking task. For example, if the experimenter briefly presented a picture of the tangram puzzle and then removed it, and participants did not have an outline that showed where each colored piece should be

placed, it is possible that the tangram puzzle may become a task that requires verbal thinking. Thus, aspects of the task may alter whether the task requires verbal or nonverbal thinking, and this is an avenue worthy of further investigation.

Finally, we did not evaluate the effects of different types of auditory distractors on responding to math problems and tangram puzzles. We played auditory distractors composed of numbers during a proportion of math problems and tangram puzzle trials. The results showed that auditory distractors disrupted performance during completion of math problems but not during the completion of tangram puzzles. It is possible that different auditory distractors (e.g., loud tones, environmental sounds) or visual distractors (e.g., other numbers) also may disrupt performance on math problems, although the extent to which specific distracting stimuli affect performance requires further investigation (Nosik et al., 2017). Nevertheless, the disruptive effects of other stimuli on responding to math problems would not negate the occurrence of covert verbal thinking. Rather, studies that investigate the effects of various auditory distractors on responding may help elucidate the results of the current study and could lead to the discovery of specific sources of faulty stimulus control of overt and covert verbal behavior.

Acknowledgements We thank Leah Bohl, Sarah Farhan, Ella Gorgan, Michelle Helms, Zhanxu Liu, Theresa Mayland, and Stephanie Zettel for their assistance with data collection.

Compliance with Ethical Standards

Conflict of Interest All authors declare that they have no conflicts of interest.

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 participants included in the study.

References

- Catania, A. C., Matthews, B. A., & Shimoff, E. (1982). Instructed versus shaped human verbal behavior: Interactions with nonverbal responding. *Journal of the Experimental Analysis of Behavior*, 38, 233–248. <https://doi.org/10.1901/jeab.1982.38-233>
- Esch, J. W., Esch, B. E., McCart, J. D., & Petursdottir, A. I. (2010). An assessment of echoic behavior in young children. *The Analysis of Verbal Behavior*, 26, 3–13.
- Esch, J. W., Mahoney, J. M., Kestner, K. M., LaLonde, K. B., & Esch, B. E. (2013). Echoic and self-echoic responses in children. *The Analysis of Verbal Behavior*, 29, 117–123.
- Frank, A. J., & Wasserman, E. A. (2005). Associative symmetry in the pigeon after successive matching-to-sample training. *Journal of the Experimental Analysis of Behavior*, 84, 147–165. <https://doi.org/10.1901/jeab.2005.115-04>
- Hayes, S. C., White, D., & Bissett, R. T. (1998). Protocol analysis and the “silent dog” method of analyzing the impact of self-generated rules. *The Analysis of Verbal Behavior*, 15, 57–63. <https://doi.org/10.1007/BF03392923>
- Iversen, I. H., Sidman, M., & Carrigan, P. (1986). Stimulus definition in conditional discriminations. *Journal of the Experimental Analysis of Behavior*, 45, 297–304. <https://doi.org/10.1901/jeab.1986.45-297>
- Leigland, S. (1997). Is a new definition of verbal behavior necessary in light of derived relational responding? *The Behavior Analyst*, 20, 3–9. <https://doi.org/10.1007/BF03392757>

- Nosik, M. R., Williams, W. L., Binder, C., & Carr, J. E. (2017). Methodological refinements in the behavior-analytic study of distraction: A preliminary investigation. *European Journal of Behavior Analysis*, 18, 71–83. <https://doi.org/10.1080/15021149.2016.1230967>
- Palmer, D. C. (2009). The role of private events in the interpretations of complex behavior. *Behavior and Philosophy*, 37, 3–19.
- Skinner, B. F. (1953). *Science and Human Behavior*. New York: The Free Press.
- Skinner, B. F. (1957). *Verbal Behavior*. Englewood Cliffs: Prentice-Hall.
- Slowiaczek, M. L., & Clifton, C. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior*, 19, 573–582.