Pergamon

Archives of Clinical Neuropsychology
19 (2004) 191–201

Archives of CLINICAL NEUROPSYCHOLOGY

The Penn Conditional Exclusion Test: a new measure of executive-function with alternate forms for repeat administration

Matthew M. Kurtz [a,b,*], J. Daniel Ragland [c], Paul J. Moberg [c], Ruben C. Gur [c]

[a] *Institute of Living, 200 Retreat Avenue, Hartford, CT 06106, USA*
[b] *Department of Psychiatry, Yale School of Medicine, New Haven, CT 06511, USA*
[c] *Department of Psychiatry, University of Pennsylvania, School of Medicine, Philadelphia, PA 19104, USA*

Accepted 26 November 2002

Abstract

Studies of change in neuropsychological function over time in both healthy and diseased populations have been hindered by the absence of alternate forms of most commonly used neuropsychological tests. The Penn Conditional Exclusion Test (PCET) was developed as a new cognitive instrument to assess "executive" functioning with four alternate forms. In Study 1, the PCET was administered in counterbalanced order to 80 healthy young adults to establish equivalent test difficulty. Results revealed that the four versions were related on categories achieved, total number of errors and total number of trials. In Study 2, the PCET was administered to 25 healthy adults of a wide age range with a battery of computerized neuropsychological tests to assess construct validity. Convergent validity was confirmed by positive correlations between the PCET and a measure of abstraction. Divergent validity was established through low, nonsignificant correlations between the PCET and measures of facial emotion recognition, word and face memory, visuospatial function, and verbal reasoning.

Keywords: Executive-function; Cognitive test; Alternate forms; Practice effects

* Corresponding author.
E-mail address: MKURTZ@HARTHOSP.ORG (M.M. Kurtz).

0887-6177/$ – see front matter © 2003 National Academy of Neuropsychology.
doi:10.1016/S0887-6177(03)00003-9

1. Introduction

Assessment of change in neurocognition over time in the same individual has played an increasingly large role in the professional activities of neuropsychologists in both clinical and research settings. With a growing elderly population, there is renewed interest in tracking the course of neurocognitive abilities across the lifespan in both healthy and diseased populations. At the same time, refined pharmacologic and behavioral interventions for a variety of neurologic and neuropsychiatric disorders have placed neuropsychologists in a central role for the assessment of treatment outcome in both randomized, treatment-control group comparison studies and in individual patients. A lack of alternate forms for most commonly used neuropsychological measures severely curtails the usefulness of neuropsychological test data in such situations.

This problem is illustrated in contemporary treatment research in schizophrenia. A wealth of data has shown that the majority of patients with schizophrenia exhibit profound neurocognitive deficits on measures of executive-function (abstraction and problem-solving), as well as attention, learning and memory, language and sensory motor skill (Heinrichs & Zakzanis, 1998; Johnson-Selfridge & Zalewski, 2001). These cognitive deficits are more closely related to measures of psychosocial status than symptoms, and are resistant to the effects of typical and atypical antipsychotic medication (Green, 1996; Purdon et al., 2000). These deficits, however, have not been the target of traditional treatment interventions.

In recent years efforts have been directed at the development of novel pharmacologic and behavioral intervention strategies. Progress in this research has been hindered by a paucity of valid and reliable neurocognitive assessment instruments with alternate forms (McCaffrey & James, 1995; McCaffrey, Ortega, Orsillo, Nelles, & Haase, 1992). This absence has been particularly striking for measures of executive-function, where practice effects are especially prevalent (e.g., Bornstein, Baker, & Douglass, 1987). In such tasks participants are asked to generate novel strategies and recognize abstract concepts. Recall of previously generated strategies from experience during the first test administration make the neurocognitive skills assessed by an executive function task after an intervention quite different from those assessed at baseline. This observation has been confirmed by recent empirical studies showing significant improvement on many commonly used measures of executive-function in the absence of treatment.

For example, Basso, Bornstein, and Lang (1999), in a study of 50 healthy males evaluated the effects of re-administration of the Wisconsin Card Sorting Test (WCST), Ruff Figural Fluency Test (RFFT), the Verbal Concept Attainment Test (VCAT), Trail Making Test (TMT), and phonemic word-list generation (F-A-S) verbal fluency after a 12-month interval. This interval was selected to mimic the longest interval most likely to be used in intervention studies. Results revealed that WCST, RFFT, and VCAT scores showed significant practice effects across the test-retest interval. Notably, the mean number of perseverative errors and perseverative responses on the WCST decreased by nearly 50% on retesting. Clearly, detecting the effects of an intervention against a background of test-retest improvements of this magnitude is difficult.

To help address these difficulties, the Penn Conditional Exclusion Test (PCET) was developed as a new measure of executive-function with four alternate forms. The task was computerized for ease of administration and scoring and to facilitate its use in future functional neuroimaging studies in which individual participants receive serial scans during completion of different forms of the task before and after experimental therapeutic interventions. This task is a variant of the "Odd Man Out" paradigm (Flowers & Robertson, 1985), in which participants are asked to decide which of four objects does not belong with the other three. Objects are selected based on one of three sorting principles for each task. Participants have to infer the correct sorting principle based on response feedback and when the subject achieves 10 consecutive correct responses, the sorting principle shifts. There are three sorting principles for each version of the task and sorting principles are unique for each of the four test versions.

The goals of this study were to: (1) establish equivalent difficulty level of the four test versions (Study 1), and (2) establish preliminary construct validity (Study 2). To achieve this final goal, performance on the PCET was compared to performance on measures from the Penn Computerized Neurocognitive Scan administered according to standard procedures (Gur, Ragland, Moberg, Turner, et al., 2001) in healthy subjects. A future study will investigate the utility of the PCET in the assessment of executive-function impairment in patients with schizophrenia.

2. Methods and materials

2.1. Participants

For Study 1, 80 healthy volunteers were recruited from undergraduate introductory Psychology classes at Drexel University, Philadelphia, PA. After informed, written consent was obtained subjects were screened with a self-report measure and were disqualified for evidence of history of Axis I or II disorder, including evidence of substance abuse and, history of medical illness that might affect brain function, including evidence of a learning disability. The sample consisted of 47 males and 33 females with an average age of 19.3 ± 2.2 years (range: 14–30 years), and 13.1 ± 1.5 years of education (range: 11–18 years).

For Study 2, healthy controls were drawn from an ongoing longitudinal investigation of brain function in schizophrenia conducted at the Schizophrenia Research Center at the University of Pennsylvania. After informed, written consent was obtained, participants underwent comprehensive screening and assessment, performed by the clinical research team (Shtasel et al., 1991). This process included screening with the Non-Patient Edition of the SCID (Spitzer, Williams, & Gibbon, 1996). Entry criteria to the Schizophrenia Research Center for healthy controls included: (1) no Axis I or II disorder, including past or present substance abuse or dependence, (2) no history of a medical illness that might affect brain function, (3) no history of neurological disorder (e.g., epilepsy, migraine, head trauma with loss of consciousness), and (4) no first-degree relatives with a diagnosis of schizophrenia or affective illness. The sample consisted of 17 males and 8 females with an average age of 31.4 ± 11.6 years (range: 18–66), and 14.7 ± 2.2 years of education (range: 10–19 years).

2.2. Experimental procedures

For Study 1, to account for potential order effects, participants were administered the four versions of the PCET in a counterbalanced Latin-square design in which each form of the test appeared in each of four possible positions an equal number of times. For Study 2, the PCET (Form 1) was administered as part of a battery of valid and reliable computerized neuropsychological tests (Gur, Ragland, Moberg, Bilker, et al., 2001; Gur, Ragland, Moberg, Turner, et al., 2001). The PCET, along with the rest of the neuropsychological battery was administered using the PowerLaboratory® platform (Chute & Westall, 1997). Total administration time for the complete battery was approximately 1.5 h.

3. Computerized neuropsychological test battery

This Penn Computerized Neurocognitive Scan (Gur, Ragland, Moberg, Bilker, et al., 2001; Gur, Ragland, Moberg, Turner, et al., 2001) was administered to healthy controls along with the PCET (Form 1) and the AIM as part of Study 2. This broader battery consisted of the Penn Continuous Performance Test (PCPT; Kurtz, Ragland, Bilker, Gur, & Gur, 2001) during which a participant is asked to respond to a set of vertical and horizontal lines whenever they form a digit; The Penn Emotion Recognition Test, a forced choice test in which a participant is asked to determine which of two faces shows greater intensity of happiness or sadness; The Penn Word Memory Test (PWMT; Gur et al., 1993) a forced choice recognition task in which participants are shown 20 target words and are asked to remember them; The Penn Face Memory Test (PFMT; Gur et al., 1993) which parallels the PWMT but consists of 20 faces; The Penn Verbal Reasoning Test (PVRT; Gur, Gur, Obrist, Skolnick, & Reivich, 1987), a test that includes 30 verbal analogies; Computerized Judgment of Line Orientation (CJOLO) a computerized adaptation of the original paper-and-pencil task (Benton, Hannay, & Varney, 1975).

3.1. PCET

Each trial of the four alternate forms of the PCET consisted of four figures oriented horizontally in which the subject is required to use the mouse to click on the figure that is different from the other three. For each trial, the participant selects one of the four items that does not belong with the other three based on one of three separate criteria. Upon selecting an item, the participant is presented with immediate feedback via a screen stating that their choice is "correct" or "incorrect" for 500 ms. These sorting criteria were nonoverlapping for the four test versions.

Test stimuli were produced in Corel WordPerfect using the drawing toolbar. Stimuli were then imported to Adobe Photoshop where images were modified. Directions for each version were read from the computer monitor and were as follows: "In this task you will be shown four objects. You will need to determine which object does not belong with the other three. When you correctly choose the object that does not belong you will be told 'correct.' If you choose an incorrect item you will be told 'incorrect."' For each presentation of four stimuli there is text at the bottom of the page that states, "click on the object that does not belong." For Form 1 the categories were: (1) line thickness, (2) shape, and (3) size (see top of Fig. 1). For Form 2, the categories were: (1) location of shape within each stimulus box, (2) color, and then (3) type of shape within the box (see bottom of Fig. 1). For Form 3, the sorting principles were: (1) angle of lines within the stimulus box, (2) dashed versus solid lines within the stimulus box, and

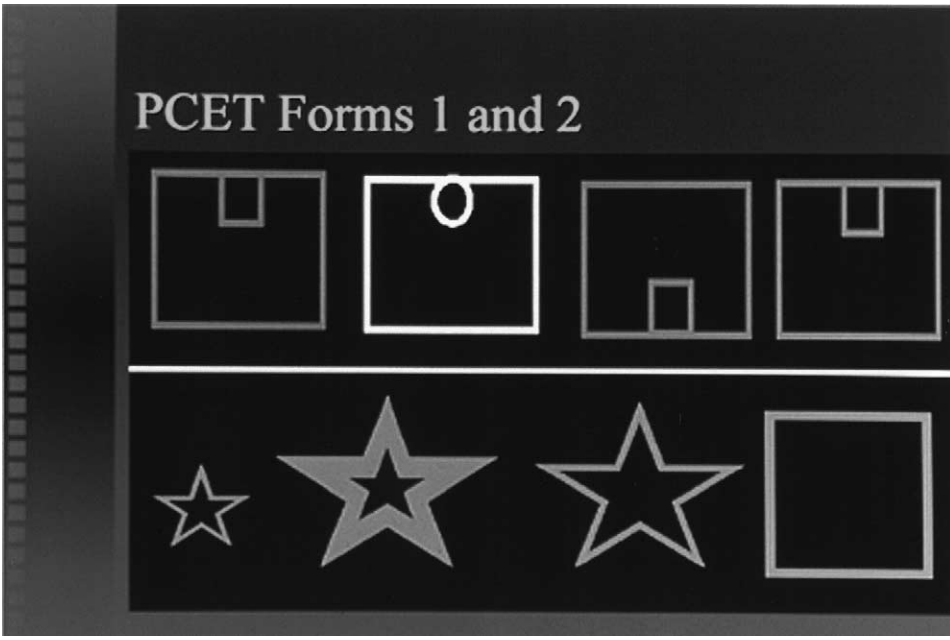

Fig. 1. Examples of stimuli from Forms 1 and 2 of the Penn Conditional Exclusion Test.

(3) presence or absence of a border on the stimulus box were the three sorting principles (see top of Fig. 2). For Form 4, the three sorting principles were: (1) sharp or rounded edges, (2) angle of stimulus box, and (3) whether the stimulus box was filled or empty (see bottom of Fig. 2). On average, each version of the PCET took 10 min to complete. If the participant is unable to achieve a single category the test ends after 144 trials.

3.2. AIM task

The AIM task consists of an Abstraction and Abstraction plus Memory subtest. The Abstraction subtest is based on a paradigm developed by Dr. Levy Rahmani (Rahmani, Geva, Rochberg, Trope, & Bore, 1990) in which participants are shown five shapes. Two shapes appear in the upper right-corner of the computer screen and two shapes in the upper left-corner of the screen. A fifth target object appears in the center of the screen, below the other stimuli. The goal is to pair the target object with one of the two sets of objects on either the left or right. Categories for pairings included: (1) all three objects have the same shape but each a different color, (2) all three objects have the same shape and two have the same color, (3) two objects have the same shape and two have the same color, (4) two have the same shape but all have different color, and (5) two objects have the same shape and two the same color, however, the target object cannot have the same color and shape as the top objects. The Abstraction plus Memory subtest is identical to the Abstraction subtest except that the target stimulus is presented for 500 ms then disappears and there is a 2.5-s delay before the appearance of the four shapes at the top of the screen. Previous research has shown significant correlations between performance on both Abstraction and

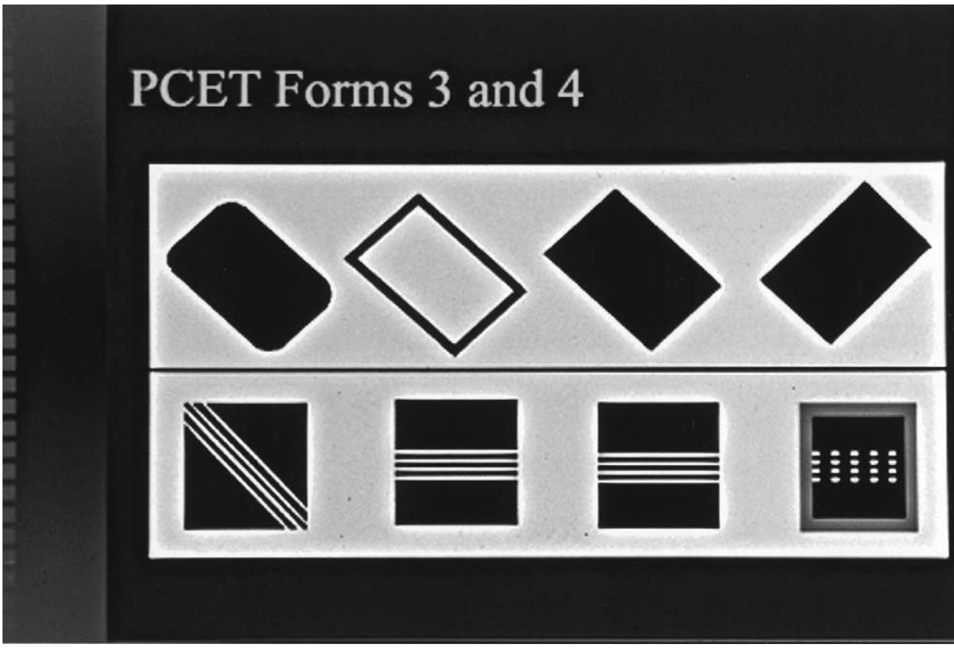

Fig. 2. Examples of stimuli from Forms 3 and 4 of the Penn Conditional Exclusion Test.

Abstraction + Memory subtests from the AIM and categories achieved and perseverative errors from the Wisconsin Card Sorting Test (WCST; Glahn, Cannon, Gur, Ragland, & Gur, 2000).

3.3. Hypothesis tests

Statistical analysis addressed the following hypotheses: (1) performance (i.e., total number of categories achieved, errors, trials to first category, and total number of trials) does not differ between the four versions of the PCET, (2) PCET performance correlates with AIM performance, (3) PCET performance does not correlate with other measures that are not linked to problem-solving, for example, visuospatial ability (Judgment of Line Orientation), word memory (Penn Word Memory Test), face memory (Penn Face Memory Test), facial emotion recognition (Penn Emotion Recognition Test), and verbal reasoning (Penn Verbal Reasoning Test) (divergent validity). Establishing convergent and divergent validity in hypotheses 2 and 3 will provide preliminary evidence of construct validity (Cronbach & Meehl, 1955).

To test the first hypothesis a series of four within-group ANOVAs were conducted on four dependent measures from the four forms of the PCET. The intraclass correlation coefficient (ICC) for each dependent measure from the PCET was computed to estimate the degree of correspondence among several forms of a test (Bartko & Carpenter, 1976). To test the second and third hypotheses, Pearson correlations were performed on PCET categories achieved and number of errors and the AIM subtests of Abstraction and Abstraction and Memory and the other measures from the Penn Computerized battery.

Table 1
Comparison of performance on four alternate forms of the Penn Conditional Exclusion Test in healthy people ($n = 80$)

	PCET 1		PCET 2		PCET 3		PCET 4				
Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	F	P	ICC
CAT	2.9	0.3	2.9	0.3	2.9	0.5	2.9	0.4	0.5	n.s.	.27*
Errors	15.3	12.9	12.4	14.2	15.3	16.9	12.0	11.6	2.1	n.s.	.38*
Trials to 1st category	18.4	12.1	16.9	16.7	14.6	5.1	14.3	7.5	3.0	.03	.21*
Trials	54.6	20.5	48.5	19.8	50.1	21.3	49.0	17.3	2.6	n.s.	.39*

Note. CAT: categories achieved; ICC: intraclass correlation; n.s.: nonsignificant.
* $P < .0001$.

4. Results

4.1. Study 1

4.1.1. Comparison of four alternate forms

The performance of healthy controls on the four versions of the PCET was compared using one-way within-subject ANOVAs and intraclass correlations as can be seen in Table 1. With the exception of trials to first category, none of the four forms of the PCET were significantly different from any of the other versions on any of the dependent measures. Mean categories achieved, mean total errors and mean total number of trials for each of the four versions were closely related. These results were confirmed by significant intraclass correlation coefficients of moderate size for each of these dependent measures (ICC = .27, $P < .0001$; ICC = .38, $P < .0001$; ICC = .39, $P < .001$, respectively). Even for trials to first category, where scores for the four forms were significantly different, a mean of only 4.1 responses separated the smallest number of responses for solving the first category (Form 4) to the largest (Form 1). The intraclass correlation coefficient for this dependent measure was significant (ICC = .21, $P < .0001$).

4.2. Study 2

4.2.1. Correlations between PCET, AIM, and other neuropsychological measures

Performance of healthy participants on the PCET and other tests in the battery is presented in Table 2. The number of variables examined from the broader neuropsychological battery was limited to nine a priori summary measures to contain Type 1 (experiment-wise) error. Categories achieved and total errors from the PCET (Form 1) were highly correlated with total correct on the Abstraction subtest from the AIM ($r = .77$, $P < .005$, and $r = -.61$, $P < .005$, respectively). Categories achieved from the PCET also correlated with total correct on the Abstraction + Memory subtest ($r = .50$, $P < .05$). Providing evidence of divergent validity, there were small and nonsignificant correlations between the PCET and a verbal memory test (Penn Word Memory Test; PWMT) and errors from the PCET and a verbal reasoning task (Penn Verbal Reasoning Test). The PCET did not correlate with measures of facial emotion recognition (Penn Emotion Recognition Task), face memory (Penn Face Memory Task), and

Table 2
Correlations between PCET variables, AIM subtests and other measures in healthy subjects ($n = 25$)

	PCET			AIM	
	Mean (S.D.)	PCET categories	PCET errors	Abstraction	Abstraction + Memory
PCET Cat	2.5 (0.8)	–	−.92**	.77**	.50*
PCET Err	32.4 (22.4)	–	–	−.61**	−.41
AIM Abst	24.0 (3.5)	–	–	–	.55**
PCPT Eff	0.07 (0.01)	.49*	−.59**	−.04	.11
EMOREC	28.0 (4.5)	.05	−.13	.12	.05
PWMT	36.3 (4.7)	.18	−.31	.06	−.04
PFMT	32.8 (3.7)	−.09	.11	.08	.51*
CJOLO	23.9 (5.0)	−.03	−.13	−.22	.16
PVRT	20.1 (5.5)	.06	−.23	.02	.18

Note. Pearson correlations between PCET variables, AIM subtests and computerized neuropsychological measures: categories achieved on the Penn Conditional Exclusion Test (PCET Cat); total errors on the Penn Conditional Exclusion Test (PCET Err), true positives on the Penn Continuous Performance Test (PCPT) divided by average reaction time (PCPT Eff), true positives for happy and sad face discrimination (EMOREC), average of true positives at an immediate test and after a 30-min delay on the Penn Word Memory Test (PWMT), average of true positives at an immediate test and after a 30-min delay on the Penn Face Memory Test (PFMT), total correct on the Computerized Judgment of Line Orientation (CJOLO), total correct on the Penn Verbal Reasoning Test (PVRT).

* $P < .05$.

** $P < .005$.

visuospatial function (Judgment of Line Orientation). Categories achieved and errors also correlated with efficiency measures from the PCPT ($r = .49$, $P < .001$, and $r = -.59$, $P < .001$, respectively) suggesting a close link between visual sustained attention and performance on the PCET.

Correlations between the PCET and AIM with the rest of the neuropsychological battery were also compared. As can be seen in Table 2, with the exception of the Penn CPT and the Penn Face Memory Test, the PCET and AIM did not correlate with any of the measures from the neuropsychological battery. Thus, the pattern of correlations between the PCET, the AIM with other neuropsychological measures was similar, and both appeared to assess specific rather than generalized abilities.

5. Discussion

Two major findings emerged from this study. Results from Study 1 indicated that the four forms of the PCET are of similar difficulty as all forms were significantly intercorrelated. With the exception of trials to first category, means on a variety of dependent measures from the PCET were highly similar. The magnitude of the intraclass correlation coefficients for the dependent measures for the four forms of the PCET were consistent with those reported for other commonly used alternate forms of neuropsychological tests, for example, The Hopkins Verbal Learning Test (Brandt, 1993). Results from Study 2 indicated that the PCET assesses executive-function and not other related neuropsychological domains in healthy people. There was evidence of both convergent and divergent validity. With respect to convergence, the PCET showed high correlations with another measure of abstraction, the Abstraction subtest of the AIM. This magnitude of correlation suggests significant overlap in the cognitive skills engaged by these tasks. With regard to divergence, the PCET did not correlate with measures of facial emotion recognition, face and word memory, visuospatial function, and verbal reasoning. With the exception of a correlation between the PCET and the Penn CPT, and one between the Abstraction and Memory subtest of the AIM and the Penn Face Memory Test, the PCET correlated with other measures in the computerized battery in a manner nearly identical to that of the AIM.

Several caveats regarding the present study should be noted. First, while the convergent and discriminant approach to validation pursued in the current study is a well-established first-step for construction validation for any new cognitive task (e.g., Cronbach & Meehl, 1955), and the construct validation reported in this study was established with neurocognitive measures that *have* been validated behaviorally in samples of patients with known brain disorders (Gur, Ragland, Moberg, Bilker, et al., 2001), this approach would have been strengthened with the addition of behavioral or functional imaging data from patients with brain disorders. This complementary, criterion-related approach to validation will be necessary before this measure may be used effectively in a clinical setting. In addition, in the absence of such data, the specific neural circuits activated by the PCET remain unknown. Second, additional normative data with equivalent numbers of both men and women and a wide range of ages and educational levels will be necessary to adequately assess impairment in patient populations on this task. Third, validation studies with a wider range of neuropsychological test instruments and larger sample sizes would strengthen the findings reported in this study. Limited sample-size in Study 2 suggests that construct validation should be replicated in larger samples of healthy people. In addition, factor analytic approaches in large sample studies could help clarify the relationship of performance on the PCET to performance on tests that tap other neuropsychological functions such as visual sustained attention. Fourth, the nature of the PCET, in which the subject initially investigates different strategies, learns the correct selection rule, and then after 10 correct responses has to shift rules based on response feedback precluded division of the test into equivalent halves for computation of an internal measure of stability (e.g., Spearman split-half reliability). Thus, the reliability of this measure is still unclear.

We have started to apply this measure to patients with schizophrenia, as well as first-degree family members. Future research with this measure will be aimed at understanding the relationship of deficits on this measure to performance in activities of daily living. While the gap between cognitive test performance and "real-world" performance is admittedly large, recent research in schizophrenia suggest that several cognitive skills, including card sorting, visual vigilance and secondary verbal memory are closely related to both clinician-rated and performance-based measures of psychosocial status (Green, 1996). Understanding this relationship will be crucial in order to identify the predictive validity of performance on the PCET for functioning in a variety of real-world settings.

There are several advantages to the PCET over existing measures of executive-function. First and most importantly, four alternate forms of similar difficulty permit multiple serial assessments of change in neuropsychological function in the same patient over time without the confounding effects of practice on the same task. Second, the brevity of the task (average

10 min in healthy participants) makes it shorter than many currently available measures of executive function. Third, as a computerized task administration and scoring is automated reducing test-retest variability associated with analogous paper-and-pencil measures. This also makes administration of the task during functional imaging procedures more feasible. Fourth, while a few simple primary measures were selected in order to minimize experiment-wise errors for the present study, additional measures such as perseverative errors, conceptual level responses and measures of loss of category may be generated for performance on the PCET in a manner analogous to the WCST. These variables may provide increased specificity regarding the nature of impairment on this task in individual patients.

Acknowledgments

We thank Brian C. Klimas, B.A., for his aid in the development of the PCET stimuli. This research was funded by grants MH-19112, MH-43880, MH-01336, and MO1RR0040 from the National Institutes of Mental Health. The first author was also supported by grant R03 MH-65377 from the National Institutes of Mental Health and gifts from the Spencer T. and Ann W. Olin Foundation to the Institute of Living, Hartford Hospital's Mental Health Network during preparation of this article.

References

Bartko, J. J., & Carpenter, W. T. (1976). On the method and theory of reliability. *Journal of Nervous and Mental Disease*, *163*, 307–317.

Basso, M. R., Bornstein, R. A., & Lang, J. M. (1999). Practice effects on commonly used measures of executive function across twelve months. *The Clinical Neuropsychologist*, *13*, 283–292.

Benton, A. L., Hannay, J. J., & Varney, N. R. (1975). Visual perception of line direction in patients with unilateral brain disease. *Neurology*, *25*, 907–910.

Bornstein, R. A., Baker, G. B., & Douglass, A. B. (1987). Short-term reliability of the Halstead–Reitan battery in a normal sample. *Journal of Nervous and Mental Disease*, *175*, 232–299.

Brandt, J. (1993). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist*, *5*, 125–142.

Chute, D. L., & Westall, R. F. (1997). *Power laboratory*. Devon, PA: MacLaboratory Incorporated.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Flowers, K. A., & Robertson, C. (1985). The effect of Parkinson's disease on the ability to maintain a mental set. *Journal of Neurology, Neursurgery, and Psychiatry*, *48*, 517–529.

Glahn, D. C., Cannon, T. D., Gur, R. E., Ragland, J. D., & Gur, R. C. (2000). Working memory constrains abstraction in schizophrenia. *Biological Psychiatry*, *47*, 34–42.

Green, M. F. (1996). What are the functional consequences of neurcognitive deficits in schizophrenia? *American Journal of Psychiatry*, *153*, 321–330.

Gur, R. C., Gur, R. E., Obrist, W. D., Skolnick, B. E., & Reivich, M. (1987). Age and regional cerebral blood flow at rest and during cognitive activity. *Archives of General Psychiatry*, *44*, 617–621.

Gur, R. C., Jaggi, J. L., Ragland, J. D., Resnick, S. M., Shtasel, D., Muenz, L., & Gur, R. E. (1993). Effects of memory processing on regional brain activation: Cerebral blood flow in normal subjects. *International Journal of Neuroscience*, *72*, 31–44.

Gur, R. C., Ragland, J. D., Moberg, P. J., Bilker, W. B., Kohler, C., Siegel, S. J., & Gur, R. E. (2001). Computerized neurocognitive scanning. II. The profile of schizophrenia. *Neuropsychopharmacology*, *25*, 777–788.

Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., Siegel, S. J., & Gur, R. E. (2001). Computerized neurocognitive scanning. I. Methodology and validation in healthy people. *Neuropsychopharmacology*, *25*, 766–776.

Heinrichs, R. W., & Zakzanis, K. K. (1998). Neurocognitive deficits in schizophrenia: A quantitative review of the evidence. *Neuropsychology*, *12*, 426–445.

Johnson-Selfridge, M., & Zalewski, C. (2001). Moderator variables of executive-function in schizophrenia: Meta-analytic findings. *Schizophrenia Bulletin*, *27*, 305–316.

Kurtz, M. M., Ragland, J. D., Bilker, W. B., Gur, R. C., & Gur, R. E. (2001). Comparison of two forms of the Continuous Performance Test, with and without working memory demands, in healthy controls and patients with schizophrenia. *Schizophrenia Research*, *48*, 307–316.

McCaffrey, R. J., & James, H. J. (1995). Issues associated with repeated neuropsychological assessments. *Neuropsychology Review*, *5*, 203–221.

McCaffrey, R. J., Ortega, A., Orsillo, S. M., Nelles, W. B., & Haase, R. F. (1992). Practice effects in repeated neuropsychological assessments. *The Clinical Neuropsychologist*, *6*, 32–42.

Purdon, S. R. E., Jones, B. D., Stip, E., Labelle, A., Addington, D., David, S., Breier, A., & Tollefson, G. (2000). Neuropsychological change in early phase schizophrenia during 12 months of treatment with olanzapine, risperidone or haloperidol. *Archives of General Psychiatry*, *57*(3), 249–258.

Rahmani, L., Geva, N., Rochberg, J., Trope, I., & Bore, B. (1990). Issues in the neurocognitive assessment and training. In E. Vakil, D. Hoffein, & Z. Groswasser (Eds.), Rehabilitation of the brain injured person: A neuropsychological perspective (pp. 43–60). New York: Freund Publishing House.

Shtasel, D. L., Gur, R. E., Mozley, P. D., Taleff, M. M., Heimberg, C., Gallacher, F., & Gur, R. C. (1991). Volunteers for biomedical research: Recruitment and screening of normal controls. *Archives of General Psychiatry, 48*, 1022–1025.

Spitzer, R. L., Williams, J. B., & Gibbon, M. (1996). *Structured clinical interview for DSM-IV: Non-patient version (SCID-NP)*. New York: New York State Psychiatric Institute.