Easier tasks can have better discriminating power:
The case of verbal fluency

Meredith R. D. Melinder, Deanna M. Barch, Gitry Heydebrand, and John G. Csernansky

Washington University
Abstract

Loren and Jean Chapman have highlighted the importance that a task’s psychometric characteristics may have when comparing different populations (Chapman & Chapman, 1973, 1978). Specifically, they have pointed out that when comparing two tasks measuring different constructs, there is a greater likelihood of finding larger deficits on a task with higher discriminating power, irrespective of what the task measures. While researchers have addressed this issue it has typically been assumed that more difficult tasks have greater discriminating power, often without actually measuring discriminating power. In this paper, we present data from two studies using phonologic and category fluency task to demonstrate that the critical factor for task selection and matching is discriminating power, not task difficulty.
Easier tasks can have better discriminating power:

The case of verbal fluency

In their seminal work in the 1970’s, Loren and Jean Chapman highlighted the importance of examining the psychometric characteristics of tasks used to compare different populations, including factors such as task difficulty and discriminating power (Chapman & Chapman, 1973, 1978). The Chapmans define discriminating power as: “the extent to which the score differentiates the more able from the less able subjects and, hence, differentiates two groups that differ in the ability measured by the test” (Chapman & Chapman, 1978). Specifically, they raised the issue that, when comparing two tasks that putatively measure different constructs, there is a greater likelihood of finding larger deficits on the task that has higher discriminating power, irrespective of what the task measures. Since the Chapmans brought this issue to the forefront, researchers have recognized that studies attempting to identify differential deficits in particular cognitive functions in individuals with schizophrenia need to use tasks matched on discriminating power (Strauss, 2001). Most discussions of this problem have used examples in which the more difficult task had greater discriminating power. As such, researchers have often assumed that larger group differences on “easier” tasks as compared to “harder” tasks can be straightforwardly interpreted as evidence of a differential deficit, without directly examining the relative discriminating power of the tasks (Bokat & Goldberg, 2003; Gourovitch, Goldberg, & Weinberger, 1996; Kremen, Seidman, Faraone, & Tsuang, 2003). The goal of the current paper is to reemphasize that the critical factor for task matching is discriminating power, not task difficulty, and that an “easier” task can sometimes have higher discriminating power if the easier task has better reliability and
true score variance (Chapman & Chapman, 1973, 1978). To illustrate this point, we provide data and analyses from two studies of phonologic and category fluency in individuals with schizophrenia.

**Verbal Fluency**

Verbal fluency measures have been frequently used in the neuropsychological literature as a means of estimating semantic system functioning by comparing performance on phonologic and category versions of the task. Phonologic fluency tasks are thought to depend on the ability to use phonologic cues as a means of accessing lexical representations, and are not thought to require the processing of semantic information. In contrast, category fluency tasks are thought to be dependent on the ability to use semantic representations to generate items, and good performance on such tasks is thought to be dependent on an intact semantic system. The rationale behind comparing phonologic and category fluency as a way of examining semantic processing has been that both tasks theoretically require the same general cognitive processes, with the exception that the latter task requires access to the semantic system (Goldberg et al., 1998). In other words, both tasks theoretically require attention, speed, strategic planning, working memory, et cetera, but only category fluency requires that the participant search through their memory based on the semantic aspects of words. Thus, by comparing an individual’s performance on category and phonologic fluency, one can “see” how the semantic system is functioning, over and above other cognitive functions inherent in the tasks.

Most studies of fluency consistently find that individuals with schizophrenia perform worse than healthy controls on both phonologic and category fluency measures.
Verbal Fluency and Discriminating Power

(Allen, Liddle, & Frith, 1993; Bokat & Goldberg, 2003; Bowie et al., 2004; Elvevag, Fisher, Gurd, & Goldberg, 2002; Elvevag, Weinstock, Akil, Kleinman, & Goldberg, 2001; Feinstein, Goldberg, Nowlin, & Weinberger, 1998; Gourovitch et al., 1996; Joyce, Collinson, & Crichton, 1996; Kremen et al., 2003; McKay et al., 1996; Paulsen et al., 1996; Robert et al., 1998; Rossell, Rabe-Hesketh, Shapleske, & David, 1999; Zalla et al., 2004). However, some researchers have found that individuals with schizophrenia show a larger deficit in category fluency as compared to phonologic fluency. Specifically, these studies have found that individuals with schizophrenia show greater impairment in the production of words to fit a semantic category than in the production of words that start with a particular letter (Feinstein et al., 1998; Goldberg et al., 1998; Gourovitch et al., 1996; Kremen et al., 2003; Rossell et al., 1999). In addition, the results of a meta-analysis conducted by Bokat and Goldberg (Bokat & Goldberg, 2003) were also consistent with a differential deficit with respect to category fluency in individuals with schizophrenia. The results of these studies have been interpreted as providing support for the hypothesis that individuals with schizophrenia have relatively specific deficits in the structure or function of their semantic system. However, not all studies have found a differential deficit with respect to semantic fluency measures compared with phonologic fluency in individuals with schizophrenia (Bowie et al., 2004; Elvevag et al., 2001; Joyce et al., 1996; Rossell, Shapleske, & David, 2000; Zalla et al., 2004).

Discriminating Power

Previous studies of verbal fluency in schizophrenia have not addressed the issue of whether phonologic and category fluency tasks are matched on discriminating power. This may not have been addressed, at least in part, because of the belief that it was not
necessary because category fluency tasks are typically “easier” for healthy controls than phonologic fluency tasks; that is, healthy controls produce more items per category than per letter. Thus, the interpretation of such studies relies on the assumption that category fluency tasks have less discriminating power because they are easier, so that greater deficits on category fluency tasks as compared to phonologic fluency tasks among individuals with schizophrenia cannot be due to a generalized deficit. However, discriminating power is essentially true-score variance (the component of variance that is replicable and not due to error or factors unrelated to the construct of interest) (Chapman & Chapman, 1973, 1978). In turn, true-score variance is the product of task reliability and observed score variance (the sum of true-score and error variance)(Chapman & Chapman, 1973, 1978).

Task difficulty can greatly influence both reliability and observed-score variance in that more difficult tasks often have greater observed score variance than easier tasks. However, more difficult tasks do not always have greater reliability or greater observed score variance. Thus, task difficulty cannot be equated with discriminating power. In fact, the Chapmans provided a number of examples in which an easier task with binary items (e.g., scored 0 or 1) had greater true-score variance than a harder task (e.g., accuracy of ~50% compared to ~20%) (Miller, Chapman, Chapman, & Collins, 1995). The situation is further complicated in tasks in which items are not binary, but instead reflect some sort of count that often has a large range. For example, the Chapmans and others have demonstrated that in reaction time data (which can be thought of as analogous to count data), there is often a linear relationship between mean latencies and standard deviations, such that slower participants (i.e., higher reaction time counts) have
larger standard deviations (i.e., observed score variance) (Chapman, Chapman, Curran, & Miller, 1994; Faust, Balota, Spieler, & Ferraro, 1999; Hale, Myerson, Smith, & Poon, 1988). Individuals typically produce more items for category than phonologic fluency tasks (e.g., higher counts). If phonological and categorical fluency tasks have similar reliabilities, but category fluency has greater observed score variance, then category fluency tasks would have greater true-score variance and discriminating power. If so, then one cannot rule out the possibility that the relatively greater deficits that individuals with schizophrenia show on category fluency tasks as compared to phonologic fluency tasks reflect a generalized deficit rather than a deficit in some specific cognitive process tapped by the category fluency task.

Few of the previous studies comparing phonologic and category fluency in schizophrenia have provided comparable reliability data for the two types of tasks. However, examination of the means and standard deviations for the control groups in previous studies does suggest that observed score variance is generally higher for category as compared to phonologic fluency. For example, Elvevag found greater variance for category fluency as compared to phonologic fluency in healthy controls, as well as a differential deficit among individuals with schizophrenia on category fluency tasks (Elvevag et al., 2001). Similarly, Rossell found greater variance for category fluency as compared to phonologic fluency, and found a differential deficit in category as compared to phonologic fluency among deluded patients (Rossell et al., 1999). Of note, not all studies that show greater variance on category fluency versus phonologic fluency among healthy controls have found differential deficits on category fluency among individuals with schizophrenia (Joyce et al., 1996; Zalla et al., 2004). However, because
these studies did not provide reliability data, the discriminating power of category fluency and phonologic fluency could not be compared across studies.

The goal of the current project was to directly compare the discriminating power for category fluency and phonologic fluency tasks and to evaluate the relationship of discriminating power to group difference effect sizes in individuals with schizophrenia and demographically similar controls. Specifically, we analyzed data from two different studies (i.e., Study 1 and Study 2) that administered versions of both category and phonologic fluency tasks. We show that when analyzed in the “typical” fashion, with phonologic and category fluency unmatched for discriminating power, category fluency elicits significantly larger group differences between individuals with schizophrenia and healthy controls than does phonologic fluency. However, when we created dependent variables for the category fluency and phonologic fluency tasks that were matched on discriminating power, there were equal deficits among individuals with schizophrenia on phonologic fluency and category fluency.

Study 1

Method

Participants

Fifty-five outpatients diagnosed with schizophrenia or schizoaffective disorder were recruited from local outpatient mental health treatment facilities to participate in this study. Diagnoses were based on the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition Text Revision (DSM-IV-TR SCID-I;(First, Spitzer, Gibbon, & Williams, 2001)), and a review of the participant's medical records. The Non-Patient version of the SCID-I was used for interviewing
healthy controls. The diagnostic interview was conducted by a Master’s or PH.D. level interviewer (MRDM, DMB, or another advanced graduate student). Subsets of ten SCID interviews were conducted by MRDM and DMB and by MRDM and another advanced graduate student to establish diagnostic interrater reliability. All medicated participants had been receiving the same medications and dosages for at least two weeks prior to testing. Forty-five healthy participants were recruited through local advertisements.

Potential participants (both healthy controls and individuals with schizophrenia) were excluded for: (1) neurological illness or history of head trauma with loss of consciousness of greater than 15 minutes; (2) mental retardation (operationalized as an estimated IQ of less than 70, based on a WRAT-III Reading standard score of 69 or less); and/or (3) being a non-native English speaker. Participants were included in the group of individuals with schizophrenia if they meet criteria for schizophrenia or schizoaffective disorder, according to the SCID-I interview, and if they were without a history of substance abuse within the last three months or substance dependence in the last six months. Twenty-one percent of the individuals with schizophrenia were taking typical antipsychotic medications, 72% were taking atypical antipsychotic medications, and 15% were taking both typical and atypical antipsychotic medications. Individuals were included in the group of healthy participants if they were without current depression or mania, a lifetime history of Axis I psychotic disorders, (including bipolar disorder, schizophrenia, or schizoaffective disorder) and without a history of substance abuse within the last three months or substance dependence in the last six months. Informed consent was obtained from all participants prior to study participation in accordance with the Washington University and Metropolitan Psychiatric Center institutional review
boards and all participants were paid for their participation.

**Tasks**

**Verbal Fluency:** For phonologic fluency, participants were asked to name as many words beginning with a certain letter of the alphabet as they could within the time limit. They were told to exclude proper nouns and different forms of the same word. Each subject completed this task for three phonologic categories: the letters F, A, and S.

For category fluency, participants were asked to name as many exemplars of a category as they could within the time limit. Each subject completed this task for the three semantic categories: Animals, fruits and body parts. Fifty participants (28 healthy controls and 22 individuals with schizophrenia) were asked to name words within a 120-second time limit, and the first 60 seconds were used for this study. Forty-two participants (17 healthy controls and 25 individuals with schizophrenia) were just asked to name words within a 60-second time limit. The performance within the first 60-seconds for participants who completed the 120-second task was not significantly different from that of the participants who only completed the 60-second task. We collected test-retest reliability data for each trial of both the phonological and category fluency tasks used in study 1 and study 2 (60-second versions) from 21 additional controls that were demographically similar in terms of age and parental education to the controls used in this study and study 2. The time interval between test 1 and test 2 was either one or two days.

**WRAT–III Reading.** The Reading subtest of the Wide Range Achievement Test-Third Edition (WRAT-III;(Wilkinson, 1993)), which requires that participants read a list of words increasing in difficulty, was administered as an estimate of premorbid IQ. The
basis for administering this test as a measure of pre-morbid IQ is the finding that reading abilities are often preserved in psychotic disorders even when other cognitive functions have declined (Dalby & Williams, 1986). The raw scores on the Reading subtest were converted to age-scaled standard scores based on the test guidelines (Wilkinson, 1993).

**Results**

The demographic and clinical ratings of both participant groups are shown in Table 1. Individuals with schizophrenia and healthy controls did not differ significantly on age ($t(90) = -0.023, p >0.98$), level of parental education ($t(90) = 1.106, p >0.27$), gender ($X^2(2) = 0.03, p >0.86$), or race ($X^2(2) = 1.21, p >0.55$).

Verbal Fluency Task Reliability

The means and standard deviation for each item in the verbal fluency tasks are shown in Table 2. We began by analyzing the reliability of the two different verbal fluency tasks. We did so only in the control participants, as the Chapmans and others have argued that control participants are the proper group on which to establish task discriminating power (Chapman & Chapman, 1978). Differences across tasks in reliability and observed score variance in the individuals with schizophrenia could be influenced by a specific deficit in one or more tasks domains, and thus would complicate the psychometric analysis of the tasks. We examined the internal consistency of each task (similarity in performance across the different subtrials within each type of fluency task) for the control participants using Cronbach’s alpha. We measured test-retest
reliability for each individual categorical and phonological fluency item, as well as the sum scores, using correlations computed on the additional test-retest sample of controls. The internal consistency of the two verbal fluency tasks was relatively similar: .78 for phonologic fluency and .70 for category fluency. In addition, the test-retest reliability for the sum of the three trials for each verbal fluency task was high and very similar. The test-retest reliability of the individual trials varied somewhat within each task, but all trials had moderate to good test-retest reliability. We then calculated the true-score variance (test-retest reliability X observed score variance) for a number of different ways of computing a dependent variable, including: 1) the average across all three trials within each task; 2) the summed total across all three trials; and 3) each individual trial of each task. As can be seen in Table 2, while both verbal fluency tasks were similar on reliability, the observed score variance for phonologic and category fluency was different. When we calculated discriminating power (reliability X observed score variance) for either the total scores (sum of the 3 trials for each type) or the average (average of the 3 trials for each type), category fluency had nearly twice the true-score variance of phonologic fluency. Thus, category fluency has greater discriminating power than phonologic fluency. As such, we can predict a priori that individuals with schizophrenia will show a larger deficit on the category fluency task than on the phonologic fluency task.

----------------------------------
Insert Table 2 about here
----------------------------------

To test the hypothesis that this difference in discriminating power might account
for observed differences between performance on the different fluency tasks in the subjects with schizophrenia, we first conducted a group comparison without taking into account differences in discriminating power. We analyzed the data using a 3-factor ANOVA, with diagnostic group (patient, control) as a between-subject factor and both fluency task type (category, phonologic) and trial (F, A, S for phonological; animals, fruits, body parts for category) as within-subject factors. As expected, there was a significant main effect for group, $F(1, 90) = 33.3, p < .001$, and a significant main effect for fluency task type, $F(1, 90) = 98.51, p < .001$. In addition, as predicted, there was a significant diagnostic group X fluency task type interaction, $F(1, 90) = 14.71, p < .001$. As shown in Table 2, there were larger group differences for category as compared to phonologic fluency. There were no significant group differences in the number of intrusions or perseverations during the phonologic or category fluency tasks. We also analyzed the data using a 2-factor ANOVA with total scores (sum of the three trials for each fluency task type) as the dependent variable, and diagnostic group as a between-subject factor and fluency task type as a within-subject factor. This analysis again revealed a main effect of group $F(1, 90) = 32.79, p < .001$, a main effect of fluency task type $F(1, 90) = 89.05, p < .001$, and a group X fluency task type interaction $F(1, 90) = 17.13, p < .001$ (Effect Size = $\text{Eta}^2 = .16$). As with the previous analyses, this interaction reflected the fact that there were larger group differences for category as compared to phonologic fluency.

If the only reason for this “differential deficit” on category versus phonologic fluency is the psychometric differences between the tasks, we should be able to eliminate this greater deficit on category fluency by matching the fluency tasks on discriminating
power. On the other hand, if this difference truly reflects a differential deficit on semantic processing as indexed by category fluency deficits, we should find that a group by fluency task type interaction remains even when the tasks are matched for discriminating power. To develop dependent measures for category and phonologic fluency tasks that were matched for discriminating power, we examined the true-score variance for individual trials of the category and phonologic fluency tasks, as well as combinations of trials. The only measure of reliability that we had for the individual trials was test-retest reliability. Thus, we computed true-score variance as the product of test-retest reliability and observed score variance for that trial. For example, as shown in Table 2, the letter “s” trial from phonologic fluency and the “fruit” trial from category fluency were the most similar in discriminating power. In addition, the sum of the letter “a” and “s” trials for phonologic fluency and the “animals” trial for category fluency were also similar in discriminating power.

To then examine whether the group by task type interaction remained significant when using these matched dependent variables, we conducted a 2-factor ANOVA with diagnostic group as a between-subject factor and fluency task type as a within subject factor, using only the letter “s” trial for phonologic fluency and the “fruit” trial for category fluency. We once again found a significant main effect of group, $F(1, 90) = 6.97, p < .01$. However, there was no longer a significant group by fluency type interaction, $F(1, 90) = 0.6, p > .4$ (Effect Size = $\eta^2 = .04$). We found similar results in a 2-factor ANOVA with group as a between-subject factor and task type as a within-subject factor, using the sum of the “a” and “s” trials for phonologic fluency, and the “animals” trial for category fluency. Specifically, there was a significant main effect of
Verbal Fluency and Discriminating Power

15

group, $F(1, 90) = 21.54, p < .001$, but no significant group by fluency type interaction, $F(1, 90) = 0.5, p > .5$ (Effect Size = $\eta^2 = .005$).

Discussion

The results of this study suggest that prior findings of differential deficits on category as compared to phonologic fluency tasks among individuals with schizophrenia may reflect, at least in part, psychometric confounds. Specifically, when we conducted analyses using the same types of dependent variables as used in prior studies, (Feinstein et al., 1998; Goldberg et al., 1998; Gourovitch et al., 1996; Kremen et al., 2003; Rossell et al., 1999) we replicated previous findings of a differential deficit on category as compared to phonologic fluency. However, analyses of discriminating power also indicated that when analyzed in the “traditional” way, category fluency had greater discriminating power than phonologic fluency. As such, greater impairment on category as compared to phonologic fluency could have reflected a generalized deficit. When we used dependent variables for category and phonologic fluency that were matched on discriminating power, we no longer found a significant differential deficit on category fluency. Nonetheless, it was still the case that the category fluency variables tended to have larger group difference effect sizes that the phonological fluency variables that were similar on discriminating power (see Table 2). However, the effect sizes of the group by fluency task type interactions were much smaller for variables similar on discriminating power ($\eta^2=.04/.005$) than the effect sizes for the original analyses ($\eta^2=.16$). Taken together, these results indicate that analyses of fluency variables matched on discriminating power provide much less support for the hypothesis that individuals with schizophrenia have a differential deficit in semantic processing that do analyses in which
the category fluency variables have much higher discriminating power than the phonological fluency variables. Further, the results highlight the fact that “easier” tasks can sometimes have better discriminating power than “harder” tasks.

**Study 2**

The goal of study 2 was to determine if we could replicate these results using a separate sample of individuals diagnosed with schizophrenia and healthy controls.

**Methods**

**Participants**

Participants included 53 individuals with DSM-IV Schizophrenia and 60 healthy controls between the ages of 18 and 59. The individuals with schizophrenia were all clinically stable inpatients or outpatients recruited through an ongoing structural imaging protocol conducted by JGC. Normal controls were recruited from the same community as the individuals with schizophrenia through local advertisements. Healthy controls were excluded if they had any lifetime history of Axis I psychiatric disorder, or any first order family member with a psychotic disorder. Potential participants (either individuals with schizophrenia or healthy controls) were also excluded for: 1) meeting DSM-IV criteria for substance abuse (severe) or dependence (any type) at any time within the past three months; 2) the presence of any clinically unstable or severe medical disorder, or a medical disorder that would confound the assessment of psychiatric diagnosis, or make participation in the research protocol unsafe; 3) present or past head injury with documented neurological sequelae or causing loss of consciousness; and/or 4) meeting DSM-IV criteria for mental retardation (mild or greater in severity). Parental socio-economic status was measured using the Hollingshead Index (Hollingshead & Redlich,
All individuals with schizophrenia were medicated, taking a variety of typical and atypical antipsychotic medications.

Schizophrenia and control diagnoses were determined using the Structured Clinical Interview for DSM-IV (SCID-IV (First et al., 2001)). The structured interviews were conducted by a M.S.W.-level research assistant who had completed SCID-IV training, and who regularly participated in ongoing diagnostic training sessions at the Metropolitan Psychiatric Center. The SCID-IV interviewer had access to all data from present and past Metropolitan Psychiatric Center hospital records, corroborative personal sources (e.g., family), and records from other hospitals. In addition, a semi-structured interview was performed by an expert clinician (in most cases, JGC), also using DSM-IV criteria. This expert clinician also had access to all available medical records and collaborative sources, but was blind to the results of the SCID-IV interview. The participant’s final diagnosis was determined by a consensus meeting between the SCID-IV interviewer and the expert clinician. Informed consent was obtained from all participants prior to study participation in accordance with the Washington University and Metropolitan Psychiatric Center institutional review boards and all participants were paid for their participation.

**Verbal Fluency Tasks**

As part of a larger neuropsychological battery, the participants were administered versions of both category and phonologic fluency tasks. In the phonologic fluency task, participants were asked to name as many words beginning with a certain letter of the alphabet as they could within the time limit. They were told to name as many words beginning with that letter of the alphabet excluding proper nouns and different forms of
the same word. Each subject completed this task for two phonologic categories: the letters: S and P. Likewise, for category fluency the subjects were asked to name as many exemplars of a category as they could within the time limit. Each subject completed this task for the one semantic category, which was "animals." All participants were asked to name the words for each trial within a 60-second time limit.

Results

The demographic and clinical characteristics of both participant groups are shown in Table 3. The groups did not differ significantly on age, \( t(111) = 0.12, p > .9 \), gender, \( X^2(2) = .25, p > .5 \), or parental SES \( t(111) = -1.0, p > .3 \).

We began the analysis of data from Study 2 by examining the relative discriminating power of the two fluency tasks, using the internal consistency estimates from Study 1. With this type of data set (two trials for fluency, one trial for category), one might choose to analyze the data by comparing the average of the two phonologic fluency trials to the one category fluency trial. As can be seen in Table 4, the observed score variance was higher for the category fluency trial as compared to the average of the two phonologic fluency trials, leading to greater true-score variance and discriminating power. As such, one would predict that there would be a larger group difference on the category as compared to phonologic fluency tasks. However, as can be seen in Table 4, the sum of the two phonologic fluency trials (as compared to the average) had a higher observed score variance. Thus true-score variance and discriminating power were more
similar to the true-score variance and discriminating power for category fluency, and so
one no longer would predict a larger group difference on category as compared to
phonologic fluency. Again, however, if there were truly a differential deficit in semantic
processing among individuals with schizophrenia, one would predict that the larger group
differences on category fluency would remain even when using dependent measures
matched for discriminating power.

To test these predictions, we first analyzed the data using the average phonologic
fluency dependent measure, using a 2-factor ANOVA with diagnostic group as a
between-subject factor and fluency task type (phonologic, category) as the within-subject
variables. This ANOVA revealed a main effect of diagnostic group, $F(1,111) = 48.0, p
< .001$, and a main effect of fluency task type, $F(1,111) = 84.5, p < .001$. In addition, as
predicted, we found a significant diagnostic group X fluency task type interaction,
$F(1,111) = 13.0, p < .001$ (Effect Size = Eta$^2$ = .11). We next analyzed the data using the
sum of the two phonologic fluency trials as the dependent measure (matched to category
fluency on discriminatory power), using a 2-factor ANOVA with diagnostic group as a
between-subject factor and fluency task type (phonologic, category) as the within-subject
variables. We again found a significant main effect of group, $F(1,111) = 40.3, p < .001$,
and a significant main effect of task type, $F(1,111) = 84.5, p < .001$. However, the group
X fluency task type interaction was no longer significant $F(1,111) = 0.3, p > .6$ (Effect
Size = Eta$^2$ = .002).
Discussion

The goal of the current study was to illustrate that an “easier” task can have higher discriminating power than a “harder” task, if the “easier” task has either higher reliability or greater observed score variance. We used a comparison of phonologic and category fluency tasks to illustrate this point. As described in the introduction, a number of previous studies have found group by fluency task type interactions that have been interpreted as reflecting a differential deficit in category fluency among individuals with schizophrenia. However, our analyses of the relative discriminating power of phonologic and category fluency tasks show that category fluency has higher discriminating power, even though healthy controls typically perform better on category as compared to phonological fluency. This higher discriminating power primarily results from greater observed score variance, as the internal consistency and test-retest reliability of the two tasks are fairly similar.

Similar to prior studies, we found significant group by fluency task type interactions for both Study 1 and Study 2 when the data were analyzed without taking into account differences in discriminating power. However, when we used dependent variables for each fluency task type that were matched for discriminating power, the effect sizes for these interactions were much reduced and the interactions were no longer significant. If individuals with schizophrenia had a differential deficit in semantic processing, as indexed by category fluency performance, one would have expected the magnitude of the group by fluency task type interaction to remain stable even when the dependent variables were matched on discriminating power. Of note, it was the case that the category fluency variables did tend to have higher group difference effect sizes than
the phonological fluency variables, even though the group by fluency task type interactions were no longer significant when examining dependent variables similar on discriminating power. Even though these studies had relatively large sample sizes and good power to detect significant interactions, it is possible that even larger sample sizes would reveal greater evidence for a differential deficit in semantic processing. In either case, the analyses presented above demonstrate that at minimum, the magnitude of any such interaction with fluency task type is dramatically reduced when comparing variables unmatched on discriminating power to variables matched on discriminating power.

In summary, the current results highlight the importance of examining relative discriminating power in all studies of cognitive and language function in special populations such as individuals with schizophrenia, even when one finds what appears to be a differential deficit on an “easier” task. Although the Chapmans (Chapman & Chapman, 1973, 1978) clearly articulated this point in their earlier writings, many researchers have relied too much on the role of task difficulty in establishing differential deficits. However, as pointed out by the Chapmans and by the results of the current study, an “easier” task can clearly have better discriminating power than a “harder” task. Therefore, without an explicit analysis of discriminating power in addition to task difficulty, one can reach misleading conclusions about the magnitude or even presence of differential cognitive deficits.
References


Dalby, J. T., & Williams, R. (1986). Preserved reading and spelling ability in psychotic disorders. *Psychological Medicine, 16*, 171-175.


Author Notes

Meredith R. D. Melinder, Department of Psychology, Washington University;
Deanna M. Barch, Departments of Psychology and Psychiatry, Washington University;
Gitry Heydebrand, Department of Psychiatry, Washington University; John G.
Csernansky, Department of Psychiatry, Washington University.

This work was supported by a NARSAD Young Investigator’s award to DMB, and NIMH grants to the DMB (MH60887) and JGC (MH56584), as well as a Conte Center for Neuroscience Research MH62130 grant. We wish to thank the staff and clients of the St. Louis Metropolitan Psychiatric Center for their kind cooperation and support.

Correspondence concerning this article should be addressed to Deanna M. Barch, Department of Psychology, Washington University, Campus Box 1125, One Brookings Drive, St. Louis, MO. 63130 (email: dbarch@artsci.wustl.edu).
Table 1:
Demographic and Clinical Characteristics for Participants in Study 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy Controls (N=45)</th>
<th>Participants with Schizophrenia (N=47)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>Age</td>
<td>37.89</td>
<td>9.15</td>
</tr>
<tr>
<td>Sex (% male)</td>
<td>57.8</td>
<td>--</td>
</tr>
<tr>
<td>Race (% Caucasian)</td>
<td>62.2</td>
<td>--</td>
</tr>
<tr>
<td>Education *</td>
<td>16.47</td>
<td>2.72</td>
</tr>
<tr>
<td>Parental Education</td>
<td>14.31</td>
<td>2.29</td>
</tr>
<tr>
<td>Age at first hospitalization</td>
<td>--</td>
<td>2.29</td>
</tr>
<tr>
<td>Number of hospitalizations</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>WRAT-III Reading standard score *</td>
<td>106.40</td>
<td>12.18</td>
</tr>
</tbody>
</table>

*p < .001 (2-tailed)
Table 2:
Study 1: Performance of Phonologic and Category Fluency

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy Controls (N=45)</th>
<th>Schizophrenia (N=47)</th>
<th>Internal Consistency</th>
<th>Test Retest Reliability (Controls)</th>
<th>Variance (Controls)</th>
<th>True Score Variance (Controls)</th>
<th>Group Difference</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phonological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>42.2</td>
<td>9.8</td>
<td>34.8</td>
<td>12.5</td>
<td>.78</td>
<td>.81</td>
<td>96.29</td>
<td>78.0</td>
</tr>
<tr>
<td>Average</td>
<td>14.1</td>
<td>3.3</td>
<td>11.5</td>
<td>4.2</td>
<td>.78</td>
<td>.81</td>
<td>10.7</td>
<td>8.7</td>
</tr>
<tr>
<td>F</td>
<td>14.6</td>
<td>3.7</td>
<td>12.0</td>
<td>4.9</td>
<td>.64</td>
<td>13.3</td>
<td>8.5</td>
<td>.56</td>
</tr>
<tr>
<td>A</td>
<td>12.8</td>
<td>4.0</td>
<td>9.6</td>
<td>4.4</td>
<td>.51</td>
<td>15.9</td>
<td>8.1</td>
<td>.72</td>
</tr>
<tr>
<td>S</td>
<td>14.8</td>
<td>4.1</td>
<td>13.0</td>
<td>5.0</td>
<td>.85</td>
<td>16.8</td>
<td>14.3</td>
<td>.41</td>
</tr>
<tr>
<td>Letters A + S</td>
<td>27.6</td>
<td>7.1</td>
<td>22.6</td>
<td>8.5</td>
<td>.76</td>
<td>50.1</td>
<td>38.1</td>
<td>.61</td>
</tr>
<tr>
<td><strong>Category</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>60.4</td>
<td>13.5</td>
<td>41.9</td>
<td>14.1</td>
<td>.70</td>
<td>.85</td>
<td>183.0</td>
<td>155.6</td>
</tr>
<tr>
<td>Average</td>
<td>20.1</td>
<td>4.5</td>
<td>14.2</td>
<td>4.3</td>
<td>.70</td>
<td>.85</td>
<td>20.3</td>
<td>17.3</td>
</tr>
<tr>
<td>Animals</td>
<td>20.6</td>
<td>5.7</td>
<td>14.6</td>
<td>5.0</td>
<td>.74</td>
<td>32.8</td>
<td>24.3</td>
<td>.98</td>
</tr>
<tr>
<td>Fruits</td>
<td>15.4</td>
<td>3.9</td>
<td>11.3</td>
<td>3.7</td>
<td>.78</td>
<td>14.9</td>
<td>11.6</td>
<td>.96</td>
</tr>
<tr>
<td>Body Parts</td>
<td>24.4</td>
<td>7.1</td>
<td>16.9</td>
<td>5.8</td>
<td>.84</td>
<td>50.0</td>
<td>42.0</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Verbal Fluency and Discriminating Power
Table 3
Demographic and Clinical Characteristics for Participants in Study 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy Controls (N=60)</th>
<th>Participants with Schizophrenia (N=53)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>37.2</td>
<td>36.9</td>
</tr>
<tr>
<td>Gender (% Male)</td>
<td>50%</td>
<td>55%</td>
</tr>
<tr>
<td>Education</td>
<td>4.0</td>
<td>4.1</td>
</tr>
<tr>
<td>Parental SES</td>
<td>11.9</td>
<td>11.5</td>
</tr>
<tr>
<td>Age of Onset</td>
<td>--</td>
<td>23.1</td>
</tr>
<tr>
<td>Number of Hospitalizations</td>
<td>--</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.0</td>
</tr>
</tbody>
</table>
Table 4:

Study 2: Performance of Phonologic and Category Fluency

<table>
<thead>
<tr>
<th>Variable</th>
<th>Healthy Controls (N=60)</th>
<th>Schizophrenia (N=53)</th>
<th>Test-Retest Reliability (Controls)</th>
<th>Variance (Controls)</th>
<th>True Score Variance (Controls)</th>
<th>Group Difference Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonological</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (S + P)</td>
<td>27.8</td>
<td>7.3</td>
<td>21.6</td>
<td>8.8</td>
<td>.85</td>
<td>52.94</td>
</tr>
<tr>
<td>Average (of S + P)</td>
<td>13.9</td>
<td>3.6</td>
<td>10.8</td>
<td>4.4</td>
<td>.85</td>
<td>13.24</td>
</tr>
<tr>
<td>Category</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animals</td>
<td>20.8</td>
<td>6.6</td>
<td>13.8</td>
<td>3.9</td>
<td>.74</td>
<td>43.74</td>
</tr>
</tbody>
</table>