Journal of the International Neuropsychological Society (2004), **10**, 301–303.
Copyright © 2004 INS. Published by Cambridge University Press. Printed in the USA.
DOI: 10.1017/S1355617704102191

LETTER TO THE EDITOR

Reliability and validity of the Delis-Kaplan Executive Function System: An update

DEAN C. DELIS,[1] JOEL H. KRAMER,[2] EDITH KAPLAN,[3] AND JAMES HOLDNACK[4]

[1]Department of Veterans Affairs Medical Center, San Diego, and University of California, San Diego, School of Medicine
[2]University of California, San Francisco, School of Medicine
[3]Suffolk University and Boston University School of Medicine
[4]The Psychological Corporation

INTRODUCTION

A critical endeavor in every health-related field is the continued development of new technologies and instrumentation for improving diagnosis and treatment. For example, the field of neuroimaging has made astonishing advances with procedures such as functional MRI, MR spectroscopy, and MR diffusion tensor imaging. The field of neuropsychology has also recognized the importance of continually striving to develop new, more sensitive clinical measures. At the same time, the development of new tests or procedures does not guarantee that the instruments represent improvements in diagnosis or treatment. For this reason, ongoing studies of reliability and validity of new diagnostic instruments are also an important part of research in all health-related fields.

It was in the spirit of advancing the instrumentation of neuropsychology that we developed and recently published nine new executive-function tests called the Delis-Kaplan Executive Function System (D-KEFS; Delis et al., 2001a). While there have been a number of research studies demonstrating the usefulness of D-KEFS tests (see below), Schmidt (2003) recently raised questions about the reliability and validity of the instruments. For example, in reviewing the D-KEFS, Schmidt (2003) stated, "Evidence regarding validity is extremely weak. Data are presented for two small clinical groups. . . ." In this paper, we address these questions.

VALIDITY STUDIES

There appears to be a misperception by Schmidt (2003) that the D-KEFS tests have little validity data to support their use. It is the case that the D-KEFS manual contains little *new* validity data. However, early in the development of the D-KEFS, we elected to publish our validity data in peer-review scientific journals rather than in the test manual. We chose this approach for several reasons. First, the development of the D-KEFS took over 10 years, and we felt that scientific journals offered a more efficient way to disseminate the validity results. Second, validity data published in peer-reviewed journals allow all neuropsychologists the opportunity to evaluate the results of the studies, not just those who have purchased the test manual. Third, and most important, validity data published in peer-review journals have the added safeguards that go with the scientific review process, namely, that one's research methods and conclusions are deemed sound and appropriate by senior researchers who serve on the editorial boards of these journals (such safeguards do not exist for validity data published only in test manuals).

The D-KEFS test manual does provide references of the validity studies that had been conducted at the time of its publication, but these references appear to have been missed or overlooked by Schmidt (2003). In addition, studies of the D-KEFS published in the literature after the publication of the test manual also appear to have been overlooked by Schmidt (2003). Rather, Schmidt (2003) seems to focus on only one section of the test manual where we presented some new pilot data from small samples of patients with Alzheimer's or Huntington's disease, and to incorrectly regard these data as our primary attempt to conduct validity studies of the D-KEFS. However, we presented these pilot data in the manual to illustrate that, even in more diffuse dementia populations, intriguing dissociations can be found in component processes of executive functions as assessed by the D-KEFS.

As an update, we are providing here a list of validity studies that have been conducted with various D-KEFS tests both before and after the publication of the test manual. These studies have demonstrated that D-KEFS tests are sensitive to the assessment of executive-function deficits in numerous clinical populations, including patients with focal frontal-lobe lesions (Baldo et al., 2001; Baldo et al.,

Reprint requests to: Dean C. Delis, Ph.D., Psychology Service (116B), V.A. Medical Center, 3350 La Jolla Village Drive, San Diego, CA 92161; ddelis@ucsd.edu

https://doi.org/10.1017/S1355617704102191 Published online by Cambridge University Press

2003; Baldo et al., in press; Delis et al., 1992; Dimitrov et al., 1999); focal ventromedial prefrontal damage (Cato et al., in press); frontal-temporal dementia (Hallam et al., in press); mild cognitive impairment (Grant et al., in press); subcortical ischemic disease (Kramer et al., 2002); lateralized right-hemisphere damage (Crouch et al., 1996); Parkinson's disease (Beatty et al., 1990; Bondi et al., 1993; Dimitrov et al., 1999); multiple sclerosis (Beatty et al., 1995); chronic alcoholics (Beatty et al., 1993); normal aging (Wecker et al, 2000); autism (Kleinhans et al., 2003); Asperger's syndrome (Kleinhans et al., 2003); schizophrenia (Beatty et al., 1994); psychopathy (Kalinian, 2003); childhood stroke (Nichols et al., 2003); attention deficit disorder (Donnelley et al., 2001); and fetal alcohol syndrome (Mattson et al., 1999; Schonfeld et al., 2001). In addition, studies have been conducted that compared the D-KEFS Sorting Test and the Wisconsin Card Sorting Test in the assessment of executive-function deficits in clinical populations (Beatty et al., 1990; 1994). These and other D-KEFS publications are listed in the reference section below (note that the D-KEFS Sorting Test was formerly called the California Card Sorting Test).

The study of a test's validity is an ongoing process, and there is always a need for further investigations of all neuropsychological tests but particularly for new tests such as the D-KEFS. We should point out that there has been a lack of validity studies conducted for one of the D-KEFS instruments, the Proverbs Test. More research is especially needed for this test regarding its sensitivity to deficits in verbal abstraction skills.

In the past, a common approach to examining the validity of an assessment instrument has been to conduct factor-analytic studies of the test's variables using the instrument's normative sample or mixed clinical populations. However, we recently published a study that demonstrated empirically serious limitations of this approach in examining the validity of process-oriented neuropsychological tests (Delis et al., 2003). In writing an invited commentary on our factor-analytic study, Larrabee (2003), a leader in the field of factor analysis, agreed with our view and pointed out other limitations in past validity studies that have used factor-analytic techniques. It was for these reasons that we did not conduct factor-analytic studies of the D-KEFS tests or develop across-test composite indices based on factor-analytic results.

RELIABILITY

We conducted numerous reliability analyses of the D-KEFS measures, and these data are published in the D-KEFS technical manual (Delis et al., 2001b). Schmidt (2003) was also critical of these data, stating that only 17% of the reliability values published in the D-KEFS manual (53 of the 316 coefficients) were above a .80 value. However, the reliability coefficients found for the D-KEFS tests are comparable to those published for a number of other commonly used neuropsychological tests, including the Wisconsin Card Sorting Test and Wechsler Memory Scale–III. For example, none of the reliability coefficients of the Wisconsin Card Sorting Test are above a .80 value (Heaton et al., 1993). It may be the case that tests such as the D-KEFS, Wisconsin Card Sorting Test, and memory instruments tap a wider spectrum of complex, effortful cognitive processes compared to more homogeneous, fundamental tasks such as the Vocabulary or Picture Completion subtests, and as a result, the potential for performance variability or measurement error may be greater for more complex tests. However, it is often the complexity of these tasks that make them so sensitive to the detection of even mild brain damage. In addition, tests of executive functions and memory often pose special problems for calculating reliability in traditional ways (e.g., item-interdependence makes it difficult to use split-half reliabilities; see Delis et al., 2000). Although caution must always be exercised in ascribing brain damage to low scores on any test, especially for more complex tasks such as the Wisconsin Card Sorting Tests, D-KEFS tests, and memory instruments, the utility of these instruments for detecting neurocognitive deficits has been demonstrated in numerous studies.

CONCLUSION

In general, all neuropsychological tests have their strengths and limitations, including the D-KEFS. The process of selecting tests is analogous to the decision whether to use CT versus MRI brain scans, with each imaging procedure being particularly useful for different conditions. For example, CT brain scans tend to be more sensitive to the detection of acute hemorrhages, whereas MRI brain scans tend to be better at revealing small, subcortical infarcts. An analogous selection process applies to neuropsychological tests. For instance, we often administer the Wisconsin Card Sorting Test to high-functioning individuals who may have a subtle deficit in concept formation and flexibility. In contrast, we prefer to administer the D-KEFS Sorting Test to patients with more moderate to severe cognitive dysfunction, because the D-KEFS task (a) is often helpful in revealing the neurocognitive mechanisms underlying the problem-solving deficit (e.g., due in part to the fact that the patient is asked to describe his or her sorting strategy); and (b) tends not to be as frustrating as the Wisconsin (e.g., due to the absent of right-wrong feedback). The D-KEFS was never intended to assess all aspects of frontal-lobe functioning (e.g., emotional or behavioral changes), and therefore, as we state in the manual, it is imperative to supplement the examination of a patient with other neuropsychological tests. It is our hope that the D-KEFS tests assist neuropsychologists by providing them with a larger armamentarium for selecting the right assessment instruments for the right patients.

REFERENCES

Baldo, J.V., Shimamura, A.P., Delis, D.C., Kramer, J., & Kaplan, E. (2001). Verbal and Design Fluency in Patients with Frontal

https://doi.org/10.1017/S1355617704102191 Published online by Cambridge University Press

Lobe Lesions. *Journal of the International Neuropsychological Society*, *7*, 586–596.

Baldo, J.V., Wilkins, D.P., Shimamura, A.P., Kramer, J., Kaplan, E., & Delis, D.C. (in press). Is it bigger than a breadbox? Performance of frontal patients on a new test of concept formation. *Archives of Clinical Neuropsychology*.

Baldo, J.V., Delis, D.C., & Kaplan, E. (2003). Role of executive functions in language: Evidence from a new verbal reasoning test. *Journal of the International Neuropsychological Society*, *9*, 183.

Beatty, W.W., Hames, K.A., Blanco, C.R., & Paul, R.H. (1995). Verbal abstraction deficit in multiple sclerosis. *Neuropsychology*, *9*, 198–205.

Beatty, W.W., Jocic, Z., Monson, N., & Katzung, V.M. (1994). Problem solving by schizophrenic and schizoaffective patients on the Wisconsin and California card sorting tests. *Neuropsychology*, *8*, 49–54.

Beatty, W.W., Katzung, V.M., Nixon, S.J., & Moreland, V.J. (1993). Problem-solving deficits in alcoholics: Evidence from the California Card Sorting Test. *Journal of Studies on Alcohol*, *54*, 687–692.

Beatty, W.W. & Monson, N. (1990) Problem solving in Parkinson's disease: Comparison of performance on the Wisconsin and California card sorting test. *Journal of Geriatric Psychiatry & Neurology*, *3*, 163–171.

Bondi, M.W., Kasniak, A.W., Bayles, K.A., & Vance, K.T. (1993). Contributions of frontal system dysfunction to memory and perceptual abilities in Parkinson's disease. *Neuropsychology*, *7*, 89–102.

Cato, M.A., Delis, D.C., Abildskov, T.J., & Bigler, E. (2004). Assessing the elusive cognitive deficits associated with ventromedial prefrontal damage: A case of a modern-day Phineas Gage. *Journal of the International Neuropsychological Society*, *10*, in press.

Crouch, J.A., Greve, K.W., & Brooks, J. (1996). The California Card Sorting Test may dissociate verbal and non-verbal concept formation abilities. *British Journal of Clinical Psychology*, *35*, 431–434.

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001a). *The Delis-Kaplan Executive Function System: Examiner's Manual*. San Antonio: The Psychological Corporation.

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001b). *The Delis-Kaplan Executive Function System: Technical Manual*. San Antonio: The Psychological Corporation.

Delis, D.C., Kramer, J.H., Kaplan, E. & Ober, B.A. (2000). *The California Verbal Learning Test–Second Edition*. San Antonio: The Psychological Corporation.

Delis, D.C., Squire, L.R., Bihrle, A.M., & Massman, P.J. (1992). Componential analysis of problem-solving ability: Performance of patients with frontal lobe damage and amnesic patients on a new sorting test. *Neuropsychologia*, *30*, 683–697.

Delis, D.C., Jacobson, M., Bondi, M.W., Hamilton, J.M., and Salmon, D.P. (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: Lessons from memory assessment. *Journal of the International Neuropsychological Society*, 9, 936–946.

Dimitrov, M., Grafman, J., Soares, A.H., & Clark, K. (1999). Concept formation and concept shifting in frontal lesions and Parkinson's disease patients assessed with the California Card Sorting Test. *Neuropsychology*, *13*, 135–143.

Donnelly JF, Carte E, Kramer JH, Zupan B, & Hinshaw S. (2001). Executive functioning in girls with subtypes of ADHD. *Journal of the International Neuropsychological Society*, *7*, 201.

Grant, D., Kramer, J.H., Rosen, H., Schauer, G., Pace-Savitsky, K., Rankin, K., Johnson, J., Weiner, M., & Miller, B.L. (abstract in press). Neuroanatomy of semantic and phonemic fluency. *Journal of the International Neuropsychological Society*.

Hallam, B. J, Kramer, J.H., Rosen, H.J., Schauer, G., Pace-Savitsky, C., Delis, D.C., Weiner, M., & Miller, B.L. (abstract in press). Frontal lobe volumes specifically predict the capacity for set shifting on executive tasks. *Journal of the International Neuropsychological Society*.

Han, D.S., Delis, D.C., Holdnack, J.A., Kramer, J.H., & Houston, W.S. (abstract in press). Relationships between IQ and performance on the Delis-Kaplan Executive Function System in children. *Journal of the International Neuropsychological Society*.

Heaton, R.K., Chelune, G.J., Talley, J.L., Kay, G.G., & Curtiss, Glenn (1993). *Wisconsin Card Sorting Test Manual*. Odessa, FL: Psychological Assessment Resources.

Kalinian, H. (2003). Executive function performance as assessed by the D-KEFS in female psychopaths. *Archives of Clinical Neuropsychology*, *18*, 760–761.

Kleinhans, N.M., Akshoomoff, N.A., & Courchesne, E. (2003) Executive functioning in autism and Asperger's syndrome: Results from the D-KEFS. *Journal of the International Neuropsychological Society*, *9*, 273.

Kramer J.H., Reed, B.R., Mungas, D., Weiner, M.W., & Chui, H.C. (2002). Executive dysfunction in subcortical ischemic vascular disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *72*, 217–220.

Mattson, S.N., Goodman, A.M., Caine, C., Delis, D.C., & Riley, E.P. (1999). Executive functioning in children with heavy prenatal alcohol exposure. *Alcoholism: Clinical and Experimental Research*, *23*, 1808–1815.

Nichols, S., Waller, S., Trauner, D., & Delis, D.C. (2003). Performance of children with early focal brain damage on a new test of concept formation. *Journal of the International Neuropsychological Society*, *9*, 139.

Schmidt, M. (2003). Hit or miss? Insight into executive functions. *Journal of the International Neuropsychological Society*, *9*, 962–964.

Schonfeld, A.M., Mattson, S.N., Lang, A.R., Delis, D.C. and Riley, E.P. (2001). Verbal and nonverbal fluency in children with heavy prenatal alcohol exposure. *Journal of Studies on Alcohol*, *62*, 239–246.

Wecker, N.S., Kramer, J.H., Wisniewski, A., Delis, D.C., & Kaplan, E. (2000). Age effects on executive ability. *Neuropsychology*, 14, 409–414.

https://doi.org/10.1017/S1355617704102191 Published online by Cambridge University Press