Evaluation of Verbal Behavior in Older Adults

Amy C. Gross, R. Wayne Fuqua, and Todd A. Merritt
Western Michigan University

Approximately 5% of older adults have a dementia diagnosis, and language deterioration is commonly associated with this disorder (Kempler, 2005). Several instruments have been developed to diagnose dementia and assess language capabilities of elderly adults. However, none of these instruments take a functional approach to language assessment as described by Skinner (1957). The purpose of this study was to develop and evaluate a function-based assessment for language deficits of older adults. Thirty-one participants were categorized into a control group ($n = 15$) and a dementia group ($n = 16$) based on their score on the Dementia Rating Scale-2. Individuals with dementia performed significantly worse on the tact assessment than those without dementia. Participants from both groups performed better on measures of tacts than intraverbals or mands, even though topographically identical responses were required in these assessments. The data provide support for Skinner's conceptualization of functionally independent verbal operants.

Key words: older adults, dementia, language assessment, verbal behavior assessment

---

Adults 65 years and older make up approximately 13% of the U.S. population. Projections indicate that this age group will reach over 16% in 2020 and over 20% in 2050 (Vincent & Velkoff, 2010). In 2000, 10.8% of older adults reported that they had a mental disability, defined as difficulty with learning, remembering, or concentrating (Waldrop & Stern, 2003). Approximately 5% of adults over 65 years of age have some form of dementia (Kempler, 2005).

Dementia is a condition characterized by memory loss and at least one other area of cognitive dysfunction, including aphasia, apraxia, agnosia, or disturbed executive functioning (American Psychiatric Association, 2000). Estimates indicate that the number of adults with dementia diagnoses worldwide will grow from 25 million in 2000 to 63 million in 2030 (Kempler, 2005). With increasing numbers of older adults diagnosed with dementia, developing appropriate assessment and treatment methods is of critical importance.

Aphasia, one of the cognitive dysfunctions included in the dementia diagnosis, is described as "deterioration of language function" (American Psychiatric Association, 2000, p. 148). Although changes in language abilities are not unique to those diagnosed with dementia (Benson & Ardila, 1996; Kempler, 2005), research suggests that individuals with dementia diagnoses perform worse on verbal tasks than similar-aged peers without dementia diagnoses (Benson & Ardila, 1996). Early and adequate assessment of language deterioration of older adults is important, because determining the degree of language deterioration relative to peers helps with the diagnosis of dementia and may allow early intervention for language deficits associated with dementia.

Traditional assessments of language focus on response topographies without addressing the function of the responses. Although topographical information may be useful for characterizing language deficits, it does not explain the cause of the deficits. Topographical information alone is insufficient for designing most effective interventions (Esch, LaLonde, & Esch, 2010). Therefore, developing assessment methods that identify the function of verbal responses may be necessary for improving interventions for older adults with deteriorating language skills.

Skinner (1957) offered a conceptual approach to the analysis of language that viewed language as operant behavior, behavior that is influenced by a range of contextual features and maintained by its consequences. He proposed a classification system of verbal operants, each controlled by a unique set of antecedents and consequences. The primary verbal operants identified by Skinner and addressed in the present study include echoics, tacts, intraverbals, and mands.

Skinner (1957) proposed that each verbal operant is functionally independent of the others, meaning that a specific response may be acquired under one set of controlling conditions, but topographically identical responses may not occur under other controlling conditions. He also proposed a functional approach to language deterioration stating, ''What has been damaged in aphasia is clearly the functional control of behavior, and the damage respects the line of control'' (p. 195). This suggests that a response that was previously in an individual's repertoire might continue to occur under certain controlling conditions, but that same response might no longer occur under other conditions. For example, a person who had used the word ''spoon'' correctly across contexts in the past may be able to currently name a photo of a spoon, yet have difficulty finishing the phrase, ''you eat with a knife, fork and _____.'' Skinner suggested that individuals can have deficits specific to speaking or listening, and he added specificity by predicting that speaker behavior might degrade in a manner that is compatible with the distinction between verbal operants. Specifically, ''The order of damage seems to follow the order of 'difficulty' deducible from the availability of a minimal repertoire. Textual and echoic behavior often survive . . . while intraverbals and tacts appear to be most vulnerable'' (Skinner, 1957, p. 219). Additionally, Skinner noted that ''Damage is usually most severe in verbal behavior receiving generalized reinforcement'' (p. 219). Mands are not reinforced with generalized reinforcers; instead, the controlling conditions for mands involve motivating operations and characteristic consequences (Michael, 1988). Therefore, it is possible that mands are more resilient against decline than other verbal operants, such as tacts or intraverbals.

Researchers have attempted to evaluate the functional independence of verbal operants as it relates to acquisition and deterioration of verbal behavior. Many studies report data consistent with functional independence during verbal skill acquisition, whereas other studies reveal transfer of training across verbal operants (for a review, see Sautter & LeBlanc, 2006). Notably, there are procedural variations across studies evaluating functional independence, which might account for these differences in findings. With regard to deterioration, a rehabilitation study for adults who had sustained a traumatic brain injury (TBI) (Sundberg, San Juan, Dawdy, & Arguelles, 1990) revealed that verbal operants were differentially affected post-injury; that is, echoics and textuals were strong but tacts, mands, and intraverbals were weak. Furthermore, these adults acquired verbal operants after their injury in a different manner than children with developmental disabilities, who were presumably acquiring the skills for the first time. Prior research had shown that direct mand training may be necessary for children with developmental disabilities to learn mands, and these children learned mands quicker than tacts. However, adults who had a sustained a TBI acquired tacts and intraverbals quicker than mands, and direct mand training was less effective for teaching mands than transfer of training from tacts or intraverbals. These findings may suggest that unique assessment and intervention strategies are needed for sophisticated versus naïve speakers, as well as for speakers who are losing, as opposed to acquiring, verbal skills.

To the authors' knowledge, researchers have not yet developed and evaluated language assessment tools for age-related language deficits based on Skinner's functional verbal behavior model. Baker, LeBlanc, and Raetz (2008) provided a taxonomy of aphasia and recommendations for how to assess verbal behavior in older adults based on a functional approach; however, empirical data were not provided. Other researchers have evaluated verbal behavior interventions with older adults with some success (e.g., Dixon, Baker, & Sadowski, 2011), but the interventions were not derived from a functional assessment. The Verbal Behavior Milestones Assessment and Placement

Program (VB-MAPP; Sundberg, 2008) is a language assessment measure that is based on Skinner's model of verbal behavior. It is used to assess language delays in young children, and the results of the assessment lead to specific intervention plans. A similar assessment model is needed for older adults who might be losing language skills. If such an assessment reveals functionally independent verbal operant deficits in older adults, then the next phase of research can be undertaken; the creation and evaluation of function-based interventions.

The current research is a first step in this process; the development and preliminary evaluation of a function-based assessment for language deficits in older adults with and without dementia. There were two primary research questions. First, in what ways do verbal behavior strengths and deficits differ between older adults with and without dementia? Second, does language deteriorate differentially across the verbal operants?

METHOD

Participants and Setting

Thirty-one individuals, aged 60 years and older, participated in the study. Researchers recruited participants from a neuropsychology clinic and a senior living facility, and sessions were conducted in their respective locations. Based on results from the Dementia Rating Scale-2 (DRS-2), which was administered during the first session, 15 participants were assigned to the control group, and 16 participants met criteria for the dementia group. The DRS-2 measures attention, initiation/perseveration, construction, conceptualization, and memory (Jurica, Leitten, & Mattis, 2001) and is used to screen and track cognitive status in older adults (Schmidt, Mattis, Adams, & Nestor, 2005). For the purposes of this study, individuals who scored in the mildly, moderately, or severely impaired range were categorized into the dementia group. Participants scoring in the below average, average, and above average range were classified as ''intact'' by the DRS-2 manual (Jurica et al., 2001) and were placed in the control group. Participants in the control group had significantly higher DRS-2 scores ($M = 10.27$, $SE = .396$) than

participants in the dementia group ($M = 6.44$, $SE = .376$), $t(29) = 7.02$, $p < .001$, which means that the dementia group showed more impairment than the control group. No participant fell in the severely impaired range on the DRS-2. This, in addition to observations and interactions during sessions, suggests that all participants were able to understand and engage in the tasks presented. All participants were Caucasian. The entire sample was made up of 7 males and 24 females, with 4 males in the control group and 3 in the dementia group. Participant ages ranged from 60 to 86 years ($M = 72.9$, $SE = 2$) in the control group and 61 to 84 years ($M = 73.9$, $SE = 1.97$) in the dementia group. The mean age of the two groups did not differ significantly $t(29) = -.36$, $p = .72$.

The study was approved by Western Michigan University's Human Subjects Institutional Review Board. Although there was an option for guardian consent and participant assent, all participants independently consented to participating in the study.

Measures

Short-form Boston Naming Test (BNT). The BNT contains 60 line-drawn pictures and it measures individuals' ability to name the items depicted in the drawings (Mitrushina, Boone, Razani, & D'Elia, 2005). For this study, we used a validated short-form of the test by administering only the even-numbered items (Fisher, Tierney, Snow, & Szalai, 1999). Because this BNT is a standard assessment often used to assess cognitive and/or language impairment, it was administered to all participants, and the results were compared to the results from the other assessment measures.

Verbal Behavior Assessment Battery (VBAB). For the current evaluation, we developed materials and procedures to assess echoics, tacts, intraverbals, and mands, as detailed in Skinner's conceptual model of verbal behavior. Materials and procedures were also developed to assess listener behavior; similar to traditional receptive language tasks. To evaluate the variables controlling verbal behavior, Sidman, Stoddard, Mohr, and Leicester (1971) recommended that evaluators hold the stimulus or the response constant across testing conditions. Therefore, the odd-numbered items on

the BNT, a previously validated measure, were incorporated into five assessment measures (listener behavior, echoic, tact, intraverbal, and vocal and selection-based mand assessment) by holding the stimulus items constant and varying the controlling variables to match those that characterize each verbal operant (see Appendix A for stimuli and Appendix B for detailed assessment instructions). The order of the 30 items was randomized across assessment measures. An additional mand assessment, the 3D mand assessment, contained five items that differed from the 30 stimuli used for the other assessment measures.

Listener Behavior Assessment. The investigator placed a set of three line drawings on the table and asked the participant to point to a certain item; an example of a listener response. The three drawings were displayed horizontally on a 10.8 × 27.9 cm piece of card stock. The three drawings were scaled and distributed equally with two vertical lines separating them. This assessment is comparable with traditional receptive language tasks (Guess & Baer, 1973). We assessed listener responses for all 30 odd-numbered BNT items. A response was scored as correct when the participant pointed to the appropriate item within 5 s of the presentation of the drawings.

Echoic Assessment. The investigator asked the participant to repeat 30 words, one at a time. A correct response occurred when the participant vocally repeated the word provided by the investigator within 5 s of the auditory model. The 30 words were the names of the odd-numbered items on the BNT.

Tact Assessment. The investigator presented the odd-numbered items from the BNT (each consisting of a simple line drawing of an object) and asked the participant to name each item. The response was scored as correct if the participant initiated a vocal response within 5 s of the presentation of the drawings and accurately named the stimulus item within 10 s of presentation.

Intraverbal Assessment. The investigator provided a brief description of each of the 30 odd-numbered items from the BNT, without showing a picture of the item, and asked the participant to name the item described. Item descriptions were developed by using definitions from the *Longman Advanced American Dictionary* (2000) and removing words that were the same or highly similar to the item name (e.g., removing ''hang'' from the description of a ''hanger''). A response was scored as correct if the participant initiated a response within 5 s of the completion of the description and named the item within 10 s.

Vocal and Selection-Based Mand Assessment. The investigator showed the participant 30 pairs of scenes. The first scene of each pair was complete; the second scene was the same but with an item missing. The scenes were line drawings, and each scene included an odd-numbered item from the BNT, which looked the same as it did in the other assessments that had visual stimuli (e.g., tact assessment). Each scene was printed on a standard sheet of card stock (i.e., 21.6 × 27.9 cm) and consisted of four to seven black and white line drawings, including the target item. For example, the first scene for the target item ''pencil'' depicted a table with a stapler, stack of three books, and the pencil. The second scene that was shown had a question mark in place of the pencil. The nontarget items in each scene were items not included in the BNT, and thus, not included as a target item in any of the trials. For each trial, the investigator presented the first complete scene for 3 s, covered it, and then presented the second scene with the missing item. For the vocal mand assessment, the investigator instructed the participant to request the missing item to complete the scene (i.e., the investigator asked, ''what do you need to complete this picture?''). The participant was given 5 s to initiate a response and 10 s to complete a response. Upon requesting the appropriate item, the investigator gave the participant the line drawing, and the participant placed it in the appropriate location. If the participant did not request the correct (i.e., missing) item given the original vocal mand assessment procedures, the investigator administered the selection-based mand assessment protocol. That is, the investigator showed the participant a set of three line drawings and instructed the participant to point to the item that belonged in the missing space. The participant was given 5 s to initiate and 10 s to complete a response. If the participant pointed to the correct picture, the investigator

gave the participant the line drawing to place in the scene. If the participant did not point to the correct picture in the selection-based mand assessment, the next item in the vocal mand assessment was presented. Thus, a selection-based mand assessment trial was implemented only when an incorrect response occurred for an item in the vocal mand assessment. Each response (i.e., vocal and selection-based) was scored separately as correct or incorrect. If the vocal mand was correct, the selection-based assessment protocol was not administered and both responses were recorded as correct. Thus, two scores were assigned each time the assessment was administered; a vocal mand assessment score and a selection-based mand assessment score.

3D Mand Assessment. The 3D mand assessment was used to probe mands for three-dimensional objects, because it was thought that more powerful motivating operations could be arranged for three-dimensional objects than were arranged for the missing items in the vocal and selection-based mand assessment. The investigator gave the participant a task to complete with three-dimensional objects, but one object needed to complete the task was missing. The investigator instructed the participant to complete the task and to ask for any additional items needed to complete the task. For example, after giving the participant a pen, but no paper, the investigator told the participant to write his or her name. If the participant asked for the missing item, the investigator presented it. If the participant did not request the missing item, the investigator removed the other items and presented the next task. The 3D mand assessment included five tasks. The participant was given 5 s to initiate a response and 10 s to complete the response. A correct response consisted of saying the name of the missing item within the time limit.

Procedure

During session 1, the investigator obtained informed consent and then administered a demographic questionnaire and the DRS-2. Next, the investigator administered the VBAB, which included the listener behavior assessment, echoic assessment, tact assessment, intraverbal assessment, vocal mand

and selection-based mand assessment, and 3D mand assessment. The order of administration of the verbal operant assessments was randomized across participants, but was held constant across sessions for each individual participant.

During session 2, which was approximately one week after session one, the investigator administered the short-form BNT and the VBAB. The purpose of re-administering the VBAB was to examine the test-retest reliability of the assessment.

Interobserver Agreement and Procedural Integrity

Interobserver agreement (IOA) data were collected on the scoring accuracy of the DRS-2, the short-form BNT, and the VBAB. Before evaluating IOA, the secondary observer was trained by viewing videotaped sessions until the secondary observer and the trainer reached over 95% agreement for the DRS-2 and the BNT and 100% agreement for the VBAB. The secondary observer viewed videotaped sessions for 32% of the sessions and scored each participant response as correct or incorrect. An agreement was defined as both the primary and secondary observer scoring the same response as correct or incorrect. Interobserver agreement percentage for each session was calculated by dividing the number of trials in which an agreement was scored by the sum of agreement and disagreement trials and multiplying by 100. Agreement for the DRS-2 and BNT were 98.9% and 99.0%, respectively, and agreement for the VBAB was 99.8%.

The secondary observer was trained to evaluate procedural integrity in the same manner, and to the same criterion, as described for IOA. The secondary observer viewed video recordings of the same sessions in which IOA was assessed and evaluated the degree to which the experimenter followed the assessment protocols. Using a checklist of appropriate procedural steps for each assessment, the secondary observer scored each assessment trial as correct if all instructions and stimuli were correctly presented or incorrect if any deviations from acceptable protocol were noted. The percentage of assessment trials that were correctly

Table 1
Assessment Score Means, Standard Deviations, and Ranges During Session 1

Assessment	Control (*n* = 15)		Dementia (*n* = 16)		Overall (*N* = 31)	
	M (SD)	Range	M (SD)	Range	M (SD)	Range
Listener Behavior	30 (0)	30	30 (0)	30	30 (0)	30
Echoic	29.9 (0.35)	29–30	29.9 (0.34)	29–30	29.9 (0.34)	29–30
Tact	28.9 (1.44)	26–30	26.6 (3.86)	16–30	27.7 (3.13)	16–30
Intraverbal	25.9 (4.37)	17–30	24.4 (4.1)	15–30	25.2 (4.21)	25–30
Vocal Mand	25.1 (5.9)	13–30	21.4 (5.86)	7–30	23.2 (6.09)	7–30
Selection-Based Mand	29.9 (0.26)	29–30	29.4 (1.1)	27–30	29.7 (0.83)	27–30
3D Mand	4.9 (0.35)	4–5	4.7 (0.48)	4–5	4.8 (0.43)	4–5

Note. 3D mand assessment has 5 items; all other assessments have 30 items.

implemented was 99.1% and 96.5% for the DRS-2 and BNT, respectively, and 99.9% for the verbal operant assessments.

Data Analysis

Between-group differences were analyzed by comparing mean scores and using Mann-Whitney tests. A nonstatistical error analysis was completed to determine whether participants made errors on the same items across all assessments (e.g., responded incorrectly on ''protractor'' across assessments) or if errors were inconsistent across assessments (e.g., responded incorrectly on ''protractor'' on one assessment but gave the correct response on the other assessments). The error analysis helped to determine if the data were consistent with functional independence of verbal operants. In addition, functional independence was evaluated by analyzing within-subjects differences using Friedman's ANOVAs. Kendall's tau correlations were used to determine test-retest reliability of the VBAB and the degree to which the short-form BNT was correlated with the DRS-2 and the VBAB. Finally, to assess changes in performance from session 1 to session 2, investigators compared mean scores and used a Wilcoxon Signed Rank Test.

For the statistical analyses, nonparametric tests (e.g., Mann-Whitney, Friedman's ANOVA, Kendall's tau, Wilcoxon Signed Ranked Test) were used because the VBAB data violated the assumptions of parametric analyses. More specifically, the data in this study were not normally distributed. Nonparametric tests evaluate ranked data rather

than raw data. Ranks are determined by giving the lowest raw data point the lowest rank (i.e., 1) and working up to higher ranks with higher raw scores. Tied ranks are acceptable (Field, 2009).

RESULTS

Results are reported for the first administration of the VBAB except when otherwise indicated. Data from the second administration were used only for evaluating test-retest reliability and determining differences in performance between the first and the second session. Listener behavior, echoic, selection-based mand, and 3D mand assessments were excluded from the between-groups comparisons, error analyses, and within-subjects evaluations due to ceiling effects.

Group differences were evaluated to help determine how verbal behavior differs between individuals with and without dementia. The control group outperformed the dementia group by at least one point on the tact, intraverbal, and vocal mand assessment, but the groups performed similarly on the echoic, listener behavior, selection-based mand, and 3D mand assessments (see Table 1 and Figure 1). Mann-Whitney tests were used to evaluate group differences on the tact, intraverbal, and vocal mand assessments. A Bonferroni correction was applied so all effects are reported at a .0167 level of significance. According to mean ranks, the control group performed better on the tact, intraverbal, and vocal mand assessments than the dementia group (see Figure 2). The control group had significantly higher scores
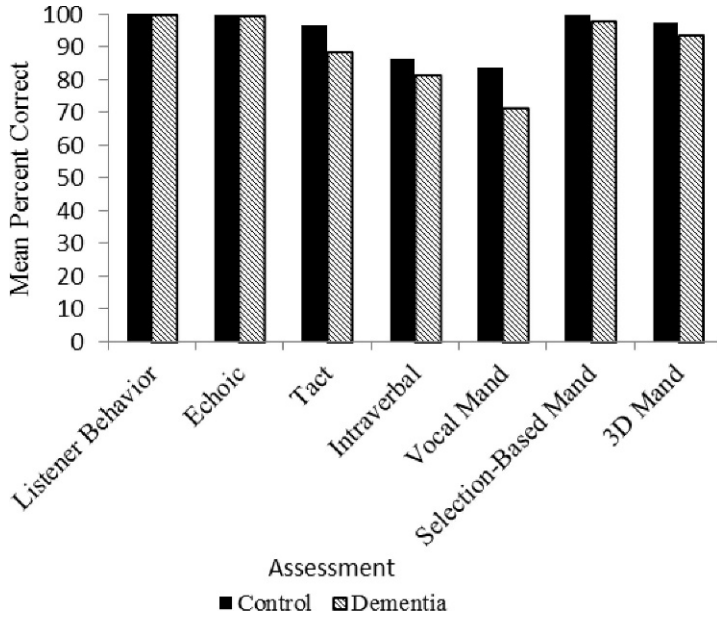
Figure 1. Verbal behavior assessment mean percent correct.

on the tact assessment than the dementia group, $U = 60.5$, $p = .007$, $r = -.43$, but the differences were not significant for the intraverbal assessment, $U = 89.0$, $p = .112$, $r = -.22$, or the vocal mand assessment, $U = 67.5$, $p = .019$, $r = -.37$. Differences on the vocal mand assessment approached significance, but using the Bonferroni correction, the groups did not differ significantly.

A nonstatistical error analysis and within-subjects analyses were used to evaluate the functional independence of verbal operant categories. The investigators performed a nonstatistical error analysis across the tact, intraverbal, and vocal mand assessments, as these assessments were not restricted by ceiling effects. If verbal behavior categories were functionally independent, one would expect to see within-subject differences across assessments on topographically identical responses. Evidence against functional independence would show that individuals respond exactly the same way across all assessments for a specific stimulus item. That is, if they make an error on a word during one assessment, they would consistently make errors on that same word across all assessments. Twelve participants consistently made errors on at least one of the words across the three assessments. In contrast, 30 participants made inconsistent errors across the assessments for at least one of the words. Overall, there were 284 instances in which a participant made inconsistent errors across the three assessments on items requiring topographically identical responses. This is compared to only 31 occasions on which a participant consistently made errors across all three assessment conditions. Based on these findings, it appears the function of the response influences responding more than does its topography. In other words, the data
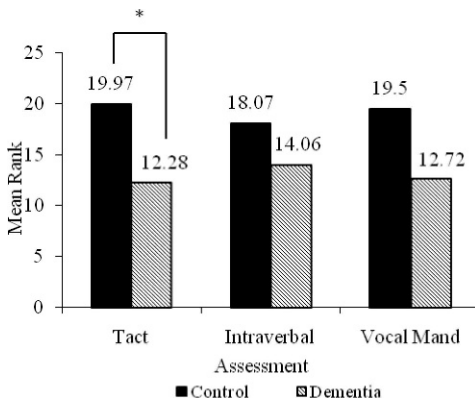

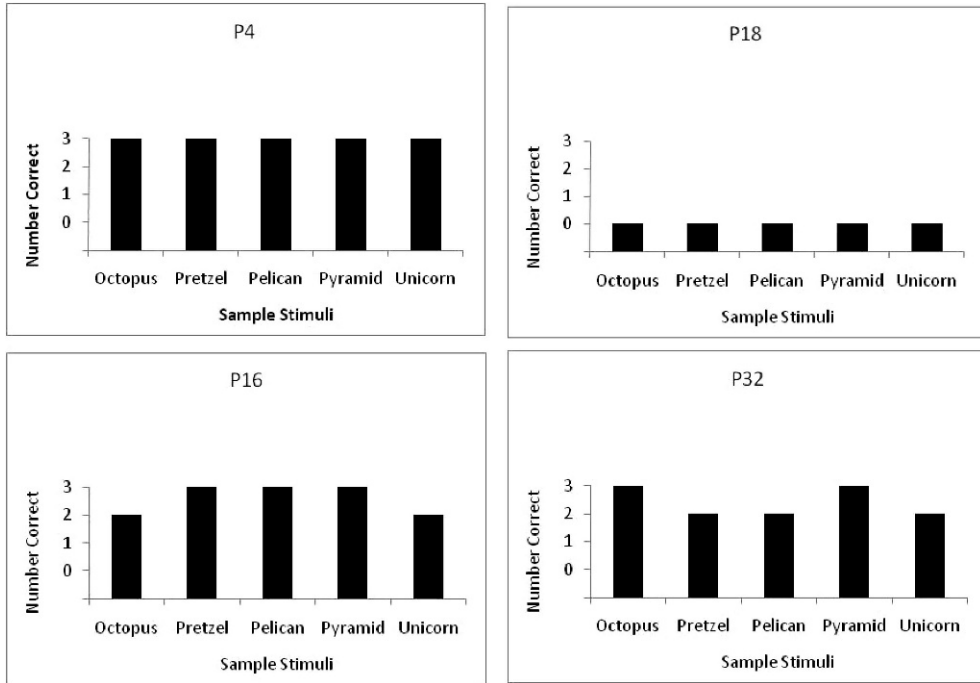*Figure 2.* Between-groups mean rank differences.

Figure 3. Performance across the tact, intraverbal, and vocal mand assessment for a subset of participants and a sample of stimuli. If the participant produced the correct response across assessments, the total score was three. If there were no correct responses, the score was zero. Scores of one or two indicate that the participant provided the correct response on some, but not all, of the assessments.

are consistent with the functional independence of verbal operants.

Figure 3 shows data from a sample of participants and a sample of stimuli, collapsed across the tact, intraverbal, and vocal mand assessments. Participant 4 was a 62-year-old male in the control group. He was the highest performer, scoring 30 out of 30 possible correct responses on each of the tact, intraverbal, and vocal mand assessments. Participant 18, a 79-year-old female, was in the dementia group and was the lowest performer. She scored 16, 15, and 15 out of 30 on the tact, intraverbal, and vocal mand assessments, respectively. Participants 16 and 32 were more typical performers. Participant 16 was a 66-year-old male in the control group who scored 28, 29, and 27 on the tact, intraverbal, and vocal mand assessments, respectively. Participant 32 was an 81-year-old female in the dementia group who scored 29, 24, and 17 on the tact, intraverbal, and vocal mand assessments, respectively. Notably, Participant 32, from the dementia group showed more variability

in performance across measures than Participant 16, who was from the control group. If the participant produced the correct response for a particular stimulus across all three assessments, the total score, shown in Figure 3, is 3. If there were no correct responses, the score is zero. Scores of 1 or 2 indicate that the participant provided the correct response on some, but not all, of the assessments. The figure illustrates the variability in performance across participants and across assessment measures on various stimuli. Many participants had performances similar to participants 16 and 32, which again supports the functional independence of verbal operants.

Friedman's ANOVAs were used to statistically evaluate the differences among the assessments for each group, and Wilcoxon tests were used to follow up on these analyses (see Figure 4). Bonferroni corrections were applied; therefore, all effects are reported at a .0167 level of significance. For the control group, scores on the tact, intraverbal, and vocal mand assessments differed significantly ($\chi^2(2) = 7.9$, $p = .019$). Participants in the
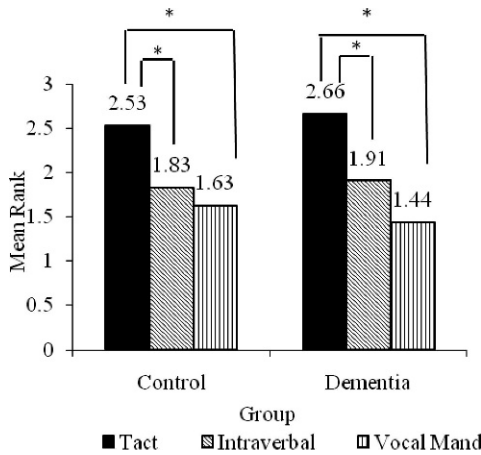
Figure 4. Within-subjects mean rank differences.

control group performed better on the tact assessment than the intraverbal assessment, $T = 8.00$, $p = .015$, $r = -.44$, and better on the tact assessment than the vocal mand assessment, $T = 5.00$, $p = .007$, $r = -.49$. There was no significant difference between the intraverbal and vocal mand assessments, $T = 34.50$, $p = .438$, $r = -.14$. For the dementia group, scores on the tact, intraverbal, and vocal mand assessments also differed significantly ($\chi^2(2) = 12.7$, $p = .002$). Participants in the dementia group performed better on the tact assessment than the intraverbal assessment, $T = 15.50$, $p = .011$, $r = -.64$. They also performed better on the tact assessment than the vocal mand assessment, $T = 12.50$, $p = .004$, $r = -.51$. There was not a significant difference between the intraverbal and vocal mand

assessments, $T = 20.00$, $p = .041$, $r = -.36$. Therefore, the pattern of findings for the control group and the dementia group was the same; performance on tacts was significantly better than on intraverbals or vocal mands, but there was no significant difference between intraverbals and vocal mands. These findings provide additional support for the functional independence of verbal operants, because performance differed across assessments within each group.

Kendall's tau was used to evaluate test-retest reliability of the verbal operant assessments. The listener behavior assessment was removed from this analysis because all participants displayed 100% accuracy across sessions, but all other verbal behavior assessments were evaluated for test-retest reliability. The first echoic assessment was not significantly correlated with the second echoic assessment, $\tau = .199$, $p = .275$. Ceiling effects probably contributed to the lack of significant findings, as all participants scored 29 or 30 out of 30 on this measure during both sessions. However, the tact, intraverbal, vocal mand, selection-based mand, and 3D mand assessments at session 1 and session 2 were all significantly correlated at the .01 level; $\tau = .403$, $p = .009$ for tact; $\tau = .719$, $p<.001$ for intraverbal; $\tau = .677$, $p<.001$ for vocal mand; $\tau = .496$, $p = .005$ for selection-based mand; and $\tau = .600$, $p = .001$ for 3D mand. Across assessment measures, mean scores were similar or improved from session 1 to session 2 for both groups (see Table 2). A Wilcoxon Signed Ranks Test was used to statistically evaluate the difference in performance between session 1 and session 2 using the total score from the

Table 2
Assessment Score Means During Session 1 and Session 2

Assessment	Control ($n = 15$)		Dementia ($n = 16$)		Overall ($n = 31$)	
	Session 1	Session 2	Session 1	Session 2	Session 1	Session 2
Listener Behavior	30	30	30	30	30	30
Echoic	29.9	29.9	29.9	29.9	29.9	29.9
Tact	28.9	29.6	26.6	28.1	27.7	28.8
Intraverbal	25.9	28.9	24.4	26.4	25.2	27.6
Vocal Mand	25.1	28.5	21.4	24.9	23.2	26.6
Selection-Based Mand	29.9	30	29.4	29.6	29.7	29.8
3D Mand	4.9	4.9	4.7	4.7	4.8	4.8

Note. The 3D mand assessment had 5 items; all other assessments had 30 items.

tact, intraverbal, and vocal mand assessments. Participants performed significantly better during session 2 than session 1, $T = 1$, $p <$ .001, $r = -.08$. Improved performance from session 1 to session 2 was likely due to practice effects, as corrective feedback was not provided to participants during either session.

Kendall's tau was also used to evaluate the relationship between the short-form BNT and the other assessment measures. Again, the listener behavior assessment was removed from the analysis due to ceiling effects. A significant correlation was found between the short-form BNT and the DRS-2, $\tau = .287$, $p = .039$. The short-form BNT was not significantly correlated with the echoic assessment or the 3D mand assessment. However, the short-form BNT was significantly correlated with the intraverbal assessment $\tau = .330$, the vocal mand assessment, $\tau = .327$, and the selection-based mand assessment, $\tau = .365$ (all *ps*, <.05). The highest correlation was between the short-form BNT and the tact assessment, $\tau = .551$, $p<.001$, which is understandable, as these assessments required the same skill of naming line drawings.

DISCUSSION

The current article reports the development, administration, and results of a verbal behavior assessment battery (VBAB) for older adults that is based on Skinner's (1957) analysis of verbal behavior. The findings provide preliminary support for the functional independence of verbal operants in older adults. Functional independence was evaluated by analyzing within-subject differences across assessments. The primary verbal operant assessments contained the same 30 items, but the controlling variables differed across assessments. If an individual ''loses'' a word from their repertoire completely, they would be expected to make errors on that item across the primary assessments. However, if the verbal operants are functionally independent, then a given response might occur under some testing conditions but not others. Data from the current investigation showed that participants from both groups performed better on the tact assessment than the vocal mand and intraverbal assessments, and performance on vocal mand and intraverbal assessments did

not differ. These findings support Skinner's conceptualization of functionally independent verbal operants, because participants' performance on the assessments differed despite the assessments containing the same 30 words. Moreover, the error analysis revealed that 30 of 31 participants were able to give topographically identical responses under some assessment conditions but not others. These 30 participants provided 284 instances in which they responded inconsistently on the same word across the tact, intraverbal, and vocal mand assessments. For example, participants were able to provide the correct response on an item during the tact assessment, but then made errors on the same item during the intraverbal or vocal mand assessments. This inconsistent and idiosyncratic loss of functional verbal behavior provides preliminary support for Skinner's claims that verbal operant repertoires are lost independently of one another. The ability to identify older adults' strengths and weaknesses across verbal operant categories may facilitate the development and evaluation of targeted, function-based interventions.

In addition to functional independence, group differences were evaluated. The VBAB was administered to older adults with and without dementia, as determined by the DRS-2. Results revealed that individuals with dementia performed significantly more poorly than those without dementia on the tact assessment. On the intraverbal and vocal mand assessments, the control group had higher average scores than the dementia group, but the magnitude of these group differences did not reach statistical significance. These findings suggest that tacts are particularly sensitive to decline in functioning, but intraverbals and mands may also be sensitive to language difficulties associated with dementia. Baker et al. (2008) recommended using existing, validated tools for assessing verbal behavior if the assessment tools measure the appropriate controlling variables. The BNT is a measure that could be used to evaluate tacts, as it requires individuals to name visual stimuli, and it was highly correlated with the tact assessment in the current study. Future evaluations should include the BNT or a similar measure, with the intention of evaluating the tact repertoire,

as tacts appear to be the verbal operant most sensitive to language difficulties associated with dementia. Given the potential sensitivity of intraverbals and mands, future assessment batteries should also include an evaluation of these verbal operant categories.

These findings are not without limitations. The current sample size was small. In addition, ceiling effects on some measures made it difficult or impossible to interpret the data. With regard to the assessment procedures, the mand assessments in this study did not allow assessing ''pure'' mands (i.e., mands controlled only by the motivating variable). Instead, the responses in these assessments may have been multiply controlled, by the motivating variable as well as by the presentation of the full stimulus scene, including the target stimulus item, in the initial part of the vocal mand assessment procedure. The vocal mand assessment procedure may have created a situation in which the verbal behavior was similar to a metonymical tact extension. A metonymical tact extension occurs when a stimulus gains control over a response because of previous associations with another stimulus upon which reinforcement is typically contingent (Skinner, 1957). For example, saying ''bed'' when seeing a bedroom scene without a bed might be due to the frequent correlation of a bed in a bedroom. Therefore, the response ''bed'' may have been a metonymical tact extension or a mand for the picture of the bed. To attempt to arrange stronger motivating operations than those in the vocal and selection-based mand assessments, the investigators developed the 3D mand assessment procedures. Yet, there were limitations to this assessment, which included the sampling of just five responses and the unknown strength of motivating operations for the selected items (i.e., paper, screwdriver, scissors, ruler, key). The 3D mand assessment might be improved in the future by selecting items that are more commonly used by older adults on a daily basis (e.g., fork) or by assessing responses naturalistically (e.g., serving lunch but not providing a fork). Another procedural consideration that is applicable to the vocal, selection-based, and 3D mand assessments is that the investigators may have reinforced responses by providing the missing items, although no verbal feedback was provided in

these or other assessment conditions. Further refinement of assessment procedures is necessary for assessing pure mands. Finally, the investigators intentionally used the same 30 items across the primary assessments with the goal of assessing the functional independence of verbal operants. With the current results providing preliminary data to support functional independence of verbal operants in older adults, future researchers might consider using different items of similar difficulty across verbal operant categories to avoid practice effects.

The findings from this study provide some support for Skinner's predictions of how verbal behavior deteriorates. Skinner (1957) suggested that ''[t]he verbal operants least likely to be forgotten are echoic and textual'' (p. 207). The current results support this suggestion, as participants consistently responded correctly during the echoic assessment. Skinner also predicted that ''[d]amage is usually most severe in verbal behavior receiving generalized reinforcement'' (p. 219). Because of this, we predicted that mands might remain in the repertoire longer than tacts or intraverbals; however, the results suggested that tacts were stronger than mands and intraverbals. As previously mentioned, the procedures used in the study may not have allowed for an assessment of pure mands, and instead, the procedures may have resulted in responses similar to metonymical tact extensions. If this was the case, then the complexity of this type of tact might explain the increased difficulty in responding during this assessment compared to the general tact assessment. Future research should continue to refine assessment procedures, particularly as they relate to the assessment of pure mands. This may allow better determination of the way in which verbal behavior deteriorates in older adults.

To the authors' knowledge, this is the first study to demonstrate the development, evaluation, and findings of a functional verbal behavior assessment battery for use with older adults. The data provide support for Skinner's analysis of functionally independent verbal operants. Results also suggest that older adults with dementia have more difficulty with verbal behavior than those without dementia. Using verbal behavior

assessments similar to those used in this study may help identify specific verbal behavior strengths and deficits of older adults, and the added precision of these assessments may allow for the development of individualized interventions that target specific verbal operant categories.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th rev. ed.) Washington, DC: Author.

Baker, J. C., LeBlanc, L. A., & Raetz, P. B. (2008). A behavioral conceptualization of aphasia. *The Analysis of Verbal Behavior*, *24*, 147–158.

Benson, D. F., & Ardila, A. (1996). *Aphasia: A clinical perspective*. New York, NY: Oxford University Press.

Dixon, M. R., Baker, J. C., & Sadowski, K A. (2011). Applying Skinner's analysis of verbal behavior to persons with dementia. *Behavior Therapy*, *41*, 120–126.

Esch, B. E., LaLonde, K. B., & Esch, J. W. (2010). Speech and language assessment: A verbal behavior analysis. *The Journal of Speech-Language Pathology and Applied Behavior Analysis*, *5*, 166–191.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications.

Fisher, N. J., Tierney, M. C., Snow, W. G., & Szalai, J. P. (1999). Odd/Even short forms of the Boston Naming Test: Preliminary geriatric norms. *The Clinical Neuropsychologist*, *13*, 359–364. doi: 10.1076/clin.13.3.359.1742

Guess, D., & Baer, D. M. (1973). An analysis of individual differences in generalization between receptive and productive language in retarded children. *Journal of Applied Behavior Analysis*, *6*, 311–329. doi: 10.1901/jaba.1973.6-311

Jurica, P. J., Leitten, C. L., & Mattis, S. (2001). *DRS-2: Dementia Rating Scale-2*. Lutz, FL: Psychological Assessment Resources.

Kempler, D. (2005). *Neurocognitive disorders in aging*. Thousand Oaks: Sage Publications.

Longman advanced American dictionary. (2000). Harlow, Essex, England: Pearson Education Limited.

Michael, J. (1988). Establishing operations and the mand. *The Analysis of Verbal Behavior*, *6*, 3–9.

Mitrushina, M., Boone, K. B., Razani, J., & D'Elia, L. F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd ed.). New York, NY: Oxford University Press.

Sautter, R. A., & LeBlanc, L. A. (2006). Empirical applications of Skinner's analysis of verbal behavior in humans. *The Analysis of Verbal Behavior*, *22*, 35–48. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2774593/pdf/anvb-22-01-35.pdf

Schmidt, K. S., Mattis, P. J., Adams, J., & Nestor, P. (2005). Alternate-form reliability of the Dementia Rating Scale-2. *Archives of Neuropsychology*, *20*, 435–441. doi: 10.1016/j.acn.2004.09.011

Sidman, M., Stoddard, T., Mohr, J. P., & Leicester, J. (1971). Behavioral studies of aphasia: Methods of investigation and analysis. *Neuropsychologia*, *9*, 119–140. doi: 10.1016/0028-3932(71)90038-8

Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copley Publishing Group.

Sundberg, M. L. (2008). Verbal behavior milestones assessment and placement program: The VB-MAPP. Concord, CA: AVB Press.

Sundberg, M. L., San Juan, B., Dawdy, M., & Arguelles, M. (1990). The acquisition of tacts, mands, and intraverbals by individuals with traumatic brain injury. *The Analysis of Verbal Behavior*, *8*, 83–99.

Vincent, G. K. & Velkoff, V. A. (2010, May). The next four decades, the older population in the United States: 2010 to 2050. *Current Population Reports*. Retrieved from http://www.census.gov/prod/www/abs/p25.html

Waldrop, J. & Stern, S. (2003, March). Disability status: 2000. *Census 2000 Briefs and Special Reports*. Retrieved from http://www.census.gov/population/www/cen2000/briefs/

Appendix A

Stimuli for Listener Behavior, Echoic, Tact, Intraverbal, and Vocal and Selection-Based Mand Assessments.

Whistle	Beaver	Saw
Pencil	Scroll	Pelican
Sphinx	Comb	Octopus
Accordion	Bed	Protractor
Latch	Escalator	Pretzel
Hammock	Helicopter	Dart
Volcano	Rhinoceros	Trellis
Hanger	Igloo	Asparagus
Dominos	Camel	Globe
Pyramid	Racquet	Unicorn

Stimuli for 3D Mand Assessment

Paper
Screwdriver
Scissors
Ruler
Key

Appendix B

Note: Bold print indicates the exact words that the investigator said. Other text indicates additional instructions for how to administer the assessments.

Listener Behavior Assessment:
Instructions: **In this part, I am going to show you groups of pictures. Then, I will ask you to point to a picture from the group.**
For each item: **Point to the _____.** (present array) Allow 5 seconds for a response. If the participant does not respond, move to the next item.

Echoic Assessment:
Instructions: **In this part, I am going to say words. I want you to repeat each one after me.**
For each item: **Say _____.** Allow 5 seconds for a response. If the participant does not respond, move to the next item.

Tact Assessment:
Instructions: **In this part, I am going to show you pictures of items. I want you to name each item.**
For each item: **Name this.** (present picture) Allow 5 seconds for a response. If the participant does not respond, turn to the next item.

Intraverbal Assessment:
Instructions: **In this part, I am going to describe a word or item. I want you to tell me what I am describing.**
For each item: Read description. Allow 5 seconds to initiate response and 10 sec to complete it.

	Description	Item
1	**This is a type of aircraft with large metal blades on top that spin very fast to make it fly.**	Helicopter
2	**This is a large water bird that catches fish for food and stores them in a deep bag of skin under its beak.**	Pelican
3	**This is an imaginary animal like a white horse with a long straight horn growing on its head.**	Unicorn
4	**This is a house made from blocks of hard snow or ice.**	Igloo
5	**This is an animal that lives in the ocean with eight tentacles, but no skeleton.**	Octopus
6	**This is an instrument usually in the shape of a half-circle, used for measuring and drawing angles.**	Protractor
7	**This is a large heavy African or Asian animal with thick skin and either one or two horns on its nose.**	Rhinoceros
8	**These are small pieces of wood or plastic with a different number of spots on each half of its top side, used in playing a game. When they are standing up and one is knocked over, all the others fall down too.**	Dominos
9	**This is a small pointed object that is thrown or shot during a game that uses a circular board with a bulls eye.**	Dart

Appendix B, *cont.*

	Description	Item
10	**This is a small metal or plastic object used to keep doors, gates, windows, etc. closed.**	Latch
11	**This is a North American animal that has thick fur, a wide flat tail, and cuts down trees with its teeth.**	Beaver
12	**This is a tool that has a flat metal blade with an edge that has teeth, used for cutting wood.**	Saw
13	**This is a round object with a map of the Earth on it.**	Globe
14	**This is a long piece of paper that can be rolled up, and is often used as an official document, especially in past times.**	Scroll
15	**This is an Egyptian image of a lion with a human head, lying down.**	Sphinx
16	**This is a flat piece of plastic or metal with a row of teeth on one side, used to make your hair look neat.**	Comb
17	**This is a hard salty type of bread baked in the shape of a stick or loose knot.**	Pretzel
18	**This is a mountain with a large hole at the top, through which lava sometimes rises into the air.**	Volcano
19	**This is a large stone building with four triangular walls that slope in to a point at the top, found especially in Egypt and Central America.**	Pyramid
20	**This is a set of stairs that move and carry people from one level within a building to another.**	Escalator
21	**This is a green vegetable shaped like a small stick with a point at one end.**	Asparagus
22	**This is a large desert animal with a long neck and one or two humps on its back.**	Camel
23	**This is a curved piece of wood, plastic, or metal with a hook on it, on which you put your clothes. It is put on a closet rod.**	Hanger
24	**This is a frame made of long narrow pieces of wood that cross each other, used to support climbing plants.**	Trellis
25	**This is a piece of equipment used for hitting the ball in games such as tennis, consisting of a stick with netted wire in a round frame.**	Racquet
26	**This is a small object that produces a high shrill sound when you blow into it and is used by referees.**	Whistle
27	**This is a musical instrument that you pull in and out to produce sounds while pushing buttons on one side to produce different notes.**	Accordion
28	**This is a large piece of material or a net that you can sleep on, that hangs between two trees or poles.**	Hammock
29	**This is a narrow pointed wooden instrument, used for writing or drawing, containing a thin stick of black or colored substance and an eraser.**	Pencil
30	**This is a piece of furniture for sleeping.**	Bed

Vocal and Selection-Based Mand Assessments:

Instructions: **In this part, I am going to show you sets of two pictures. The first picture will be a complete picture. The second picture will be the exact same picture, but one piece will be missing. I want you to tell me what you need to complete the picture. Once you do, I will give you the item to place on the picture to complete it**.

For each item: **Look at the first picture of this set.** (present picture 1, wait 3 seconds, cover it, and present picture 2) **Look at the second. What do you need to complete this picture?** Allow 5 seconds to initiate response and 10 seconds to complete it. (if correct, give missing item; if incorrect, move to selection-based mand).

Selection-Based Mand: **Point to what you need to complete this picture.** (present array) Allow 5 seconds to initiate response and 10 seconds to complete it. (if correct, give missing item; if incorrect move to next item on vocal mand assessment).

3D Mand Assessment:

Instructions: **In this part I am going to give you jobs to do, and I want you to complete each job. If you are missing any items that you need to complete the job, tell me what you need and I will give it to you.**

For each item: Present each job. Allow 5 seconds to initiate a response and 10 seconds to complete the response.

	Job	Item
1	**For this job, I want you to write down your name. Tell me if you need anything else to complete the job.** (present a pen)	Paper
2	**For this job, I want you to put this one in the wood so that it looks like this one. Tell me if you need anything else to complete the job.** (present wood piece with two screws, one just started and one completely put into the wood, point to each screw when reading the description)	Screwdriver
3	**For this job, I want you to cut out this circle. Tell me if you need anything else to complete the job.** (present paper with outline of circle)	Scissors
4	**For this job, I want you to tell me how long this line is, in inches. Tell me if you need anything else to complete the job.** (present paper with printed line)	Ruler
5	**For this job, I want you to open the bag and hand me the item inside. Tell me if you need anything else to complete the job.** (present bag with lock on it)	Key