UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Modifications to the CVLT-II Recognition Discriminability Indices to Enhance the Characterization of Recognition Memory Impairment in Healthy Aging and Neurodegenerative Disease

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Clinical Psychology

by

Lisa V. Graves

Committee in charge:

University of California San Diego

Professor Mark W. Bondi
Professor Lisa Delano-Wood
Professor Dean C. Delis

San Diego State University

Professor Paul E. Gilbert, Chair
Professor Sarah N. Mattson

2019

The Dissertation of Lisa V. Graves is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

San Diego State University

2019

DEDICATION

This work and all my related efforts are dedicated to my beloved late grandmother, Sylvia Hewitt Graves, whose life, love, and endurance despite the numerous and unfathomable hardships of Alzheimer's disease have inspired my commitment to better understand and appreciate the human memory.

May our memories of her live on always.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I must thank my mentor and Dissertation Committee Chair, Dr. Paul Gilbert, for the invaluable guidance and support he has provided to me throughout the course of this dissertation project and all of my graduate training. His clear talent and passion for mentoring and teaching are in my mind unparalleled, and I feel extremely fortunate to have benefited from his encouragement, wisdom, trust, and great sense of humor. In all honesty, it is impossible to find the right words to fully express how grateful I am for his mentorship and the impact he has had on my professional and personal development over the past several years. Paul – you are my mentor for life, and if I can be just a fraction of the person you are, I will consider myself extremely successful.

I must also thank my other Dissertation Committee Members, each of whom provided significant contributions to this work and played an integral role in shaping my rewarding experience as a doctoral student. Dr. Dean Delis – thank you for sharing your contagious enthusiasm for neuropsychological research, and for welcoming our collaborations on multiple research projects involving the California Verbal Learning Test, including those reflected in this dissertation project. Dr. Mark Bondi and Dr. Lisa Delano-Wood – thank you both not only for your significant contributions to this work, but for the impactful mentorship and support you have provided to me as members of my Guidance Committee. To Dean, Mark, and Lisa – thank you also for your invaluable clinical supervision. Dr. Sarah Mattson – thank you for your genuine enthusiasm, thoughtfulness, and feedback on this dissertation project, which undoubtedly improved the quality of this work.

I must also acknowledge other individuals who put forth significant efforts in the production of the manuscripts associated with this dissertation project: Dr. David P. Salmon, Dr. Jody Corey-Bloom, Dr. Steven Paul Woods, Dr. Heather M. Holden, and Ms. Emily Van Etten. To Heather and Emily – I will never forget the many laughs and bittersweet moments we shared in Suite 201.

From the bottom of my heart, I thank my loved ones (Mom, Dad, Christo, Matty, Jenny, Grandpa, and Tyler, especially) and my graduate school cohort (Allie, Joel, Tam, Holly, Tonya, Steph, Angela, Amy, and Mary) for the incredible patience, understanding, encouragement, and support they have provided to me throughout my time in graduate school. It truly takes a village, and I am so proud of and grateful for mine.

Last, but certainly not least, I sincerely thank each research participant and patient with whom I have had the pleasure of working with. Thank you all for sharing your experiences with me, and for reminding me why I chose to pursue this line of work.

Please note that Chapter 2, in full, is a reprint of the material as it appears in the *Journal of Clinical and Experimental Neuropsychology*, 2017, Taylor & Francis Group; Chapter 3, in full, is a reprint of the material as it appears in *Aging, Neuropsychology, and Cognition,* 2017, Taylor & Francis Group; and Chapter 4, in full, is a reprint of the material as it appears in the *Journal of the International Neuropsychological Society*, 2018, Cambridge University Press. I would like to thank Drs. Gilbert, Holden, Delano-Wood, Bondi, Salmon, Corey-Bloom, Woods, and Delis, as well as Ms. Emily Van Etten, for their contributions as coauthors to the manuscripts reflected in these chapters.

VITA

EDUCATION

San Diego State University (SDSU)/University of California, San Diego (UC San Diego)
Joint Doctoral Program (JDP) in Clinical Psychology – APA-Accredited

Ph.D.	Clinical Psychology – Major Area of Study: Neuropsychology	2019
M.S.	Clinical Psychology	2015

SDSU Master of Arts Program in Psychology

M.A.	Psychology – Emphasis: Behavioral and Cognitive Neuroscience	2014

UC San Diego

B.S.	Animal Physiology and Neuroscience	2009
B.A.	Psychology	2009

PROFESSIONAL INTERESTS

- Cognitive and functional performance in normal aging and neurodegenerative disease
- Multimodal approaches (integration of neuropsychological methods with biomarker, genetic, and neuroimaging data) to modeling cognitive and functional decline in older age
- Early detection and timely intervention of individuals at risk for dementia
- Neuropsychological assessment of adults with known or suspected neurological injury, with an emphasis on differential diagnosis of dementia

AWARDS

JDP Incentive Awards for Travel to International Neuropsychological Society Meeting | 2016–19

JDP Publication Award | 2018

SDSU Center for Clinical and Cognitive Neuroscience Student Travel Award | 2015 – 2016

James & Mary Crouch Memorial Scholarship | 2015 – 2016

Instructionally Related Activities Travel Award | 2012

PEER-REVIEWED PUBLICATIONS

Graves, L. V., Simone, S., Williams, M. E., Courville, T., Mattson, S. N., Delano-Wood, L., Bondi, M. W., Salmon, D. P., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2019). Revisiting total recognition discriminability in Huntington's and Alzheimer's disease: New insights from the CVLT-3. *Applied Neuropsychology: Adult.* Advance online publication. doi: 10.1080/23279095.2019.1605993

Graves, L. V., Holden, H. M., Van Etten, E. J., Delano-Wood, L., Bondi, M. W., Corey-Bloom, J., Salmon, D. P., Gilbert, P. E., & Delis, D. C. (2019). New intrusion

analyses on the CVLT-3: Clinical utility in distinguishing the memory disorders of Alzheimer's versus Huntington's disease. *Journal of the International Neuropsychological Society*. Advance online publication. doi: 10.1017/S1355617719000407

Williams, M. E., **Graves, L. V.**, DeJesus, S. Y., Holden, H. M., DeFord, N. E., & Gilbert, P. E. (2019). Spatial memory ability during middle age may depend on level of spatial similarity. *Learning & Memory, 16*(1), 20-23. doi:10.1101/lm.048280.118

DeFord, N. E., DeJesus, S. N., Holden, H. M., **Graves, L. V.**, Lopez, F. V., & Gilbert, P. E. (2019). Young and older adults may utilize different cognitive abilities when performing a spatial recognition memory test with varying levels of similarity. *The International Journal of Aging and Human Development*. Advance online publication. doi:10.1177/0091415019831443

Graves, L. V., Holden, H. M., Van Etten, E. J., Delano-Wood, L., Bondi, M. W., Salmon, D. P., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2018). New Yes/No Recognition memory analysis on the California Verbal Learning Test-3: Clinical utility in Alzheimer's and Huntington's disease. *Journal of the International Neuropsychological Society, 24*(8), 833-841. doi:10.1017/S1355617718000474

Graves, L. V., Van Etten, E. J., Holden, H. M., Delano-Wood, L., Bondi, M. W., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2017). Refining CVLT-II recognition discriminability indices to enhance the characterization of recognition memory changes in healthy aging. *Aging, Neuropsychology, and Cognition, 25*(5), 767-782. doi:10.1080/13825585.2017.1372358

Graves, L. V., Moreno, C. C., Seewald, M., Holden, H. M., Van Etten, E. J., Uttarwar, V., McDonald, C. R., Delano-Wood, L., Bondi, M. W., Woods, S. P., Delis, D. C., & Gilbert, P. E. (2017). Effects of age and gender on recall and recognition discriminability. *Archives of Clinical Neuropsychology, 32*(8), 972-979. doi:10.1093/arclin/acx204

Seewald, P. M., DeJesus, S. Y., **Graves, L. V.**, Moreno, C. C., Mattson, S. N., & Gilbert, P. E. (2017). Age-related differences on a new test of temporal order memory for everyday events. *Aging, Neuropsychology, and Cognition, 28*(3), 319-332. doi:10.1080/13825585.2017.1298716

Graves, L. V., Holden, H. M., Delano-Wood, L., Bondi, M. W., Woods, S. P., Corey-Bloom, J., Salmon, D. P., Delis, D. C., & Gilbert, P. E. (2017). Total recognition discriminability in Huntington's and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 39*(2), 120-130. doi:10.1080/13803395.2016.1204993

DeFord, N. E., Landy, K. M., Pirogovsky-Turk, E., Van Etten, E. J., **Graves, L. V.**, Salmon, D. P., Filoteo, J. V., & Gilbert, P. E (2016). The effect of interference on temporal order memory in individuals with Parkinson's disease. *Brain and Cognition, 107*, 30-36. doi:10.1016/j.bandc.2016.05.008

Sheppard, D. P., **Graves, L. V.**, Holden, H. M., Delano-Wood, L., Bondi, M. W., & Gilbert, P. E. (2016). Spatial pattern separation varies in older adult carriers for the

apolipoprotein E epsilon4 allele. *Neurobiology of Learning and Memory, 129,* 113-119. doi:10.1016/j.nlm.2015.04.011

Green, A. J., Cervantez, M., **Graves, L. V.**, Morgan, C. D., & Murphy, C. (2013). Age and Apolipoprotein E ε4 effects on neural correlates of odor memory. *Behavioral Neuroscience, 127*(3), 339-349. doi:10.1037/a0031891

MANUSCRIPTS UNDER REVIEW

Holden, H. M., Tierney, S. M., **Graves, L. V.**, Woods, S. P., Herndon, A., Delis, D. C., Corey-Bloom, J., & Gilbert, P. E. (in revision). Subtle verbal memory deficits are detectable in premanifest Huntington's disease using the California Verbal Learning Test-II.

Williams, M. E., **Graves, L. V.,** Delano-Wood, L., Bondi, M. W., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (in revision). The emergence of age-related changes in recognition memory in healthy middle-aged adults using the CVLT-II.

Van Etten, E. J., **Graves, L. V.**, Taylor, B. P., Holden, H. M., Lopez, F. V., Pirogovsky-Turk, E., Corey-Bloom, J., Filoteo, J. V., & Gilbert, P. E. (under review). Recall and recognition discriminability in Parkinson's disease and Huntington's disease.

MANUSCRIPTS IN PREPARATION

Graves, L. V., Drozdick, L., Holdnack, J., Courville, T., Farrer, T. F., Gilbert, P. E., & Delis, D. C. (in preparation). Cohort differences between the CVLT3 and CVLT-II: Differential positive and negative Flynn effects across age groups.

Van Etten, E. J., Sumida, C., Holden, H. M., **Graves, L. V.**, & Gilbert, P. E. (in preparation). Who, when, and where? Evidence for changes in middle age on a new test of episodic memory.

PRESENTATIONS

Graves, L. V., Holden, H. M., Van Etten, E. J., Delano-Wood, L., Bondi, M. W., Corey-Bloom, J., Salmon, D. P., Gilbert, P. E., & Delis, D. C. (2019, February). *New intrusion analyses on the CVLT-3: Clinical utility in distinguishing patients with Alzheimer's and Huntington's disease.* Poster presented at the 47th Annual Meeting of the International Neuropsychological Society, New York City, NY.

Simone, S., **Graves, L. V.**, Williams, M. E., Mattson, S. N., Delano-Wood, L., Bondi, M. W., Salmon, D. P., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2019, February). *Total recognition discriminability in Huntington's and Alzheimer's disease: New insights from the CVLT-3.* Poster presented at the 47th Annual Meeting of the International Neuropsychological Society, New York City, NY.

Williams, M. E., **Graves, L. V.**, Delano-Wood, L., Bondi, M. W., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2019, February). *An exploration of memory performance in amnestic mild cognitive impairment and premanifest Huntington's*

disease using the CVLT-II. Poster presented at the 47th Annual Meeting of the International Neuropsychological Society, New York City, NY.

Williams, M. E., **Graves, L. V.**, Van Etten, E. J., Holden, H. M., Delano-Wood, L., Bondi, M. W., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2018, June). *The emergence of age-related changes in recognition memory in healthy middle-aged adults using the CVLT-II.* Poster presented at the 16th Annual Meeting of the American Academy of Clinical Neuropsychology, San Diego, CA.

Graves, L. V., Holden, H. M., Van Etten, E. J., Delano-Wood, L., Bondi, M. W., Salmon, D. P., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2018, February). *Novel recognition discriminability in Alzheimer's and Huntington's disease: New insights from the CVLT-3.* Poster presented at the 46th Annual Meeting of the International Neuropsychological Society, Washington, DC.

Van Etten, E. J., Sumida, C. A., **Graves, L. V.**, Holden, H. M., Lopez, F. V., & Gilbert, P. E. (2018, February). *Age-related differences on a new memory test for "Who, When, and Where" begin in middle age.* Poster presented at the 46th Annual Meeting of the International Neuropsychological Society, Washington, DC.

Endres, H. O., Holden, H. M., Tierney, S. M., **Graves, L. V.**, Corey-Bloom, J., & Gilbert, P. E. (2017, April). *Recall and recognition discriminability in premanifest Huntington's disease.* Poster presented at the 97th Annual Meeting of the Western Psychological Association, Sacramento, CA.

Olvera, S., Seewald, P. M., Moreno, C., Yandall, S., **Graves, L. V.**, Gilbert, P. E. (2017, April). *Age-related difference in temporal order memory for a self-generated list of everyday events.* Poster presented at the 97th Annual Meeting of the Western Psychological Association, Sacramento, CA.

Ursa, M., **Graves, L. V.,** Moreno, C., Seewald, P. M., Holden, H. M., Van Etten, E. J., Uttarwar, V., McDonald, C., Delano-Wood, L., & Gilbert, P. E. (2017, April). *Differential effects of age on recall and recognition discriminability in men and women.* Poster presented at the 97th Annual Meeting of the Western Psychological Association, Sacramento, CA.

Graves, L. V., Holden, H. M., Delis, D. C., & Gilbert, P. E. (2017, February). *Modifications to the CVLT-II novel recognition discriminability measure to enhance the detection of memory decline in normal aging.* Poster presented at the 45th Annual Meeting of the International Neuropsychological Society, New Orleans, LA.

Taylor, B. P., Van Etten, E. J., Holden, H. M., **Graves, L. V.**, Lopez, F. V., Nguyen, L., Pirogovsky-Turk, E., Corey-Bloom, J., Filoteo, J. V., & Gilbert, P. E. (2017, February). *CVLT-II performance in Huntington's disease and Parkinson's disease.* Poster presented at the 45th Annual Meeting of the International Neuropsychological Society, New Orleans, LA.

Van Etten, E. J., **Graves, L. V.**, Taylor, B. P., Holden, H. M., Lopez, F. V., Nguyen, L., Pirogovsky-Turk, E., Corey-Bloom, J., Filoteo, J. V., & Gilbert, P. E. (2017, February). *Recall and Recognition Discriminability in Parkinson's Disease and Huntington's Disease.* Poster presented at the 45th Annual Meeting of the International Neuropsychological Society, New Orleans, LA.

Van Etten, E., Sumida, C., Wagner G., Hileman J., Holden, H., **Graves, L. V.**, & Gilbert, P. E. (2016, April). *Age-related differences in short- and delayed-memory for sequences of face-place pairs.* Poster presented at the 96th Annual Meeting of the Western Psychological Society, Long Beach, CA.

DeJesus, S. N., DeFord, N. E., Holden, H. M., **Graves, L. V.**, Lopez, F. V., & Gilbert, P. E. (2016, April). *Examining the effects of interference on spatial recognition memory in young and older adults.* Poster presented at the 23rd Annual Meeting of the Cognitive Neuroscience Society, New York, NY.

Moreno, C., **Graves, L. V.**, Holden, H. M., Woods, S. P., Delano-Wood, L., Bondi, M., Salmon, D. P., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2016, February). *Recall and recognition discriminability in healthy aging.* Poster presented at the 44th Annual Meeting of the International Neuropsychological Society, Boston, MA.

Graves, L. V., Holden, H. M., Woods, S. P., Delano-Wood, L., Bondi, M., Corey-Bloom, J., Salmon, D. P., Delis, D. C., & Gilbert, P. E. (2016, February). *Recognition discriminability in Huntington's disease and Alzheimer's disease.* Poster presented at the 44th Annual Meeting of the International Neuropsychological Society, Boston, MA.

Moreno, C., **Graves, L. V.**, Holden, H. M., Tierney, S., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2015, April). *Recognition discriminability changes during premanifest stages of Huntington's disease: Evidence from the California Verbal Learning Test-II.* Poster presented at the 95th Western Psychological Association Annual Meeting, Las Vegas, NV.

Jacobson, A., Green, E., Oleson, S., **Graves, L. V.**, McIntosh, E., Buncic, A., Szajer, J., Gramling, L., Bitterlin, J. L., Robinson, J., & Murphy, C. (2014, May). *Body mass index affects brain activation to taste stimuli in young adults.* Poster presented at the 26th Annual Meeting of the Association for Psychological Sciences, San Francisco, CA.

Jacobson, A., Green, E., McIntosh, E., **Graves, L. V.**, Oleson, S., Anguiano, P., & Murphy, C. (2013, November). *Trait anxiety and its relationship to gustatory and reward processing.* Poster presented at the 43rd Annual Meeting of the Society for Neuroscience, San Diego, CA.

McIntosh, E., Jacobson, A., Green, E., Oleson, S., **Graves, L. V.**, & Murphy, C. (2013, November). *Associations between depressive symptoms and reward processing: Nutritive versus non-nutritive sweeteners.* Poster presented at the 43rd Annual Meeting of Society for Neuroscience, San Diego, CA.

Oleson, S., Green, E, Jacobson, A., McIntosh, E., **Graves, L. V.**, & Murphy, C. (2013, November). *Anterior cingulate activation is correlated with subjective pleasantness ratings for saccharin in individuals with metabolic syndrome.* Poster presented at the 43rd Annual Meeting of Society for Neuroscience, San Diego, CA.

Cervantez, M., **Graves, L. V.**, Green, A. J., Morgan, C. D., & Murphy, C. (2013, April). *Waist to hip ratios predict odor recognition memory processing speed in carriers of the Apolipoprotein E ε4 allele.* Poster presented at the 35th Annual Meeting of the Association for Chemoreception Sciences, Huntington Beach, CA.

Graves, L. V., Cervantez, M., Green, A. J., Morgan, C. D., & Murphy, C. (2013, April). *OERP scalp topography as a function of age and Apolipoprotein E ε4 during encoding of olfactory information.* Poster presented at the 20th Annual Meeting of the Cognitive Neuroscience Society, San Francisco, CA.

Szajer, J., **Graves, L. V.**, & Murphy, C. (2013, April). *Sensory modality influences episodic metamemory accuracy in healthy aging and Alzheimer's disease***.** Poster presented at the 35th Annual Meeting of the Association for Chemoreception Sciences, Huntington Beach, CA.

Graves, L. V., Cervantez, M., Green, A. J., Morgan, C. D., & Murphy, C. (2013, February). *The effects of age and the Apolipoprotein E ε4 allele on OERP scalp topography in odor recognition memory.* Poster presented at the 41st Annual Meeting of the International Neuropsychological Society, Waikoloa, HI.

Cervantez, M., **Graves, L. V.**, Green, A. J., Morgan, C. D., & Murphy, C. (2012, November). *Differences in P2 latency as a function of age, response type, and electrode site.* Poster presented at the 65th Annual Meeting of the Gerontological Society of America, San Diego, CA.

Graves, L. V., Cervantez, M., Green, A. J., Morgan, C. D., & Murphy, C. (2012, November). *Differences in OERP activity during retrieval using odors versus retrieval using odor labels as cues in an odor memory task.* Poster presented at the 65th Annual Meeting of the Gerontological Society of America, San Diego, CA.

Cervantez, M., **Graves, L. V.**, Green, A. J., Morgan, C. D., & Murphy, C. (2012, April). *Differences in olfactory event-related potentials (OERPs) among Apolipoprotein ε4 positive and negative individuals in an odor-memory retrieval task.* Poster presented at the 34th Annual Meeting of the Association for Chemoreception Sciences, Huntington Beach, CA.

Graves, L. V., Cervantez, M., Green, A. J., Morgan, C. D., & Murphy, C. (2012, April). *OERP differences as a function of age, the Apolipoprotein E (ApoE) ε4 allele, electrode site, and response type during retrieval using odor labels in an odor-recognition memory task.* Poster presented at the 34th Annual Meeting of the Association for Chemoreception Sciences, Huntington Beach, CA.

CLINICAL EXPERIENCE

APA-Accredited Predoctoral Internship in Clinical Psychology
VA Palo Alto Health Care System | Clinical Neuropsychology Track | 2018 – 2019

Memory Clinic and Interdisciplinary Cognitive Assessment Clinic

Polytrauma Rehabilitation Center/Comprehensive Rehabilitation Center

Polytrauma Transitional Rehabilitation Center

Trauma Recovery Services – Women's Trauma Recovery Program

Neuropsychological Assessment Experience

UC San Diego Memory, Aging, and Resilience Center | UC San Diego | 2018

Delis Neuropsychological Center | Encinitas, CA | 2017 – 2018

Neuropsychological Assessment Clinic in Oncology | UC San Diego | 2016 – 2018

Neuropsychological Assessment Unit | VASDHS | 2016 – 2017

UC San Diego Neuropsychology Laboratory | UC San Diego | 2016 – 2017

Psychodiagnostic Assessment and Cognitive-Behavioral Intervention Experience

Military Sexual Trauma & Interpersonal Trauma Clinic | VASDHS | 2015 – 2016

Dialectical Behavior Therapy Program | VASDHS | 2015 – 2016

SDSU Psychology Clinic | SDSU | 2014 – 2015

Other Clinical Experience

At-Risk Support Group Facilitator | Huntington's Disease Society of America | 2017 – 2018

PROFESSIONAL AND ACADEMIC SERVICE

Invited Presentations

Acute and Post-Acute Rehabilitation for Severe Traumatic Brain Injury | VAPAHCS | 2019

Emotion Regulation Skills Training in Post-Acute Rehabilitation | VAPAHCS | 2019

Assessment of Mild Cognitive Impairment and Dementia | VAPAHCS | 2019

Dialectical Behavior Therapy | SDSU | 2018

Dialectical Behavior Therapy | SDSU | 2017

Dialectical Behavior Therapy | UCSD | 2017

Dialectical Behavior Therapy for Borderline Personality Disorder | SDSU | 2016

Somatoform and Dissociative Disorders | SDSU | 2012

Graduate Teaching Assistantships

Statistical Methods in Psychology | SDSU | 2012 – 2013

Abnormal Psychology | SDSU | 2012

Drugs and Behavior | SDSU | 2011

Psychology of Personality | SDSU | 2011

Committees

Student Representative for SDSU/UC San Diego JDP in Clinical Psychology | UC San Diego School of Medicine Psychiatry Education & Training Committee | 2017 – 2018

Student Representative | SDSU/UC San Diego JDP in Clinical Psychology Selection Committee | 2017

Student Representative for SDSU/UC San Diego JDP in Clinical Psychology| Association of Neuropsychology Students & Trainees | APA Division 40 | 2016 – 2018

Student Representative | SDSU/UC San Diego JDP in Clinical Psychology Student Council | 2015 – 2018

SELECTED SEMINARS AND WORKSHOPS

Introduction to R | Instructor: Jonathan Helm, Ph.D. | SDSU | 2018

Linear Mixed-Effects Models in R | Instructor: Michael Thomas, Ph.D. | UC San Diego | 2018

PROFESSIONAL MEMBERSHIPS

APA Division 40: Society for Clinical Neuropsychology | 2014 – 2019

International Neuropsychological Society | 2016 – 2019

Cognitive Neuroscience Society | 2012

Gerontological Society of America | 2012

Association for Chemoreception Sciences | 2011

ABSTRACT OF THE DISSERTATION

Modifications to the CVLT-II Recognition Discriminability Indices to Enhance the Characterization of Recognition Memory Impairment in Healthy Aging and Neurodegenerative Disease

by

Lisa V. Graves

Doctor of Philosophy in Clinical Psychology

University of California San Diego, 2019
San Diego State University, 2019

Professor Paul E. Gilbert, Chair

Rationale: Although studies using the California Verbal Learning Test (CVLT) have shown that recognition memory declines with age, age-related differences on more nuanced aspects of recognition have not been explored in detail. CVLT research also has

shown that Alzheimer's disease (AD) is associated with impaired recall and recognition that reflect an encoding deficit stemming from neuropathology in medial temporal cortices. Huntington's disease (HD) is also associated with impaired recall but, in contrast to AD, in early stages shows less compromised recognition, reflecting primarily a retrieval deficit due to frontal-system dysfunction. Although original CVLT research demonstrated that recognition deficits are less severe in HD than in AD, more recent CVLT-II evidence has yielded mixed findings, highlighting the need for more refined measures to elucidate the extent of recognition memory impairment in AD and HD and thereby enhance characterizations of the profiles of memory loss associated with these neurodegenerative conditions. **Method and Results:** *Study 1* examined whether AD and HD differences on the CVLT-II Total Recognition Discriminability (RD) index (which assesses the ability to distinguish List A [original word list] targets from all distractors on the Yes/No Recognition trial) varied across nonparametric (used in CVLT) and parametric (used in CVLT-II) calculations of Total RD. Comparisons of group differences on nonparametric versus parametric Total RD indices could only be done in the context of raw scores (the CVLT-II does not by default include a nonparametric Total RD index, and no standardized nonparametric Total RD scores were therefore available), and the magnitudes of group differences were not shown to significantly differ across nonparametric and parametric methods. However, in contrast to what was observed in prior CVLT studies, analyses indicated that relative to the AD group, the HD group exhibited comparable standardized parametric Total RD scores (despite higher raw nonparametric and parametric Total RD scores [after accounting for demographic factors when appropriate]). A potential explanation for the discrepancy in performance of

individuals with HD on the CVLT versus CVLT-II recognition trials was that an increased proportion of prototypical distractors (i.e., those that are semantically related to targets) on the CVLT-II relative to the CVLT disproportionately amplified the difficulty of the recognition trial for individuals with frontal-system dysfunction (e.g., HD). *Study 2* investigated the utility of more refined RD indices generated to parse the degree of semantic association between targets and distractors (on Source and Novel RD indices that assess the ability to distinguish List A targets from List B [interference word list] and novel distractors, respectively) in characterizing nuanced aspects of yes/no recognition memory in healthy older and young adults. Although older adults performed worse than young adults on all RD indices, age group differences were smaller on more refined RD indices that excluded semantically-related distractors in RD calculations. Building on findings from Studies 1 and 2, *Study 3* examined performance in individuals with AD and HD on Total RD and List A vs. Novel/Unrelated RD indices. The latter is a new index that isolates the ability to distinguish List A targets from distractors that are novel and semantically unrelated to targets. The List A vs. Novel/Unrelated RD index is therefore less subject than the Total RD index to source memory difficulties and semantic confusion, both of which are often seen in individuals with frontal-system dysfunction (e.g., HD), and may yield more purified assessments of yes/no recognition memory in HD. Analyses indicated that the List A vs. Novel/Unrelated RD index yielded more robust AD versus HD differences than the Total RD index, providing greater differentiation between individuals whose memory disorder is primarily at the encoding/storage level (e.g., AD) versus at the retrieval level (e.g., early HD).

Relevance: Given the expanding older population and that memory loss is a hallmark

feature of cognitive impairment in both healthy aging and neurodegenerative disease, more refined memory indices are needed to elucidate the nature of memory changes that may be associated with normal aging and neurodegenerative processes. Collectively, Studies 1, 2, and 3 utilized innovative psychometric approaches with the CVLT-II to explore age-related differences on more nuanced aspects of yes/no recognition memory, as well as elucidate the extent of yes/no recognition memory impairment in AD and HD.

CHAPTER 1:

Integrated Introduction

The human life expectancy in the United States has continuously increased over the last several decades. Currently, age is the greatest known risk factor for neurodegenerative disease. Thus, as individuals live longer, their risk of developing dementia due to neurodegenerative disease is expected to increase. Alzheimer's disease (AD) is the most common form of dementia, accounting for 60-80% of all cases (Alzheimer's Association, 2016). In 2016, an estimated $236 billion was spent in the U.S. on care for individuals with AD. Costs are expected to rise to over $1.1 trillion by 2050 unless fruitful efforts to identify and treat those with, or at risk for, AD are made. Huntington's disease (HD), another form of dementia, is a debilitating neurodegenerative condition associated with a triad of motor, cognitive, and psychiatric symptoms. Although information available on the direct costs of HD is more limited, research has shown that annual costs of care increase with disease progression (Divino et al., 2013). Given that memory loss is a hallmark feature of cognitive decline in both healthy aging and dementia, more refined measures of memory are needed to further elucidate the nature of memory changes that may accompany normal aging and neurodegenerative disease.

The California Verbal Learning Test (CVLT) is a standardized neuropsychological measure that provides a multitude of verbal learning and memory indices (Delis, Kramer, Kaplan, & Ober, 1987, 2000, 2017). The original and second editions of the CVLT have been widely used in research and clinical settings, to characterize memory function and decline in healthy aging as well as in various

neurodegenerative populations, including AD and HD. Recall, and to a lesser extent recognition memory, have been shown decline with age (Craik & McDowd, 1987; Danckert & Craik, 2013). Additionally, research has shown that the different patterns of neurodegeneration associated with AD and HD yield distinct profiles of memory loss (Delis et al., 1991; Hodges, Salmon, & Butters, 1990; Moss, Albert, Butters, & Payne, 1986; Salmon & Bondi, 2009; Salmon & Filoteo, 2007; Troster et al., 1993). AD is associated with medial temporal lobe damage, particularly in the hippocampal formation, subsequent damage to cortical association areas, and relative sparing of most subcortical structures (Braak & Braak, 1991; Hyman, Van Hoesen, Damasio, & Barnes, 1984). Individuals with AD usually have pervasive memory deficits characterized by poor learning, rapid forgetting, and poor recognition (Budson & Kowall, 2013), a profile of memory loss thought to reflect an encoding/storage deficit. In contrast, HD is associated with early damage to basal ganglia structures (Vonsattel, 2000; Vonsattel et al., 1985) that have extensive projections to the frontal lobes (Alexander, Crutcher, & DeLong, 1990; Crosson et al., 2003; Cummings, 1993), followed by more diffuse involvement of other cortical and subcortical regions and networks. Patients with early stage HD often have significant deficits in recall with less compromised recognition (Butters, Wolfe, Granholm, & Martone, 1986; Butters, Wolfe, Martone, Granholm, & Cermak, 1985; Lundervold, Reinvang, & Lundervold, 1994; Martone, Butters, Payne, Becker, & Sax, 1984; Massman, Delis, Butters, Levin, & Salmon, 1990), a profile of memory loss thought to primarily reflect a retrieval deficit. Although recognition is less impaired than recall in early HD, recognition is often significantly impaired in the later stages of

disease, raising the possibility that encoding processes are also compromised to at least some degree (see Montoya et al., 2006 for review).

Rationale for Study 1

In general, studies using the original CVLT have shown that deficits in verbal learning and recall are less severe in HD than in AD (Delis et al., 1991; Kramer et al., 1988; Kramer, Levin, Brandt, & Delis, 1989). Additionally, individuals with HD (particularly those in milder stages of the disease) have been shown to perform better than those with AD on the CVLT index of Recognition Discriminability (RD), which captures the ability to distinguish List A targets from all distractors on the Yes/No Recognition trial (Delis et al., 1991; Kramer et al., 1988). In particular, evidence suggests that better RD performance in HD relative to AD corresponds with lower false positive (FP) error rates in the former versus the latter (Kramer et al., 1988).

Only one study prior to Study 1 of this staple dissertation project directly compared HD and AD performance on the CVLT-II Yes/No Recognition trial (Fine et al., 2008). The study found that individuals with HD performed significantly better than those with AD on CVLT-II measures of Total RD and Novel RD (which captures the ability to distinguish List A targets from novel distractors), but comparably on Source RD (which captures the ability to distinguish List A targets from List B distractors). Thus, studies using the CVLT and CVLT-II that were conducted prior to Study 1 collectively demonstrated that Total RD is less impaired in HD than in AD.

The original and second editions of the CVLT utilized nonparametric and parametric formulas, respectively, to compute a Total Recognition Discriminability (Total RD) index, which assesses the ability to distinguish List A target items from all

distractor types on the Yes/No Recognition trial. Although CVLT and CVLT-II studies conducted prior to Study 1 demonstrated that Total RD performance is less impaired in HD than in AD (Delis et al., 1991; Fine et al., 2008; Kramer et al., 1988; Kramer, Levin, Brandt, & Delis, 1989), the degree to which AD and HD differences on Total RD may vary across applications of nonparametric and parametric methods for calculating Total RD had not been explored. Insight into whether such variation occurs would potentially inform efforts to interpret and compare CVLT and CVLT-II findings regarding yes/no recognition memory in HD and AD. Additionally, insight into whether the nonparametric and parametric formulas for Total RD differ in the extent to which they capture FP error rates would assist in elucidating the nature of yes/no recognition memory impairment in these populations. Moreover, limited evidence regarding HD performance on the CVLT-II relative to a demographically similar comparison group was available prior to Study 1, and existing studies were based on relatively small samples. Accordingly, *Study 1* of this staple dissertation project examined whether AD and HD differences on Total RD varied across applications of nonparametric (used in original CVLT) and parametric (used in CVLT-II) formulas for calculating Total RD in a relatively large sample of individuals with HD, individuals with AD, and two demographically similar healthy comparison groups (Graves et al., 2017).

Rationale for Study 2

Although not affected to the same degree as recall, recognition memory has been shown to decline with age (Craik & McDowd, 1987; Danckert & Craik, 2013). Accordingly, older adults have been shown to exhibit worse performance relative to young adults on CVLT indices of recall, and to a lesser extent, recognition (Delis et al.,

1987; Ebert & Anderson, 2009; Kausler, 1994; Turner & Pinkston, 1993; Van der Linden, Philippot, & Heinen, 1997; Woodruff-Pak & Finkbiner, 1995). Moreover, studies have shown that the effect of aging is greater on source memory than on item memory (Bayer et al. 2011; Dennis et al., 2008; Glisky & Kong, 2008; Hashtroudi, Johnson, & Chrosniak, 1989; McIntyre & Craik, 1987; Naveh-Benjamin & Craik, 1995; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; Spaniol, Madden, & Voss, 2006; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999). In addition to Total RD, the CVLT-II provides an index of Source RD, which assesses the ability to distinguish List A target items from List B distractor items on the CVLT-II Yes/No Recognition trial. Thus, the Source RD index, although not a direct measure of source memory per se, taps into aspects of source memory by measuring one's ability to distinguish whether a word was included on List A or List B. The CVLT-II also includes an index of Novel RD, which assesses the ability to distinguish List A targets from novel distractors (i.e., items that were not previously presented during test administration prior to the Yes/No Recognition trial). Thus, the Novel RD index represents recognition memory in a more traditional sense, providing a measure of one's ability to distinguish "old" stimuli (i.e., target items) from "new" stimuli (i.e., novel distractor items). Half of the List B distractors as well as novel distractors are prototypical, or semantically related to targets, rendering them perhaps more challenging to identify as distractors than the other half of the List B and novel distractors, which are semantically unrelated to targets (e.g., Baldo et al., 2002). Thus, more refined indices that isolate the ability to distinguish List A targets from semantically-unrelated List B and novel distractors may yield purer assessments of Source and Novel RD, respectively.

Study 2 of this staple dissertation project sought to examine between- and within-group differences on the four false positive (FP) error subtypes found on the CVLT-II Yes/No Recognition trial (prototypical List B, unrelated List B, prototypical novel, and unrelated novel) in healthy older and young adults. Additionally, FP error rates were used to generate and compare *d'* scores between healthy older and young adults on multiple variations of Source and Novel RD that differed in the degree of semantic association between targets and distractors. It was generally hypothesized that more refined Source and Novel RD indices, that minimized levels of semantic interference, would yield smaller age group differences on Yes/No Recognition testing, thereby elucidating the nature or degree of age-related differences on more nuanced aspects of yes/no recognition memory.

Rationale for Study 3

Studies using the original CVLT consistently demonstrated that deficits on the Yes/No Recognition trial are less severe in mildly demented individuals with HD than in equally demented individuals with AD (Delis et al., 1991; Kramer et al., 1988; Kramer, Levin, Brandt, & Delis, 1989). This difference was shown with the original CVLT Total RD index that measures the ability to distinguish List A targets from all distractors on the Yes/No Recognition trial (Delis et al., 1991; Kramer et al., 1988). In contrast, studies that compared individuals with AD and those with HD on the CVLT-II Yes/No Recognition trial have produced inconsistent results. While one study found that individuals with HD obtained higher standardized scores than those with AD on the CVLT-II Total RD index (Fine et al., 2008), another study with a larger sample found that AD and HD groups performed comparably on this measure (Study 1; Graves et al., 2017). One implication of

this pattern of results is that individuals with HD may have worse Yes/No Recognition performance on the CVLT-II than on the original CVLT. This possibility is consistent with the clinical observation that Total RD scores of patients with HD are generally lower on the CVLT-II than on the original CVLT (Dean C. Delis, personal communication, October 26, 2017).

Reasons for differences in the performance of individuals with HD on the Yes/No Recognition trials of the two versions of the CVLT may lie in differences in how Total RD is calculated. The Yes/No Recognition trial on the original CVLT included only half (i.e., eight) of the 16 List B items as distractors (Delis et al., 1987). Due to a ceiling effect in cognitively normal individuals, the difficulty of the trial was increased on the CVLT-II by including all 16 List B items as distractors (Dean C. Delis, personal communication, September 26, 2017). Although this had the intended effect of making the Yes/No Recognition trial more difficult, it potentially made the test more sensitive to deficits in source memory. Individuals with frontal-system dysfunction (e.g., HD) may have particular difficulty in identifying the source of each previously-presented item (List A or List B) during the Yes/No Recognition trial when asked whether or not an item had been on List A (Fine et al., 2008), and increasing the number of List B distractors on the CVLT-II Yes/No Recognition trial may have amplified this difficulty.

The CVLT-II Yes/No Recognition trial also had an increased proportion of distractors that are semantically related to List A target items (eight of 28 distractors for the CVLT versus 16 of 32 distractors for the CVLT-II). Research has shown that patients with frontal-system dysfunction are prone to making semantic intrusion or semantic confusion errors due to impaired inhibition of activation within semantic networks (e.g.,

Baldo et al., 2002). A deficit in inhibition of the semantic network during the CVLT may lead individuals with HD to have greater difficulty in rejecting distractors that share obvious semantic associations with targets than in rejecting distractors that do not (the same deficit could lead to semantically-related intrusion errors during free recall trials). This would have a greater adverse effect on the CVLT-II than the CVLT for individuals with HD due to the increased proportion of semantically-related distractors. Increasing the proportion of semantically-related distractors may not have the same effect on individuals with AD since their severe recognition memory deficits reflect a profound encoding/storage deficit that can be attributed to more extensive neuropathology targeting the medial temporal lobes and cortical association areas. Thus, individuals with AD are likely to exhibit relatively comparable levels of difficulty in rejecting novel distractors whether or not they share obvious semantic associations with targets.

While the CVLT-II included eight novel unrelated distractor items on the Yes/No Recognition trial, it did not provide a separate index that assessed the ability of individuals to endorse List A targets while rejecting those novel unrelated distractors. The second and third editions of the CVLT (CVLT-II and CVLT-3, respectively) contain the same target words on the recall trials and the same targets and distractors on the Yes/No Recognition trial (in fact, the only word-item changes that were made to the CVLT-3 are on the Forced Choice Recognition trial). However, the CVLT-3 (Delis, Kramer, Kaplan, & Ober, 2017) includes a more refined RD index, List A vs. Novel/Unrelated RD, that isolates the ability to distinguish List A targets from distractors that were not previously presented during test administration and do not share obvious semantic associations with targets. Thus, the new List A vs. Novel/Unrelated RD index

minimizes any potential influences of source and semantic interference, and is therefore thought to provide a purer assessment of yes/no recognition memory.

Study 3 of this staple dissertation project sought to elucidate the nature and extent of AD and HD differences in yes/no recognition memory by examining performance in individuals with AD and HD on CVLT-3 indices of Total RD and List A vs. Novel/Unrelated RD. It was hypothesized that although both AD and HD would be associated with deficits on Yes/No Recognition testing, individuals with HD would perform better than those with AD, particularly on the new List A vs. Novel/Unrelated RD index. In other words, the List A vs. Novel/Unrelated RD index, in minimizing any potential influences of source and semantic interference, was expected to yield a purer assessment of yes/no recognition memory in HD and thereby exhibit greater differentiation than the Total RD index between individuals with HD and those with AD.

CHAPTER 2:

Study 1

The content within this section, titled "Chapter 2: Study 1," reflects material from a paper that has been published in the *Journal of Clinical and Experimental Neuropsychology*. The proper citation is as follows:

Graves, L. V., Holden, H. M., Delano-Wood, L., Bondi, M. W., Woods, S. P., Corey-Bloom, J., Salmon, D. P., Delis, D. C., & Gilbert, P. E. (2017). Total recognition discriminability in Huntington's and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 39*(2), 120-130. doi:10.1080/13803395.2016.1204993

Abstract

Both the original and second editions of the California Verbal Learning Test (CVLT) provide an index of total recognition discriminability (TRD) but respectively utilize nonparametric and parametric formulas to compute the index. However, the degree to which population differences in TRD may vary across applications of these nonparametric and parametric formulas has not been explored. We evaluated individuals with Huntington's disease (HD), individuals with Alzheimer's disease (AD), healthy middle-aged adults, and healthy older adults who were administered the CVLT-II. Yes/no recognition memory indices were generated, including raw nonparametric TRD scores (as used in CVLT-I) and raw and standardized parametric TRD scores (as used in CVLT-II), as well as false positive (FP) rates. Overall, the patient groups had significantly lower TRD scores than their comparison groups. The application of nonparametric and parametric formulas resulted in comparable effect sizes for all group comparisons on raw TRD scores. Relative to the HD group, the AD group showed comparable standardized parametric TRD scores (despite lower raw nonparametric and parametric TRD scores), whereas the previous CVLT literature has shown that standardized TRD scores are lower in AD than in HD. Possible explanations for the similarity in standardized parametric TRD scores in the HD and AD groups in the present study are discussed, with an emphasis on the importance of evaluating TRD scores in the context of other indices such as FP rates in an effort to fully capture recognition memory function using the CVLT-II.

Keywords: California Verbal Learning Test, California Verbal Learning Test – Second Edition, recognition discriminability, Huntington's disease, Alzheimer's disease

Total Recognition Discriminability in Huntington's and Alzheimer's Disease

The California Verbal Learning Test (CVLT; Delis, Kramer, Kaplan, & Ober, 1987, 2000) is a standardized neuropsychological test that provides a multitude of verbal learning and memory indices. The original and second editions of the CVLT (CVLT-I and CVLT-II, respectively) are widely used in research and clinical settings and have been utilized in efforts to characterize memory function and decline in various populations.

Total Recognition Discriminability (TRD) on the CVLT

Both versions of the CVLT include a TRD index, which is a single score that reflects the ability of the examinee to endorse target items *and* reject distractor items. The CVLT-I used the following nonparametric formula (see Underwood, 1974) to compute a TRD index that was interpreted as a percentage, with 100 (%) set as the maximum possible score:

$$TRD_{CVLT\text{-}I,nonp} = [1 - (\text{total FPs} + \text{total misses}) / 44] \times 100.$$

Thus, the nonparametric TRD index incorporates an examinee's total number of FPs into a ratio or percentage TRD score. In addition to differentiating patient and control populations, this index has been useful in distinguishing patients with different profiles of memory loss, particularly those with primarily cortical (e.g., AD) versus subcortical (e.g., HD) degeneration (see Delis et al., 2000 for review).

The CVLT-I was developed between 1979 to 1981, prior to the availability of personal computers. As was discussed in the CVLT-II manual (Delis et al., 2000), the nonparametric measure of TRD was employed in the CVLT-I because it allowed for a relatively quick, convenient calculation of recognition discriminability by hand and still

correlated strongly with more complex parametric signal-detection measures such as *d'*. However, as noted by Corwin (1994), the nonparametric formula was less able than other measures to account for response bias on recognition memory tasks with an unequal number of target and distractor items, such as the yes/no recognition memory task on the CVLT-II. Alternatively, *d'* is calculated independently of response bias, rendering it better suited for tests with an unequal number of target and distractor items. By the time the CVLT-II was developed, personal computers were widely available and facilitated the efficient application of more complex mathematical methods for assessing recognition memory function, further strengthening the rationale for employing the parametric *d'* measure to compute the TRD index on the CVLT-II.

A raw *d'* score reflects the absolute difference in standard deviation units between an examinee's hit rate and FP rate and is therefore analogous to a contrast *z* score (Delis et al., 2000; Macmillan & Creelman, 1991):

$$TRD_{CVLT\text{-}II,d'} = z(\text{hit rate}) - z(\text{FP rate}).$$

Thus, in contrast to the nonparametric TRD score, which more generally reflects an examinee's percentage of correct responses, the parametric TRD score more specifically reflects an examinee's hit rate relative to their FP rate. In this regard, the parametric formula for TRD might better capture recognition memory function in cases where there are unequal numbers of target and distractor items. However, whether the parametric formula for TRD sufficiently captures the full magnitude of FP errors, particularly in individuals who are susceptible to committing very high FP rates (e.g., those with AD) is not entirely clear and warrants further consideration. Additional distractor items were included on the CVLT-II relative to the CVLT-I to increase test difficulty and lower the

ceiling effect that is often found on recognition memory tests, including that on the CVLT-I (Dean C. Delis, personal communication, December 5, 2015). Specifically, the CVLT-II includes all 16 List B distractor items, whereas the CVLT-I included only 8 List B distractor items. Thus, the proportion of distractor items that are from List B is larger on the CVLT-II (16/32) than on the CVLT-I (8/28). In shifting from the use of a nonparametric TRD formula in the CVLT-I to the use of a parametric TRD formula in the CVLT-II, it is unclear whether an attempt to accommodate the imbalance between the number of target and distractor items on the yes/no recognition memory test comes at the cost of not fully capturing the true possible range of FP rates that may occur in certain neurodegenerative populations.

HD, AD, and Profiles of Memory Loss

HD is a neurodegenerative disorder caused by expanded repetitions of the cytosine-adenine-guanine (CAG) trinucleotide on the huntingtin gene located on the short arm of chromosome 4 (Huntington's Disease Collaborative Research Group, 1993). HD is characterized by an array of motor, cognitive, and psychiatric changes. Motor changes include chorea in addition to bradykinesia, rigidity, and ataxia (Ross et al., 2014). Cognitive deficits associated with HD include impaired episodic memory, executive functioning, attention, and visuospatial processing (Dumas, van den Bogaard, Middelkoop, & Roos, 2013). On the other hand, AD is characterized by early deficits in episodic memory followed by later decline in other cognitive domains including language, executive functioning, and visuospatial processing (Salmon & Bondi, 2009). In contrast to HD, motor functioning is relatively preserved in the context of AD.

The classic profile of episodic memory loss in AD is thought to be one of poor encoding and retention of information, which lead to rapid forgetting and result in impaired recall and recognition (Budson & Kowall, 2013; Dickerson & Atri, 2014; Salmon & Bondi, 2009). Conversely, earlier evidence suggests that individuals with HD exhibit what has been referred to as a "subcortical profile" of episodic memory decline that includes poor recall and improvements in recognition. This profile is thought to reflect impaired retrieval processes but relatively intact encoding and maintenance mechanisms and is particularly evident in earlier stages of the disease (Butters, Wolfe, Martone, Granholm, & Cermak, 1985; Butters, Delis, & Lucas, 1995; Delis et al., 1991). More recent evidence suggests that recognition memory is indeed compromised in HD, but to a lesser extent than recall (see Montoya et al., 2006 for review). Indeed, the extant literature strongly suggests that, although individuals with HD and AD have both been shown to exhibit recall deficits, recognition memory is thought to be less impaired in HD, at least in the earlier stages of the disease.

Studies using the CVLT-I and CVLT-II have generally shown that individuals with HD and AD are impaired in various aspects of verbal learning and memory, and that they differ from other populations (both healthy and impaired) and from each other in their profiles of memory loss (see Elwood, 1995 for review of CVLT-I literature).

HD and AD performance on the CVLT. Research suggests that individuals with HD perform worse than a demographically similar comparison group on multiple measures of verbal learning and recall on the CVLT-I (Kramer et al., 1988; Massman, Delis, Butters, Levin, & Salmon, 1990; Massman, Delis, Butters, Dupont, & Gillin, 1992; Massman, Delis, & Butters, 1993). Some studies have shown that those with HD

demonstrate worse recognition discriminability (Kramer et al., 1988; Lundervold, Reinvang, & Lundervold, 1994; Massman et al., 1990), and other evidence suggests that they exhibit greater improvement on recognition discriminability relative to Trial 5 recall (Massman et al., 1992). Various studies also have indicated that individuals with AD perform worse than healthy older adults on multiple indices of verbal learning and recall on the CVLT-I (Delis et al., 1991; Deweer et al., 1994; Kramer et al., 1988; Massman et al., 1993; Mendez & Ashla-Mendez, 1991; Simon, Leach, Winocur, & Moscovitch, 1994). Additionally, those with AD have been shown to demonstrate higher FP rates or a positive response bias (Delis et al., 1991; Deweer et al., 1994; Kramer et al., 1988), worse recognition discriminability (Delis et al., 1991; Deweer et al., 1994; Kramer et al., 1988), and no improvement on recognition testing relative to free recall (Delis et al., 1991).

No prior studies to the authors' knowledge have examined CVLT-II performance in individuals with HD relative to a demographically similar comparison group (i.e., cognitively healthy middle-aged adults). One study investigated the effects of rivastigmine treatment on cognitive function in early stage HD and included a group of healthy controls to account for practice effects between baseline and follow-up assessments (Sesok, Bolle, Kobal, Bucik, & Vodusek, 2014). However, no direct comparisons between patients and controls in performance on neuropsychological measures (including the CVLT-II) were made. Consistent with CVLT-I findings, studies using the CVLT-II have shown that individuals with AD perform worse than healthy older adults on most measures of recall (Delis et al., 2005; Duarte et al., 2006; Sherod et al., 2009) as well as recognition discriminability (Duarte et al., 2006).

HD/AD comparisons on the CVLT. In general, studies using the CVLT-I have

shown that deficits in verbal learning and recall are less severe in individuals with HD relative to those with AD (Delis et al., 1991; Kramer et al., 1988; Kramer, Levin, Brandt, & Delis, 1989). Additionally, evidence suggests that individuals with HD – particularly in milder stages of the disease – demonstrate better recognition discriminability than those with AD on the CVLT-I (Delis et al., 1991; Kramer et al., 1988) that is reflected by lower FP rates in the absence of differences in hit rates (Kramer et al., 1988). Also, compared to individuals with AD, those with HD have shown greater improvement on recognition discriminability relative to Trial 5 recall (Delis et al., 1991). Some evidence suggests that individuals with mild HD have a smaller positive response bias than those with AD on the CVLT-I (Kramer et al., 1988), whereas other findings have not found such group differences (Delis et al., 1991).

Only two studies have examined the relative performance of individuals with HD and AD on the CVLT-II. The first study revealed that whereas those with HD and AD do not differ on immediate and delayed recall measures when using the traditional measure of target recall, those with AD perform significantly worse than those with HD on short-delay free recall, short-delay cued recall, and long-delay cued recall when using a new index called "recall discriminability" that analyzes target recall relative to intrusion rate (Delis et al., 2005). The second study revealed that those with AD perform significantly worse than those with HD on CVLT-II measures of total and novel recognition discriminability, but comparably on measures of source recognition discriminability (Fine et al., 2008). These studies were based on the same sample of 16 individuals with HD and 17 individuals with AD.

Although the two versions of the CVLT utilize different formulas to compute the TRD index, no prior study has directly assessed the degree to which population differences in TRD vary across applications of these nonparametric and parametric methods. Insight into whether or not such variation occurs would inform efforts to interpret and compare CVLT-I and CVLT-II findings regarding recognition memory function in HD and AD in particular. Additionally, insight into whether the nonparametric and parametric formulas for TRD differ in the extent to which they capture FP errors would be helpful in improving efforts to accurately characterize recognition memory function in these populations in research and clinical settings. Moreover, limited evidence exists regarding HD performance relative to a demographically similar comparison group on the CVLT-II, and what is available is based on relatively small samples. Therefore, the purpose of the present study was to compare nonparametric and parametric assessments of TRD using the CVLT-II in a relatively large sample of individuals with HD and AD and two demographically similar comparison groups.

Method

Participants

Participants included 66 individuals with HD, 33 individuals with AD, 68 healthy middle-aged adults (comparison sample for the HD group), and 35 healthy older adults (comparison sample for the AD group). Individuals with HD were recruited from the Huntington's Disease Clinical Research Program (HDCRP) at the University of California, San Diego (UCSD), which follows a cohort of individuals with HD who have participated in longitudinal clinical studies and undergone annual evaluations of cognitive

and motor symptoms (a portion of the HD group came from the same sample used in the studies by Delis et al., 2005 and Fine et al., 2008). The HD group was administered the Unified Huntington's Disease Rating Scale (UHDRS; Huntington Study Group, 1996) by a senior staff neurologist at the HDCRP. Individuals with HD were diagnosed with definite HD on the basis of unequivocal motor signs on the UHDRS and a positive family history of HD. In addition, all HD participants had a CAG repeat length greater than 39, indicating that all carried the fully penetrant genetic mutation for HD. Exclusionary criteria for individuals with HD and healthy middle-aged adults included the following: a diagnosis of any neurological disorder (with the exception of HD in the HD group), a diagnosis of any major medical condition (e.g., cancer), a diagnosis of any psychiatric disorder (with the exception of a mood or anxiety disorder in the HD group, for which any current symptoms were managed with medication), a history of traumatic brain injury, and a history of substance abuse. All participants provided informed written consent and this portion of the study was approved by the Institutional Review Boards of San Diego State University (SDSU) and UCSD.

CVLT-II data from the 33 individuals with AD and 35 healthy older adults were extracted from an archival database that included data from a larger battery of neuropsychological tests administered at the Shiley-Marcos Alzheimer's Disease Research Center (ADRC) in La Jolla or the Veterans Affairs San Diego Healthcare System (VASDHS) (a portion of the AD group came from the same sample used in the studies by Delis et al., 2005 and Fine et al., 2008). Participants at both sites were administered a standardized battery of tests by trained research assistants or psychometrists. Diagnoses of individuals with probable AD were consistent with the

criteria established by the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (ADRDA) workgroup (McKhann et al., 1984; McKhann et al., 2011).

Healthy middle-aged and older adults were recruited from the San Diego community by the HIV Neurobehavioral Research Center (HNRC) and the Bondi Laboratory at UCSD, respectively, using flyers (posted with approval by public sites/institutions) and outreach to senior centers. Efforts were made to target healthy populations with demographic characteristics similar to those of the patient groups.

Measures

Dementia Rating Scale-2 (DRS-2). Individuals with HD and AD completed the DRS-2 (Jurica, Leitten, & Mattis, 2001), a measure of global cognitive functioning, as part of a larger neuropsychological battery.

CVLT-II and TRD indices. The CVLT-II was administered as part of a larger neuropsychological battery to all participants using standard administration procedures outlined by Delis and colleagues (Delis et al., 2000). CVLT-II data were collected between May 2002 and July 2013. The CVLT-II involves the presentation of word-lists and provides a multitude of verbal learning and memory indices, including immediate recall, free and cued recall over short and long delays, and recognition memory. The TRD indices that were of primary interest to the present study were derived from the yes/no recognition memory portion of the CVLT-II. In the present study, short- and long-delay tests of recall were separated by an interval of approximately 20 minutes, during which other nonverbal neuropsychological measures were administered. CVLT-II data were scored using CVLT-II scoring software (Delis & Fridlund, 2000).

Nonparametric and parametric TRD scores were calculated using CVLT-II data. Raw nonparametric TRD scores were computed using the following formula:

$$TRD_{CVLT\text{-}II,nonp} = [1 - (\text{total FPs} + \text{total misses}) / 48] \times 100.$$

Note that the CVLT-II contains 48 total items in the yes/no recognition memory test, whereas the CVLT-I contained 44 items. Raw and standardized parametric TRD scores were computed by CVLT-II software. Raw nonparametric and parametric TRD scores in addition to standardized parametric TRD scores were analyzed. An analysis of standardized nonparametric TRD scores could not be conducted, as the normative data that would be required to do so have not been published or made available otherwise.

Statistical Analyses

Analyses were conducted in Statistical Package for the Social Sciences (SPSS) Version 22. Prior to examining group differences in TRD scores, chi-square analyses were conducted to determine whether groups differed in gender. In addition, analysis of variance (ANOVA) tests were conducted to determine whether groups differed in age, education, or DRS-2 scores.

Shapiro-Wilk tests of normality revealed that TRD scores were non-normally distributed in the present sample ($p < .05$). Therefore, Mann-Whitney *U* tests were conducted to examine differences in raw nonparametric and parametric TRD scores in the comparison of the HD group to healthy middle-aged adults and in the comparison of the AD group to healthy older adults. For the HD and AD group comparisons on raw nonparametric and parametric TRD scores, nonparametric and parametric TRD scores were ranked, and two-way ANOVA tests were conducted to examine group differences in the ranked TRD scores, while including gender as a second between-subjects factor

(Akritas, 1990; Baguley, 2012). Mann-Whitney *U* tests were conducted to examine group differences in standardized parametric TRD scores in all comparisons of interest.

Effect size values for each group comparison in each analysis of TRD scores were calculated. Following Mann-Whitney *U* tests, *r* values for the group effect were computed by dividing the *Z* value associated with the *U* statistic by the square root of N (Fritz, Morris, & Richler, 2012). Following two-way ANOVA tests on ranked data, *r* values for the group effect were computed manually using sum of squares (SS) error terms from the SPSS output and the following formula: $r_{effect} = \sqrt{(SS_{effect} / SS_{total})}$. All *r* values were converted to Cohen's *d* effect size estimates using the following formula: $d = 2r/\sqrt{(1-r^2)}$. Fisher's *r* to *z* transformation analyses were conducted to determine whether group differences in TRD scores significantly differed between applications of nonparametric and parametric methods. Spearman rank correlation analyses and Fisher's *r* to *z* transformation analyses also were conducted in a set of exploratory analyses involving FP rates, TRD scores, and source recognition discriminability (SoRD; the endorsement of List A target items and rejection of List B distractor items) scores.

Although a comparison of standardized nonparametric and parametric TRD scores could not be made, standardized parametric TRD scores were still analyzed to provide researchers and clinicians with the opportunity to make relevant inferences with the present data, as the norms for standardized TRD scores are available in the CVLT-II manual and are widely utilized in research and clinical settings.

Results

Demographic information for the HD and AD groups and their respective comparison groups is provided in Table 2.1. A chi-square analysis revealed that there

were no differences in the percentage of men versus women between the HD group and healthy middle-aged adults or between the AD group and healthy older adults (*p*s > .05). However, compared to the HD group, the AD group had a significantly greater percentage of men relative to women, χ^2 (N = 91) = 9.06, p < .01. Thus, subsequent analyses examining differences in raw nonparametric and parametric TRD scores between the HD and AD groups included gender as a between-subjects factor.

A one-way ANOVA test revealed that there were no differences in age between the HD group and healthy middle-aged adults or between the AD group and healthy older adults (*p*s > .05). However, individuals with AD were significantly older than those with HD, $t(89) = 14.15$, $p < .001$, which was expected given known differences in the age of disease onset. This highlights that age is systematically confounded with group in the comparison of HD and AD, which renders including age as a covariate in subsequent analyses examining differences in raw TRD scores between the HD and AD groups a statistically invalid method for parceling out age effects on raw TRD scores. This issue is inherently present in studies involving the comparison of raw scores between groups of individuals with HD and AD. Accordingly, age was not included as a factor in ANOVA tests examining differences in raw TRD scores between the HD and AD groups in the present study. Moreover, the size of the present sample would not accommodate alternative analyses that might otherwise address this issue.

A one-way ANOVA test revealed no differences in education between the HD and AD groups; the HD group and healthy middle-aged adults; or the AD group and healthy older adults (*p*s > .05). The HD and AD groups did not differ in mean DRS-2 scores (p > .05), suggesting that the groups were comparable in terms of overall cognitive

impairment. Moreover, the variation in DRS-2 scores within each of the patient groups was minimal.

As shown in Table 2.2, Mann-Whitney *U* tests revealed that the HD and AD groups performed significantly worse than their respective comparison groups on all TRD indices. According to Fisher's *r* to *z* transformation analyses, the effect sizes associated with the nonparametric and parametric formulas were comparable in all comparisons of patient groups to their respective comparison groups (*p*s > .05; see Table 2.3).

As shown in Table 2.2, ANOVA tests on ranked data revealed that the AD group performed significantly worse than the HD group on *raw* nonparametric and parametric TRD indices. In the comparison of the HD and AD groups using raw scores, the difference in effect sizes associated with the nonparametric and parametric formulas was negligible and was not statistically significant according to a Fisher's *r* to *z* transformation analysis ($p > .05$; see Table 2.3). However, a Mann-Whitney *U* test revealed that the AD group performed comparably to the HD group on *standardized* parametric TRD, despite the observation that the AD group had significantly lower *raw* parametric TRD scores than the HD group. Table 2.2 displays the inferential and descriptive statistics for all planned group comparisons on TRD scores.

Exploratory analyses were conducted in an attempt to elucidate the observation that the AD group had standardized parametric TRD scores that were comparable to those in the HD group, despite lower raw nonparametric and parametric TRD scores. Specifically, additional analyses involving FP rates, TRD scores, and SoRD scores were conducted.

Mann-Whitney *U* tests were conducted to examine differences between the HD and AD groups in raw and standardized FP rates. The Mann-Whitney *U* tests revealed that the AD group (Mdn_{FPraw} = 14.00; Mdn_{FPz} = 2.50) had significantly higher raw, U = 339.50, $p < .001$, and standardized FP rates, U = 611.50, $p < .01$, than the HD group (Mdn_{FPraw} = 5.00; Mdn_{FPz} = 1.00). Spearman rank correlation analyses then were conducted to examine correlations between raw FP rates and raw nonparametric and parametric TRD scores in the HD and AD groups. The analyses revealed significant negative correlations between raw FP rates and raw nonparametric and parametric TRD scores in the HD ($r_{s\_FP(nonp)}$ = -.73, $p < .001$; $r_{s\_FP(d')}$ = -.57, $p < .001$) and AD ($r_{s\_FP(nonp)}$ = -.91, $p < .001$; $r_{s\_FP(d')}$ = -.64, $p < .001$) groups. Moreover, Fisher's *r* to *z* transformation analyses revealed that the correlation between raw FP rates and raw nonparametric TRD scores was significantly larger than the correlation between raw FP rates and raw parametric TRD scores in the AD group (z = 2.99, $p < .01$) but not the HD group (z = 1.46, p = .14). However, it is important to note that because FP errors are incorporated in the calculation of the TRD index, the reported correlations between FP rates and TRD scores may be influenced by a certain degree of circularity.

Additionally, exploratory analyses were conducted to examine differences between the HD and AD groups in SoRD scores. A Mann-Whitney *U* test revealed that the AD group (Mdn = 0.80) had significantly lower raw SoRD scores than the HD group (Mdn = 1.75), U = 370.50, $p < .001$. However, the HD (Mdn = -1.50) and AD (Mdn = -2.00) groups were comparable on standardized SoRD scores, U = 874.50, p = .49.

Discussion

In the present study, nonparametric and parametric formulas were applied in the assessment of TRD using the CVLT-II in a relatively large sample of individuals with HD and AD and healthy adults. As expected, the HD and AD groups performed worse than their respective comparison groups on nonparametric (raw) and parametric (raw and standardized) indices of TRD. It is worth noting that the effect size for the comparison of AD and healthy older adults on standardized parametric TRD scores was larger than the effect size for the HD and healthy middle-aged adults comparison. However, this difference is consistent with empirical evidence suggesting that individuals with HD exhibit rather heterogeneous cognitive abilities, with memory deficits that are typically less severe than those observed in individuals with AD.

Relative to the HD group, the AD group had comparable standardized parametric TRD scores despite lower raw nonparametric and parametric TRD scores (even after adjusting for gender, which is corrected for in the standardization of scores on the CVLT-II). The examination of raw scores in research may be informative, yet clinical judgments about the nature of recognition memory abilities and dysfunction by default rely on the analysis and interpretation of standardized scores, as did previous efforts to characterize and distinguish profiles of memory loss in neurodegenerative populations using the CVLT. Interpreting standardized TRD scores from the present study in isolation would lead to the conclusion that individuals with AD show comparable deficits in TRD relative to those with HD. This is in contrast with findings from previous studies in which the CVLT-I (which employs the nonparametric formula) was used to assess TRD. These studies showed that individuals with AD exhibited worse recognition discriminability than those with HD (Delis et al., 1991; Kramer et al., 1988). Additionally, Fine et al.

(2008) used the CVLT-II to assess TRD in individuals with HD and AD and also found that those with AD performed worse than those with HD based on standardized scores, albeit using a smaller sample than the present study sample.

A possible explanation for the similarity in standardized parametric TRD scores in the HD and AD groups in the present study, despite earlier evidence for individuals with AD performing worse than those with HD on standardized TRD scores using the CVLT, may involve the extent to which the nonparametric and parametric formulas for TRD capture high FP rates. Namely, the unequal number of target and distractor items on the CVLT-II (although better accounted for by the use of the parametric *d'* formula to calculate the TRD index) may be an important factor to consider when assessing TRD in individuals with a tendency to commit high FP rates, such as individuals with AD. Consistent with previous findings (Kramer et al., 1988), the AD group in the present study committed significantly more FP errors than the HD group. In addition, exploratory analyses revealed significant negative correlations between raw FP rates and raw nonparametric and parametric TRD scores in both the HD and AD groups. Moreover, Fisher's *r* to *z* transformation analyses revealed that the correlation between raw FP rates and raw nonparametric TRD scores was significantly larger than the correlation between raw FP rates and raw parametric TRD scores in the AD group but not the HD group. These observations highlight that the nonparametric TRD formula may more fully capture the contribution of FP errors to a TRD score and, as a result, provide important information regarding an examinee's recognition memory function that may otherwise be lost in the application of the parametric TRD formula and the standardization of parametric TRD scores. The present data and findings suggest that TRD scores may be

somewhat overestimated (i.e., the impact of FP errors on TRD scores may be inadvertently reduced) in individuals with AD, leading them to appear to perform comparably to individuals with HD. However, as previously noted, it is important to note that because FP errors are incorporated in the calculation of the TRD index, the reported correlations between FP rates and TRD scores may be influenced by a certain degree of circularity. This point should be taken into consideration when interpreting the present findings.

Another potentially important factor when considering the similarity in standardized parametric TRD scores in the HD and AD groups in the present study involves the greater number of List B distractor items included in the yes/no recognition memory test on the CVLT-II relative to the CVLT-I. Research has shown that individuals with HD and other individuals with frontal system pathology are susceptible to source memory deficits (Baldo et al., 2002; Fine et al., 2008; Pirogovsky et al., 2007). On the CVLT, these deficits may manifest in the endorsement of List B distractor items on the yes/no recognition memory test in particular (see Fine et al., 2008). Consequently, it may be argued that individuals with HD (including those in the present study) are likely to exhibit lower TRD scores on the CVLT-II than they would on the CVLT-I given the opportunity to endorse more List B distractor items on the CVLT-II, which could potentially result in comparable TRD scores to individuals with AD. However, an exploratory analysis of SoRD scores (i.e., the ability to endorse List A target items and reject List B distractor items) in the present study revealed that although the AD group had significantly lower raw SoRD scores than the HD group, the groups were comparable on standardized SoRD scores, which is consistent with previous findings (Fine et al.,

2008). This suggests that the discrepancy between the present findings and those reported in previous CVLT studies regarding standardized parametric TRD scores in HD and AD is not likely due to differences in the number of List B distractor items included in the yes/no recognition memory test across the two versions of the CVLT.

Taken together, the primary and exploratory findings of the present study highlight the importance and utility of examining nonparametric TRD scores and other recognition memory indices (e.g., FP rates) in addition to (not instead of) standardized parametric TRD scores when using the CVLT-II to characterize recognition memory function.

Limitations

There are some limitations to the present study that deserve acknowledgement and discussion. First, age was systematically confounded with group in the comparison of HD and AD. The mean ages of the two patient groups differed by more than three standard deviations. This rendered including age as a covariate an insufficient method for parceling out the effects of age on raw TRD scores in subsequent analyses examining differences in raw TRD scores between the HD and AD groups. Age was therefore not included in ANOVA tests examining differences in raw TRD scores between the HD and AD groups. Moreover, the size of the present sample would not accommodate alternative analyses that might otherwise address this issue. Although this limits the interpretation of direct comparisons between the HD and AD groups, it reflects an inherent issue in studies involving the comparison of raw scores between individuals with HD and AD because of the known difference in the average age of disease onset. In light of the issue, it is important to emphasize that the HD and AD groups were equivalent in terms of overall

cognitive impairment based on DRS-2 scores, which are commonly used in neuropsychological studies to evaluate and compare the severity of disease in neurological populations. Second, CVLT-II data were collected between May 2002 and July 2013, and individuals were diagnosed with probable AD in alignment with the criteria established by the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (ADRDA) workgroup (McKhann et al., 1984; McKhann et al., 2011). More specifically, some individuals were diagnosed with AD using the criteria established in 1984, while other individuals were diagnosed with AD using the criteria that were updated in 2011. Although the general framework of probable AD dementia from the 1984 criteria were retained in the 2011 criteria, the 2011 criteria emphasize documenting cognitive decline in persons who meet the core clinical criteria for probable AD dementia to increase the certainty of diagnosis. We believe it is important to acknowledge that the present study sample consists of individuals with AD who were diagnosed using either the 1984 or 2011 criteria, and to encourage readers to take this into consideration in the evaluation of the present findings. Third, the size of the AD group in the present study was relatively small compared to the HD group. However, compared to previously published studies involving the CVLT-II, the AD group in the present study was substantially larger. Fourth, the study sample was relatively well educated and may not fully represent the population. However, it could be hypothesized that the observed magnitude of deficits in TRD in the patient groups actually may be increased in a sample of individuals with less cognitive reserve. Finally, we would like to acknowledge that although it would have been helpful to include an analysis and discussion of performance validity data derived

from the forced-choice recognition test, these data are not available for all participants in the study sample.

Conclusion

The present study found that, relative to individuals with HD, individuals with AD had comparable standardized parametric TRD scores despite lower raw nonparametric and parametric TRD scores, which is in contrast with what has been previously reported in the CVLT literature. A possible explanation for this difference in findings between the present and previous studies involves potential differences in the extent to which the nonparametric and parametric formulas for TRD capture high FP rates. A comprehensive approach to evaluating recognition memory function that includes the examination of other indices in addition to (not instead of) standardized TRD scores, which are relied upon by default for making clinical judgments about the nature of recognition memory abilities and dysfunction, is encouraged. The present findings may have important implications when making comparisons between CVLT-I and CVLT-II findings regarding TRD in HD and AD and in improving efforts to accurately characterize recognition memory function in these populations.

Acknowledgements

Dean C. Delis, Ph.D. is an author of the CVLT and receives royalties for the test. This work was supported in part by NIH grant R01 AG034202 to Paul E. Gilbert; NIH grant K24 AG026431 to Mark W. Bondi; the Shiley-Marcos ADRC grant AG05131 to David P. Salmon; and the UCSD Huntington's Disease Society of America (HDSA) Center of Excellence grant to Jody Corey-Bloom.

Dissertation Author's Acknowledgements

I, Lisa Graves (primary investigator and author of this material), would like to thank all coauthors – Drs. Gilbert, Holden, Delano-Wood, Bondi, Woods, Corey-Bloom, Salmon, and Delis – for their contributions to this work. I also would like to thank the *Journal of Clinical and Experimental Neuropsychology* for accepting this work for publication, and the Taylor & Francis Group for publishing this work.

Tables

Table 2.1. *Gender distribution and mean values (standard deviations) of age and education for individuals with Huntington's disease (HD), healthy middle-aged adults (MA), individuals with Alzheimer's disease (AD), and healthy older adults (OA).*

Variable	HD	MA	AD	OA
n	58	68	33	35
% Female	57%	47%	24%	46%
Age (years)	48.03 (9.58)	43.63 (15.56)	76.55 (8.60)	75.80 (8.82)
Education (years)	14.33 (2.15)	15.15 (2.17)	15.00 (2.69)	16.06 (1.86)
DRS-2 (total score)	123.78 (12.94)	N/A	119.55 (7.61)	N/A

Note: DRS-2 = measure of global cognitive functioning.

Table 2.2. *Inferential and descriptive statistics for all planned group comparisons on TRD indices.*

	Index	HD	MA	*Test Statistic*	*p*	*r*	*d*
HD vs MA	Nonp	76.90 (12.40)	94.67 (7.91)	$U = 386.50$	<.001	.70	1.94
	d'	1.71 (0.88)	3.40 (0.73)	$U = 328.00$	<.001	.72	2.08
	$z\ d'$	-1.78 (1.25)	0.44 (0.95)	$U = 354.50$	<.001	.71	2.02
	Index	AD	OA	*Test Statistic*	*p*	*r*	*d*
AD vs OA	Nonp	58.96 (12.24)	90.65 (6.50)	$U = 0.50$	<.001	.86	3.37
	d'	0.68 (0.61)	2.83 (0.68)	$U = 7.50$	<.001	.85	3.21
	$z\ d'$	-2.14 (0.89)	0.34 (0.79)	$U = 19.00$	<.001	.84	3.04
	Index	HD	AD	*Test Statistic*	*p*	*r*	*d*
HD vs AD	Nonp	76.90 (12.40)	58.96 (12.24)	$F(1,88) = 37.41$	<.001	.28	0.16
	d'	1.71 (0.88)	0.68 (0.61)	$F(1,88) = 34.27$	<.001	.27	0.15
	$z\ d'$	-1.78 (1.25)	-2.14 (0.89)	$U = 807.50$	.21	.13	0.26

Note: Nonp = raw nonparametric TRD score; d' = raw parametric TRD score; $z\ d'$ = standardized parametric TRD score.

Table 2.3. *Fisher's r to z transformations of group effects on raw nonparametric versus parametric TRD scores.*

		Nonp	*d'*		
Comparison	n	*r1*	*r2*	*z*	*p*
HD vs MA	126	.70	.72	0.32	.75
AD vs OA	68	.86	.85	0.21	.83
HD vs AD	91	.28	.27	0.07	.94

References

Akritas, M. G. (1990). The rank transform method in some two-factor designs. *Journal of the American Statistical Association, 85*(409), 73-78.

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences.* Basingstoke, United Kingdom: Palgrave Macmillan.

Baldo, J. V., Delis, D., Kramer, J., & Shimamura, A. P. (2002). Memory performance on the California Verbal Learning Test-II: Findings from patients with focal frontal lesions. *Journal of the International Neuropsychological Society, 8,* 539-546.

Budson, A. E., & Kowall, N. W. (2013). *Handbook of Alzheimer's disease and other dementias.* Hoboken, NJ: Wiley-Blackwell.

Butters, N., Wolfe, J., Martone, M., Granholm, E., & Cermak, L. S. (1985). Memory disorders associated with Huntington's disease: Verbal recall, verbal recognition, and procedural memory. *Neuropsychologia*, *23,* 729–743.

Butters, N., Delis, D. C., & Lucas, J. (1995). Clinical assessment of memory disorders in amnesia and dementia. *Annual Review of Psychology*, *46,* 493–523.

Corwin, J. (1994). On measuring discrimination and response bias: Unequal number of targets and distractors and two classes of distractors. *Neuropsychology*, *8*(1), 110–117.

Delis, D. C. & Fridlund, A. J. (2000). *CVLT-II Comprehensive Scoring System and Computerized Report.* San Antonio, TX: The Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test.* San Antonio, TX: The Psychological Corporation.

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

Delis, D. C., Massman, P. J., Butters, N., Salmon, D. P., Cermak, L. S., & Kramer, J. H. (1991). Profiles of demented and amnesic patients on the California Verbal Learning Test: Implications for the assessment of memory disorders. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 19–26.

Delis, D. C., Wetter, S. R., Jacobson, M. W., Peavy, G., Hamilton, J., Gongvatana, A., … & Salmon, D. P. (2005). Recall discriminability: Utility of a new CVLT-II measure in the differential diagnosis of dementia. *Journal of the International Neuropsychological Society, 11*(6), 708-715.

Deweer, B., Ergis, A. M., Fossati, P., Pillon, B., Boller, F., Agid, Y., & Dubois, B. (1994). Explicit memory, procedural learning and lexical priming in Alzheimer's disease. *Cortex, 30,* 113–126.

Dickerson, B., & Atri, A. (2014). *Dementia: Comprehensive principles and practices.* Oxford University Press.

Duarte, A., Hayasaka, S., Du, A., Schuff, N., Jahng, G.-H., Kramer, J., … & Weiner, M. (2006). Volumetric correlates of memory and executive function in normal elderly, mild cognitive impairment and Alzheimer's disease. *Neuroscience Letters, 406,* 60–65.

Dumas, E. M., van den Bogaard, S. J., Middelkoop, H. A., & Roos, R. A. (2013). A review of cognition in Huntington's disease. *Frontiers in Bioscience, 1,* 1–18.

Elwood, R. W. (1995). The California Verbal Learning Test: Psychometric characteristics and clinical application. *Neuropsychology Review, 5*(3), 173–200.

Fine, E. M., Delis, D. C., Wetter, S. R., Jacobson, M. W., Hamilton, J. M., Peavy, G., … & Salmon, D. P. (2008). Identifying the "source" of recognition memory deficits in patients with Huntington's disease or Alzheimer's disease: Evidence from the CVLT-II. *Journal of Clinical and Experimental Neuropsychology, 30*(4), 463–470.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141*(1), 2-18.

Huntington's Disease Collaborative Research Group. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell, 72,* 971–983.

Jurica, P. J., Leitten, S., Mattis, S. (2001). *Dementia Rating Scale-2: Professional manual.* Lutz, FL: Psychological Assessment Resources.

Kramer, J. H., Delis, D. C., Blusewicz, M. J., Brandt, J., Ober, B. A., & Strauss, M. (1988). Verbal memory errors in Alzheimer's and Huntington's dementias. *Developmental Neuropsychology, 4,* 1–15.

Kramer, J. H., Levin, B. E., Brandt, J., & Delis, D. C. (1989). Differentiation of Alzheimer's, Huntington's, and Parkinson's disease patients on the basis of verbal learning characteristics. *Neurology, 3,* 111–120.

Lundervold, A. J., Reinvang, I., & Lundervold, A. (1994). Characteristic patterns of verbal memory function in patients with Huntington's disease. *Scandinavian Journal of Psychology, 35,* 38–47.

Macmillan, N. A., & Creelman, D. C. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Massman, P. J., Delis, D. C., & Butters, N. (1993). Does impaired primacy recall equal impaired long-term storage? Serial position effects in Huntington's disease and Alzheimer's disease. *Developmental Neuropsychology, 9,* 1–15.

Massman, P. J., Delis, D. C., Butters, N., Levin, B. E., & Salmon, D. P. (1990). Are all subcortical dementias alike?: Verbal learning and memory in Parkinson's and Huntington's patients. *Journal of Clinical and Experimental Psychology, 12*(5), 729–744.

Massman, P. J., Delis, D. C., Butters, N., Dupont, R. M., & Gillin, J. C. (1992). The subcortical dysfunction hypothesis of memory deficits in depression: Neuropsychological validation in a subgroup of patients. *Journal of Clinical and Experimental Neuropsychology*, *14,* 687–706.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. Neurology, 34, 939–944.

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, Jr., C. R., Kawas, C. H., … & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging and the Alzheimer's Association workgroup. *Alzheimer's & Dementia, 7*(3), 263-269.

Mendez, M. F., & Ashla-Mendez, M. (1991). Differences between multi-infarct dementia and Alzheimer's disease on unstructured neuropsychological tasks. *Journal of Clinical and Experimental Neuropsychology, 13,* 923–932.

Montoya, A., Pelletier, M., Menear, M., Duplessis, E., Richer, F., & Lepage, M. (2006). Episodic memory impairment in Huntington's disease: A meta-analysis. *Neuropsychologia, 44,* 1984–1994.

Pirogovsky, E., Gilbert, P. E., Jacobson, M., Peavy, G., Wetter, S., Goldstein, G., … & Murphy, C. (2007). Impairments in source memory for olfactory and visual stimuli in preclinical and clinical stages of Huntington's disease. *Journal of Clinical and Experimental Neuropsychology, 29,* 395-404.

Ross, C. A., Aylward, E. H., Wild, E. J., Langbehn, D. R., Long, J. D., Warner, J. H., … & Tabrizi, S. J. (2014). Huntington disease: Natural history, biomarkers and prospects for therapeutics. *Nature Reviews Neurology, 10,* 204–216.

Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual Review of Psychology, 60,* 257–282.

Sesok, B., Bolle, N., Kobal, J., Bucik, V., & Vodusek, D. B. (2014). Cognitive function in early clinical phase Huntington disease after rivastigmine treatment. *Psychiatria Danubia, 26*(3), 239–248.

Sherod, M. G., Griffith, H. R., Copeland, J., Belue, K., Krzywanski, S., Zamrini, E. Y., … & Marson, D. C. (2009). Neurocognitive predictors of financial capacity across the dementia spectrum: Normal aging, mild cognitive impairment, and Alzheimer's disease. *Journal of the International Neuropsychological Society, 15*(2), 258–267.

Simon, E., Leach, L., Winocur, G., & Moscovitch, M. (1994). Intact primary memory in mild to moderate Alzheimer disease: Indices from the California Verbal Learning Test. *Journal of Clinical and Experimental Neuropsychology* 16, 414–422.

Underwood, B. J. (1974). The role of the association in recognition memory. *Journal of Experimental Psychology Monographs, 102*, 917–939.

CHAPTER 3:

Study 2

The content within this section, titled "Chapter 3: Study 2," reflects material from a paper that has been published in *Aging, Neuropsychology, and Cognition*. The proper citation is as follows:

Graves, L. V., Van Etten, E. J., Holden, H. M., Delano-Wood, L., Bondi, M. W., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2017). Refining CVLT-II recognition discriminability indices to enhance the characterization of recognition memory changes in healthy aging. *Aging, Neuropsychology, and Cognition, 25*(5), 767-782. doi:10.1080/13825585.2017.1372358

Abstract

The present study examined age-related changes in the four false positive (FP) error subtypes found on the California Verbal Learning Test—Second Edition (CVLT-II) yes/no recognition memory trial and the influence of the different subtypes on calculations of *d'* scores on indices of source and novel recognition discriminability (SoRD and NRD, respectively). Healthy young (n = 57) and older (n = 55) adults exhibited different patterns of FP errors; nonetheless, older adults generally made more FP errors than young adults. Accordingly, older adults also performed worse than young adults on all SoRD and NRD indices. Moreover, the magnitudes of observed differences between and within the two age groups on SoRD and NRD indices were shown to vary depending on the manner in which FP error subtypes were incorporated into calculations of *d'* scores. The present findings underline the importance of examining FP errors in the assessment of recognition memory abilities, and using refined indices of recognition discriminability to further elucidate the nature of age-related recognition memory impairment. Furthermore, the present findings highlight the potential for these refined indices to demonstrate clinical utility through enhancing characterizations of memory loss in more cognitively impaired populations.

Keywords: aging, source memory, item memory, recognition discriminability, California Verbal Learning Test

Refining CVLT-II Recognition Discriminability Indices to Enhance the Characterization of Recognition Memory Changes in Healthy Aging

The human life expectancy in the United States has continuously risen over the last several decades. Age is currently the greatest known risk factor for neurodegenerative disease. As the human life expectancy continues to rise, the burden of cognitive decline in older age and the prevalence of dementia due to neurodegenerative disease are expected to increase. The development and use of more refined assessments will be important for enhancing characterizations of the cognitive strengths and weaknesses associated with healthy aging.

Memory loss is one of the most common cognitive issues that arise in older age. Although memory loss, generally speaking, is associated with aging, evidence suggests that not all aspects of memory show an equal rate or magnitude of age-related decline. For example, several studies have shown that the effect of aging is greater on source memory than on item memory (Bayer et al. 2011; Dennis et al., 2008; Glisky & Kong, 2008; Hashtroudi, Johnson, & Chrosniak, 1989; McIntyre & Craik, 1987; Naveh-Benjamin & Craik, 1995; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; Spaniol, Madden, & Voss, 2006; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999). Source memory relates to the context from which information was learned or acquired, whereas item memory relates to content of such information regardless of its source. In other words, item memory refers to the ability to remember *what* happened, whereas source memory refers to the ability to remember *where, when, and how* it happened (Dennis et al., 2008). It has been suggested that impaired encoding of contextual information accounts for poorer performance among older adults on source memory tasks (Johnson,

Hashtroudi, & Lindsay, 1993). Specifically, age-related dysfunction may result in the inability to engage mnemonic processes for integrating contextual information with item memory during encoding. Moreover, older adults may possess only enough cognitive resources to encode the stimulus itself, at the expense of also encoding contextual information (i.e., are stimulus bound), resulting in poorer recall of such contextual (or source) information (Glisky & Kong, 2008; Johnson et al., 1993).

Memory for the context (i.e., source) and content of an episodic event may rely on different brain regions. Neuroimaging studies and studies involving patients with focal brain lesions have shown that source memory may rely on the functional integrity of both the frontal and temporal lobes (Awipi & Davachi, 2008; Cansino, Maquet, Dolan, & Rugg, 2002; Ekstrom & Bookheimer, 2007; Janowsky, Shimamura, & Squire, 1989; Kirwan, Wixted, & Squire, 2008; Mitchell, Raye, Johnson, & Greene, 2006; Peters, Koch, Schwarz, & Daum, 2007; Peters, Suchan, Koster, & Daum, 2007). Accordingly, age-related pathology of the frontal and temporal regions may account for the source memory decline that is often observed in normal aging (Dennis et al., 2008; Fan, Snodgrass, & Bilder, 2003; Glisky & Kong, 2008; Glisky, Rubin, & Davidson, 2001; Henkel, Johnson, & De Leonardis, 1998; Mitchell et al., 2006). Other studies have indicated that the frontal lobes, in particular, are strongly implicated in source memory (Craik, Morris, Morris, & Loewen, 1990; Fan et al., 2003; Glisky, Polster, & Routhieaux, 1995; Glisky et al., 2001; Janowsky et al., 1989; Schacter, Harbluk, & McLachlan, 1984), whereas the medial temporal lobes may be more involved in item memory (Shimamura & Squire, 1987; Stark & Squire, 2000, 2003).

Recognition memory is a component of declarative memory that involves the ability to recognize previously encountered stimuli. Although not affected to the same degree as recall, recognition memory has been shown to decline with age (Craik & McDowd, 1987; Danckert & Craik, 2013). The original and second editions of the California Verbal Learning Test (CVLT-I and CVLT-II, respectively; Delis, Kramer, Kaplan, & Ober, 1987, 2000) are widely used in research and clinical settings and have been utilized in efforts to characterize memory function and decline in healthy aging. In general, older adults have been shown to exhibit worse performances relative to young adults on indices of recall, and, to a lesser extent, recognition (Delis et al., 1987; Ebert & Anderson, 2009; Kausler, 1994; Turner & Pinkston, 1993; Van der Linden, Philippot, & Heinen, 1997; Woodruff-Pak & Finkbiner, 1995).

Further exploration of more nuanced aspects of recognition memory function may provide additional valuable insight into the cognitive changes that accompany healthy aging. Since the mid to late twentieth century, signal detection theory (SDT) has been applied in studies of recognition memory as a gold standard for assessing recognition memory function that takes sensitivity and response bias into account. Delis and colleagues included a recognition discriminability index on the CVLT-I (Delis et al., 1987), and introduced additional subtypes of recognition discriminability on the CVLT-II (Delis et al., 2000). These CVLT-II indices are calculated using *d'* (Macmillan & Creelman, 1991). In addition to total recognition discriminability, the CVLT-II provides an index of source recognition discriminability (SoRD), which captures the ability to distinguish List A target items from List B distractor items on the CVLT-II yes/no recognition memory trial. Thus, the SoRD index, although not a direct measure of source

memory per se, taps into aspects of source memory by measuring one's ability to distinguish whether a word was included on List A or List B. Half of the List B distractor items are prototypical, or semantically related to target items, rendering them perhaps more challenging to identify as distractors than the other half of the List B distractor items, which are semantically unrelated (e.g., Baddeley, 1966). Thus, a SoRD index that excludes contributions from FP errors related to prototypical distractors and therefore more specifically captures the ability to distinguish List A target items from List B distractor items that are not semantically related to target items may yield a more refined assessment of SoRD.

The CVLT-II also includes an index of novel recognition discriminability (NRD), which captures the ability to distinguish List A target items from novel (i.e., non-List B) distractor items. Thus, the NRD index represents recognition memory in a more traditional sense, providing a measure of one's ability to distinguish "old" stimuli (i.e., target items) from "new" stimuli (i.e., novel distractor items). Half of the novel distractor items are prototypical, or semantically related to target items, rendering them more challenging to identify as distractors than the other half of the novel distractor items, which are semantically unrelated. Thus, a NRD index that excludes contributions from FP errors related to prototypical distractors and therefore more specifically captures the ability to distinguish List A target words from novel distractor items that are semantically unrelated to target items may provide a more refined assessment of NRD.

In addition to SoRD and NRD, the CVLT-II provides a third subtype of recognition discriminability called semantic recognition discriminability (SeRD) that captures the ability to distinguish List A target items from distractor items that are

semantically related to target items, including those that are from List B as well as those that are novel (Delis et al., 2000). In contrast, the SoRD and NRD indices reflect the ability to distinguish targets from List B and novel distractors, respectively, without parsing the contributions of prototypical, or semantically related distractors from those of semantically unrelated distractors.

On that premise, the extent to which between- and within-group differences in SoRD and NRD performances may be influenced by the degree of semantic association between targets and distractors found on the CVLT-II yes/no recognition memory trial has not been explored. Thus, the present study has two main objectives. First, between- and within-group differences in FP errors in each of the four subtypes that are found on the CVLT-II yes/no recognition memory trial (prototypical List B, unrelated List B, prototypical novel, and unrelated novel) will be examined in healthy older (n = 55) and young (n = 57) adults. Second, between- and within-group differences in *d'* scores were examined on three variations of the SoRD and NRD indices: 1) original SoRD and NRD (which include both prototypical and semantically unrelated List B and novel distractors in *d'* calculations), 2) SoRD-prototypical and NRD-prototypical (which include prototypical List B and novel distractors only in *d'* calculations), and 3) SoRD-unrelated and NRD-unrelated (which include semantically unrelated List B and novel distractors only in *d'* calculations). Older adults are expected to make more FP errors than young adults, although it is hypothesized that the two age groups may exhibit different patterns of FP errors across the four subtypes. Additionally, older adults are expected to perform worse than young adults on all SoRD and NRD indices, albeit to a lesser extent on indices that exclude prototypical distractors. In particular, the SoRD-unrelated and NRD-

unrelated indices are expected to be associated with smaller group differences than the SoRD and NRD indices. Furthermore, both age groups are expected to demonstrate better performances on SoRD-unrelated and NRD-unrelated indices than on SoRD and NRD indices, and older adults are generally expected to exhibit worse performances on SoRD than on NRD. Findings from the present study may help to elucidate the nature of recognition memory function and changes in healthy aging with the use of more refined measures of recognition discriminability.

Method

Participants

Study participants included 57 healthy young adults (18-25 years of age) and 55 healthy older adults (65 years of age or older). Older adults were characterized as cognitively healthy based on Dementia Rating Scale-2 (DRS-2; Jurica, Leitten, & Mattis, 2001) scores (130 or above). Exclusionary criteria for all healthy adult participants included the following: a diagnosis of any neurological disorder, a diagnosis of any major medical condition (e.g., cancer), a diagnosis of any psychiatric disorder (with the exception of a mood disorder, for which any current symptoms must be well managed), a history of traumatic brain injury, and a history of substance abuse. All participants provided informed written consent and the study was approved by the Institutional Review Boards of San Diego State University (SDSU) and/or the University of California, San Diego (UCSD).

Healthy young adults were recruited from the San Diego community by the Center for Healthy Aging and Neurodegenerative Disease Research (CHANDR) at SDSU and the Huntington's Disease Clinical Research Program (HDCRP) at UCSD. Healthy

older adults were recruited from the San Diego community by CHANDR at SDSU, the Normal Aging Laboratory at UCSD, and the HDCRP at UCSD. Participants were administered a standardized battery of neuropsychological tests by trained research assistants or psychometrists. CVLT-II data from the subset of healthy older adults recruited by the Normal Aging Laboratory were extracted from an archival database that included data from a larger battery of neuropsychological tests administered at the Shiley-Marcos Alzheimer's Disease Research Center (ADRC) in La Jolla and the Veterans Affairs San Diego Healthcare System (VASDHS).

CVLT-II and Recognition Discriminability (RD) Indices

The CVLT-II was administered using standard procedures outlined by Delis and colleagues (2000). The CVLT-II is a list-learning test that provides a multitude of verbal learning and memory indices, including immediate recall, free and cued recall over short and long delays, and recognition memory. The RD indices that were of primary interest in the present study were generated using variables derived from the yes/no recognition memory trial on the CVLT-II. Short- and long-delay tests of recall were separated by an interval of approximately 20 minutes, during which other nonverbal neuropsychological measures were administered. CVLT-II data were scored using CVLT-II scoring software (Delis & Fridlund, 2000). Raw scores on hits, the four FP error subtypes [prototypical List B (used in calculating the SoRD-prototypical index), unrelated List B (used in calculating the SoRD-unrelated index), prototypical novel (used in calculating the NRD-prototypical index), and unrelated novel (used in calculating the NRD-unrelated index)], and the six RD indices (SoRD, SoRD-prototypical, SoRD-unrelated, NRD, NRD-prototypical, and NRD-unrelated) were examined.

SoRD and NRD indices are calculated using the following formulas (Delis et al., 2000):

1. SoRD (*d'*) = *z*(hits) – *z*(FP errors associated with prototypical List B distractors + FP errors associated with semantically unrelated List B distractors)
2. NRD (*d'*) = *z*(hits) – *z*(FP errors associated with prototypical novel distractors + FP errors associated with semantically unrelated novel distractors)

SoRD-prototypical, SoRD-unrelated, NRD-prototypical, and NRD-unrelated indices were generated using the following formulas:

3. SoRD-prototypical (*d'*) = *z*(hits) – *z*(FP errors associated with prototypical List B distractor items only)
4. SoRD-unrelated (*d'*) = *z*(hits) – *z*(FP errors associated with unrelated List B distractor items only)
5. NRD-prototypical (*d'*) = *z*(hits) – *z*(FP errors associated with prototypical novel distractor items only)
6. NRD-unrelated (*d'*) = *z*(hits) – *z*(FP errors associated with unrelated novel distractor items only)

Raw *d'* scores are computed by calculating inverse proportions of hits and respective FP errors and subtracting respective FP error rates from hit rates (see Macmillan & Creelman, 1991).

Statistical Analyses

Analyses were conducted in the Statistical Package for the Social Sciences (SPSS) Version 24. Prior to examining age group differences in hits, the four FP error

subtypes, and the six RD indices of interest, chi-square analyses and one-way analysis of variance (ANOVA) tests were conducted to determine whether gender and education, respectively, were significant predictors of the outcome variables (hits, FP errors, RD indices). Gender and education were not significant predictors of the particular outcome variables of interest in the present study and therefore were not controlled for in the primary analyses.

Shapiro-Wilk tests of normality revealed that all outcome variables were non-normally distributed (ps < .05). Thus, nonparametric analyses were conducted to address the aims of the present study.

Analysis of hits. A Mann-Whitney *U* test was conducted to examine the age group difference in the number of hits on the CVLT-II yes/no recognition memory trial.

Analyses of FP error subtypes. Due to the substantial number of zero FP errors across subtypes and individuals, separate chi-square analyses were conducted to examine differences between the two age groups in the number of individuals who made zero FP errors versus one or more FP errors in each of the four subtypes. Additionally, separate chi-square tests of independence were conducted to make pairwise comparisons of the four FP error subtypes within each age group. Effect size values (r) were calculated to quantify and compare the magnitudes of significant between- and within-group differences in FP error subtypes [$r = \sqrt{(\chi^2/\mathrm{N})}$].

Analyses of RD indices. Six separate Mann-Whitney *U* tests were conducted to examine age group differences in *d'* scores on the six RD indices of interest (SoRD, SoRD-prototypical, SoRD-unrelated, NRD, NRD-prototypical, and NRD-unrelated). Additionally, two separate Friedman tests were conducted to examine the effect of RD

index type on *d'* scores within each group. If a significant omnibus effect of RD index type was observed, nine follow-up Wilcoxon signed-rank tests were conducted to make the following pairwise comparisons within a particular age group: 1) SoRD vs. NRD, 2) SoRD-prototypical vs. NRD-prototypical, 3) SoRD-unrelated vs. NRD-unrelated, 4) SoRD vs. SoRD-prototypical, 5) SoRD vs. SoRD-unrelated, 6) SoRD-prototypical vs. SoRD-unrelated, 7) NRD vs. NRD-prototypical, 8) NRD vs. NRD-unrelated, and 9) NRD-prototypical vs. NRD-unrelated. Effect size values (*r*) were calculated to quantify and compare the magnitudes of significant between- and within-group differences in *d'* scores on RD indices ($r = Z/\sqrt{N}$).

False discovery rate adjustment. Adjustments for a false discovery rate (FDR) of .05 (see Benjamini & Hochberg, 1995) were applied in the analyses of between- and within-group differences on FP error subtypes and RD indices. Original *p* values are presented in the study tables, and asterisks indicate which *p* values retained significance following FDR adjustments.

Results

Demographic Information

A chi-square analysis revealed no difference between the older (49.09% women) and young (57.89% women) adult groups in their proportions of men and women, χ^2 (2, $N = 112$) = 0.87, $p = .45$. A one-way ANOVA revealed that older adults ($M = 16.36$, $SD = 2.08$) completed more years of education than young adults ($M = 14.28$, $SD = 2.21$), $F(1, 110) = 31.64$, $p < .001$. All older adults had DRS-2 scores of 130 or higher ($M = 140.62$, $SD = 2.96$).

Analysis of Hits

A Mann-Whitney *U* test revealed that older adults (*mean rank* = 50.15, *sum of ranks* = 2758.50) had significantly fewer hits on the CVLT-II yes/no recognition memory trial than young adults (*mean rank* = 62.62, *sum of ranks* = 3569.50), $U = 1218.50$, $p < .05$.

Analyses of FP Error Subtypes

Age group differences in FP error subtypes. Descriptive and inferential statistics for age group differences in the number of individuals who made zero versus one or more FP errors in each of the four subtypes are provided in Table 3.1. Proportions of older and young adults who made one or more FP errors in each of the four subtypes are illustrated in Figure 3.1. Chi-square analyses revealed that the extent to which the number of individuals who made zero FP errors was higher than the number of individuals who made one or more FP errors was smaller in the older adult group than in the young adult group in three of the four FP error subtypes: prototypical List B, unrelated List B, and prototypical novel (i.e., the proportion of individuals who made one or more FP errors was larger in the older adult group than in the young adult group in the three aforementioned subtypes). No age group difference in the extent to which the number of individuals who made zero FP errors was higher than the number of individuals who made one or more FP errors was observed in the unrelated novel subtype.

Within-group differences in FP error subtypes. Inferential statistics for comparisons within each age group in the number of individuals who made zero versus one or more FP errors across subtypes are provided in Table 3.2. Proportions of individuals who made one or more FP errors across subtypes within each age group are

illustrated in Figure 3.2. Chi-square analyses revealed different patterns of FP errors within the older adult, χ^2 (3, N = 55) = 30.48, p <.001, and young adult, χ^2 (3, N = 57) = 14.15, p <.01, groups. In the older adult group, the proportion of individuals who made one or more FP errors was 1) greater for the prototypical List B and prototypical novel subtypes than the unrelated List B and unrelated novel subtypes, respectively, 2) greater for the unrelated List B subtype than the unrelated novel subtype, 3) greater for the prototypical novel subtype than the unrelated List B subtype, and 4) greater for the prototypical List B subtype than the unrelated novel subtype. In the young adult group, the proportion of individuals who made one or more FP errors was greater for the prototypical novel subtype than the unrelated novel subtype; however, no other comparisons within the young adult group were significant.

Analyses of RD Indices

Age group differences on RD indices. Mean and standard deviation values as well as 25th, 50th (median), and 75th percentile values of the older and young adult groups on all six RD indices are provided in Table 3.3. Descriptive and inferential statistics for age group differences on RD indices are provided in Table 3.4. Mann-Whitney *U* tests revealed that older adults performed significantly worse than young adults on all six RD indices.

Comparisons of effect sizes for age group differences on RD indices. Although analyses revealed that older adults performed significantly worse than young adults on all RD indices, effect sizes associated with the observed age group differences on RD indices were compared to elucidate the extent to which incorporating FP errors associated with prototypical distractors only, unrelated distractors only, or both prototypical and

unrelated distractors in calculating SoRD and NRD scores impacted observed age group differences. The effect size associated with the age group difference on the SoRD-unrelated index (List A targets vs. unrelated List B distractors only) was 24.24% smaller than the effect size associated with the age group differences on the SoRD (List A targets vs. all List B distractors) and SoRD-prototypical (List A targets vs. prototypical List B distractors only) indices, which were comparable. Additionally, the effect size associated with the age group difference on the NRD-unrelated index (List A targets vs. unrelated novel distractors only) was 38.89% smaller than the effect size associated with the age group difference on the NRD index (List A targets vs. all novel distractors), but was comparable to the effect size associated with the age group difference on the NRD-prototypical index (List A targets vs. prototypical novel distractors only). In sum, the extent to which older adults performed worse than young adults was smaller on the SoRD-unrelated and NRD-unrelated indices than on the SoRD and NRD indices.

The effect size associated with the age group difference on the SoRD index was 8.33% smaller than the effect size associated with the age group difference on the NRD index. In contrast, the effect size associated with the age group difference on the SoRD-prototypical index was 50.00% larger than the effect size associated with the age group difference on the NRD-prototypical index. Finally, the effect size associated with the age group difference on the SoRD-unrelated index was 13.64% larger than the effect size associated with the age group difference on the NRD-unrelated index. Thus, a larger age group difference on SoRD relative to NRD was observed on SoRD and NRD indices that included either FP errors associated with prototypical distractors only (i.e., SoRD-prototypical and NRD-prototypical) or semantically unrelated distractors only (i.e.,

SoRD-unrelated and NRD-unrelated), although the difference was substantially smaller in the context of the latter indices.

Within-group differences on RD indices. Descriptive and inferential statistics for within-group differences on RD indices are provided in Table 3.5. Friedman tests revealed a significant effect of RD index type within both the older adult, χ^2 (5, N = 55) = 104.77, p <.001, and young adult, χ^2 (5, N = 57) = 161.39, p <.001, groups. Wilcoxon signed-rank tests revealed different patterns of performances on RD indices within the older and young adult groups.

In the older adult group, scores were higher on the SoRD-unrelated index than on the SoRD and SoRD-prototypical indices; however, scores were comparable on the latter two indices. Additionally, scores were higher on the NRD-unrelated index than on the NRD and NRD-prototypical indices, although scores were comparable on the latter two indices (after an FDR adjustment). Furthermore, in the older adult group, scores were comparable on the SoRD and NRD indices, on the SoRD-prototypical and NRD-prototypical indices, and on the SoRD-unrelated and NRD-unrelated indices (after an FDR adjustment).

Performances on RD indices in the young adult group largely mirrored the pattern of performances that was observed in the older adult group. For example, scores were higher on the SoRD-unrelated index than on the SoRD index; however, in contrast to the older adult group, scores also were higher on the SoRD-prototypical index than on the SoRD index, and scores on the SoRD-prototypical and SoRD-unrelated indices were comparable. Additionally, scores were higher on the NRD-unrelated index than on the NRD and NRD-prototypical indices, although scores were comparable on the latter two

indices. Furthermore, in the young adult group, scores were comparable on the SoRD and NRD indices, on the SoRD-prototypical and NRD-prototypical indices, and on the SoRD-unrelated and NRD-unrelated indices.

Discussion

The present study demonstrated that FP errors associated with prototypical distractors substantially influence calculations of SoRD and NRD scores on the CVLT-II yes/no recognition memory trial. The examination of age group differences in FP errors revealed that, compared to young adults, a greater proportion of older adults made FP errors associated with prototypical List B, prototypical novel, and unrelated List B distractors. However, the two age groups did not differ in proportions of individuals who made FP errors associated with unrelated novel distractors. This finding is expected given that these items are generally conceptualized as the least challenging to identify as distractors or "non-targets" due to being both novel (i.e., were not presented at any point during task administration) and semantically unrelated – and therefore less similar – to target items.

Analyses also demonstrated that the two age groups yielded different patterns in proportions of individuals who made FP errors across the four subtypes. In particular, the pattern of FP errors within the older adult group suggests an age-related vulnerability to the effect of semantic interference from prototypical items on yes/no recognition testing, over and above an effect of source interference from List B items. In the older adult group, there was a greater proportion of individuals who made FP errors associated with distractors that are prototypical, or semantically related to targets (regardless of whether the items were from List B or novel) than there was of those who made FP errors

associated with non-prototypical, or semantically unrelated distractors. Moreover, there was a greater proportion of individuals who made FP errors associated with prototypical novel distractors (which present only semantic interference) than there was of those who made FP errors associated with unrelated List B distractors (which present only source interference). Furthermore, the proportions of individuals who made FP errors associated with prototypical List B distractors (which present *both* semantic and source interference) and prototypical novel distractors (which present *only semantic* interference) were comparable. These findings may imply that healthy older adults are even more vulnerable to semantic interference than source interference (i.e., experience even greater difficulty in identifying prototypical items as distractors, than in identifying List B items as distractors as a result of age-related source memory impairment) on the CVLT-II yes/no recognition memory trial. However, in the analysis of semantically unrelated distractors, there was a greater proportion of individuals who made FP errors associated with items that were from List B than those that were novel, which is not surprising given the research literature on age-related source memory impairment. Taken together, this set of findings regarding FP errors in the older adult group suggests that 1) older adults are particularly susceptible to inaccurately endorsing prototypical distractors over and above experiencing difficulty in identifying List B items as distractors, and 2) in the context of semantically unrelated distractors only, continue to exhibit difficulty in identifying List B items as distractors. These findings provide more evidence of age-related source memory impairment as well as highlight that prototypical items, by introducing semantic interference, are an additional source of confusion or difficulty for healthy older adults on yes/no recognition memory testing.

In the young adult group, there was a greater proportion of individuals who made FP errors associated with prototypical distractors than there was of those who made FP errors associated with semantically unrelated distractors, only with regard to novel distractor items. Thus, young adults also may be prone to inaccurately endorsing prototypical distractors, albeit to a lesser extent than older adults based on an examination of effect sizes (see Table 3.2), although, in contrast to older adults, they are less likely to experience difficulty in aspects of yes/no recognition memory testing that rely on source memory (i.e., identifying List B items as distractors).

The examination of age group differences on SoRD, SoRD-prototypical, SoRD-unrelated, NRD, NRD-prototypical, and NRD-unrelated indices revealed that older adults performed significantly worse than young adults on all indices. Moreover, effect sizes associated with the observed age group differences on RD indices were compared to elucidate the extent to which incorporating FP errors associated with prototypical distractors only, unrelated distractors only, or both prototypical and unrelated distractors in calculating SoRD and NRD scores impacted observed age group differences. A particular emphasis was made on comparing the degree to which older and young adults differed on new, more refined SoRD and NRD indices that exclude FP errors associated with prototypical distractors (i.e., SoRD-unrelated and NRD-unrelated) relative to existing CVLT-II SoRD and NRD indices that include FP errors associated with both prototypical and semantically unrelated distractors. As expected, the effect sizes associated with age group differences on SoRD-unrelated and NRD-unrelated indices were smaller than the effect sizes associated with age group differences on SoRD and NRD indices, respectively. The reduction in age group differences is likely driven by the

notion that older adults showed greater improvements relative to young adults on indices that exclude FP errors associated with prototypical distractors, based on an examination of effect sizes. Analyses also revealed that the effect size associated with the age group difference on the SoRD index was smaller than the effect size associated with the age group difference on the NRD index. In contrast, the effect size associated with the age group difference on the SoRD-unrelated index was larger than the effect size associated with the age group difference on the NRD-unrelated index (this pattern was even more evident in the context of SoRD-prototypical and NRD-prototypical indices). In sum, this set of findings indicates that, by excluding contributions from FP errors associated with prototypical distractor items that are semantically related to target items in the calculation of SoRD and NRD scores, 1) age group differences on SoRD and NRD are smaller in magnitude, and 2) the extent to which older adults perform worse than young adults is greater on SoRD than on NRD, which further supports the notion that, relative to item memory, source memory is particularly vulnerable to age-related decline (Bayer et al., 2011; Dennis et al., 2008; Glisky & Kong, 2008; Hashtroudi, Johnson, & Chrosniak, 1989; McIntyre & Craik, 1987; Naveh-Benjamin & Craik, 1995; Schacter, Kaszniak, Kihlstrom, & Valdiserri, 1991; Spaniol, Madden, & Voss, 2006; Trott, Friedman, Ritter, Fabiani, & Snodgrass, 1999). A possible limitation of these findings is that the older adult group in the study sample was relatively well educated and may not fully represent the general population of cognitively healthy older adults. However, it is reasonable to suspect that observed age group differences on RD indices would be larger in a sample of individuals with less cognitive reserve. Moreover, the present findings highlight the potential for these refined RD indices to demonstrate clinical utility in the assessment and

characterization of recognition memory deficits in more cognitively impaired populations, such as individuals with neurodegenerative disease.

Analyses revealed different patterns of within-group differences on RD indices across the two age groups. In the older adult group, performances were higher on the SoRD-unrelated index than on the SoRD and SoRD-prototypical indices, whereas performances were comparable on the latter two indices. These findings suggest a significant influence of FP errors related to prototypical List B distractors on SoRD performances in older adults, and further highlight the cumulative effects of source and semantic interference on increasing the difficulty of identifying distractor items for older adults. Similarly, performances were higher on the NRD-unrelated index than on the NRD and NRD-prototypical indices, whereas performances were comparable on the latter two indices, suggesting a significant influence of FP errors related to prototypical novel distractors on NRD performances in older adults, and further highlighting the impact of semantic interference on yes/no recognition testing in older adults. Performances on the SoRD-prototypical and NRD-prototypical indices (and on the SoRD and NRD indices) were comparable, which is not surprising given that the proportions of older adults who made FP errors associated with prototypical List B and prototypical novel distractors were comparable. However, performances on the SoRD-unrelated and NRD-unrelated indices also were comparable despite the observation that the proportion of older adults who made FP errors associated with unrelated List B distractors was greater than the proportion of older adults who made FP errors associated with unrelated novel distractors. Nonetheless, this set of findings collectively suggests that older adults do benefit from the exclusion of FP errors associated with prototypical distractor items in

the calculation of *d'* scores for SoRD and NRD. Moreover, the findings suggest that disproportionate source memory impairments in older adults may be more evident in the close examination of FP errors, rather than through comparisons of scores on SoRD indices relative to NRD indices.

In the young adult group, performances were higher on the SoRD-unrelated and NRD-unrelated indices than on the SoRD and NRD indices, respectively. Moreover, performances were comparable on the SoRD and NRD indices, on the SoRD-prototypical and NRD-prototypical indices, and on the SoRD-unrelated and NRD-unrelated indices. Taken together, these results suggest that young adults also do benefit from the exclusion of FP errors associated with prototypical distractor items in the calculation of *d'* scores for SoRD and NRD. Nonetheless, disproportionate source or novel recognition memory impairments among young adults were not observed in analyses of FP errors or RD indices, which is not surprising given that relative weaknesses in source or item memory are not typically observed in young adulthood.

Overall, the present findings yield evidence for improved performances among both older and young adults on SoRD and NRD indices that exclude FP errors associated with prototypical distractor items in the calculation of *d'* scores. Moreover, the present findings highlight the important role of FP errors in the assessment of recognition discriminability and efforts to characterize recognition memory function and changes in healthy aging.

Conclusion

The present study examined the impact of different FP error subtypes on assessments of SoRD and NRD using the CVLT-II in a cognitively healthy sample, and

the degree of age-related differences on SoRD and NRD indices that exclude contributions from FP errors associated with prototypical distractors (i.e., SoRD-unrelated and NRD-unrelated) relative to age-related differences on original SoRD and NRD indices that include FP errors associated with both prototypical and semantically unrelated distractors. Both age groups demonstrated better performances on SoRD-unrelated and NRD-unrelated indices than on SoRD and NRD indices, respectively. Although older adults performed worse than young adults on all RD indices, age group differences were smaller in magnitude on the more refined SoRD-unrelated and NRD-unrelated indices relative to original SoRD and NRD indices, although older adults were shown to perform disproportionately worse than young adults on SoRD in the context of refined indices. Although CVLT-II indices of SoRD and NRD in their current form can reliably demonstrate age-related differences on these aspects of recognition memory function (i.e., those pertaining to source and item memory), the refined indices utilized in the present study may be used to further elucidate the extent to which healthy older and young adults differ on these particular constructs. Furthermore, the present findings highlight the potential for these refined RD indices to exhibit clinical utility in improving assessments and characterizations of recognition memory deficits in more cognitively impaired populations.

Acknowledgements

Dean C. Delis, Ph.D. is an author of the CVLT and receives royalties for the test. This work was supported in part by the National Institutes of Health (NIH) under grant number R01 AG034202 to P. E. G. and grant number K24 AG026431 to M. W. B.

Dissertation Author's Acknowledgements

I, Lisa Graves (primary investigator and author of this material), would like to thank all coauthors – Drs. Gilbert, Holden, Delano-Wood, Bondi, Corey-Bloom, and Delis, as well as Ms. Emily Van Etten – for their contributions to this work. I also would like to thank the journal of *Aging, Neuropsychology, and Cognition* for accepting this work for publication, and the Taylor & Francis Group for publishing this work.

Figures

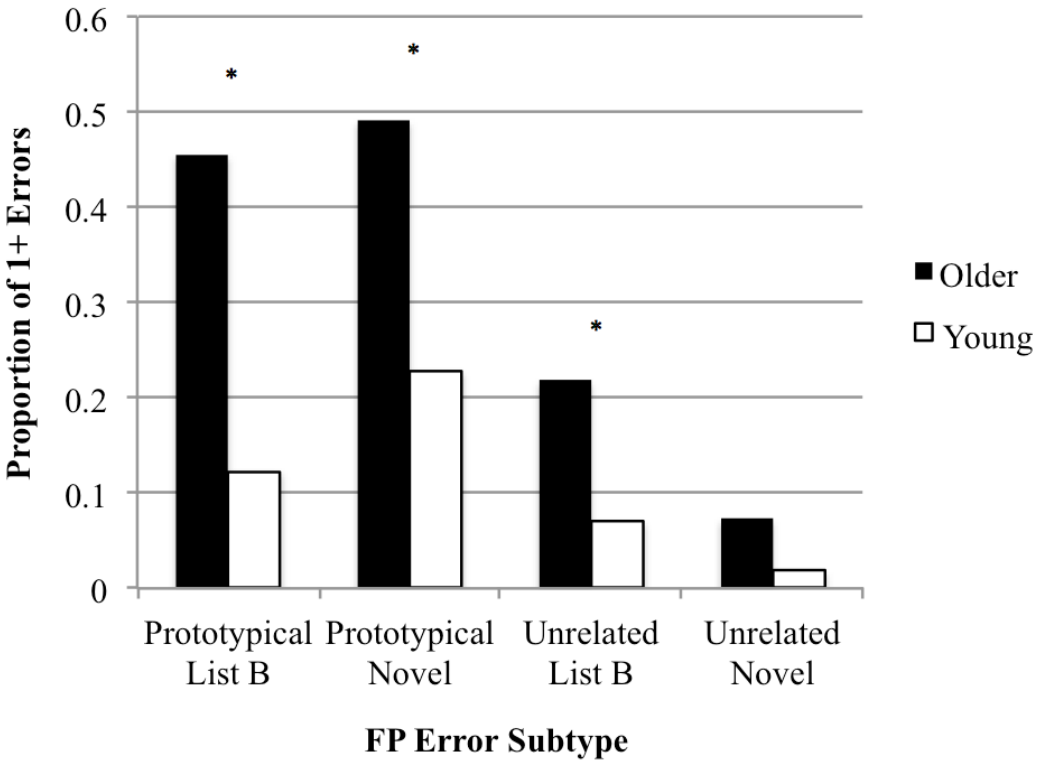

Figure 3.1. *Proportions of older and young adults who made one or more (1+) FP errors in each of the four subtypes: prototypical List B, prototypical novel, unrelated List B, and unrelated novel. Asterisks (*) indicate significant group differences*.

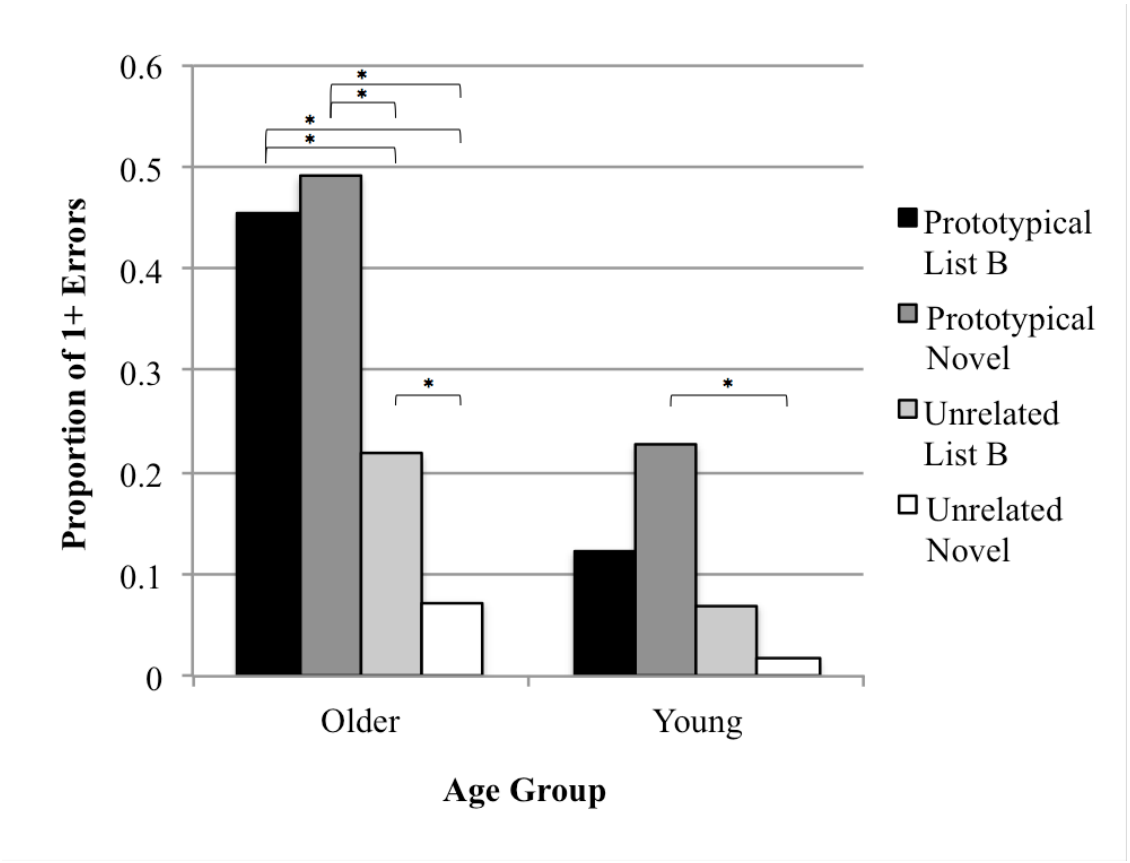

Figure 3.2. *Proportions of individuals who made one or more (1+) FP errors across subtype (prototypical List B, prototypical novel, unrelated List B, and unrelated novel) within each age group. Asterisks (*) indicate significant pairwise comparisons within groups.*

Tables

Table 3.1. *Age group differences in the number of individuals who made zero (0 FP) versus one or more (1+ FP) FP errors in each of the four subtypes, based on chi-square analyses.*

FP Error Subtype	Older			Young			χ^2	*p*	*r*
	0 FP	1+ FP	Proportion of 1+ FP	0 FP	1+ FP	Proportion of 1+ FP			
Prototypical List B	30	25	.455	50	7	.123	15.09	<.001*	.37
Prototypical Novel	28	27	.491	44	13	.228	8.42	<.01*	.27
Unrelated List B	43	12	.218	53	4	.070	5.01	<.05*	.21
Unrelated Novel	51	4	.073	56	1	.018	2.00	.202	.13

*p value retains significance following FDR adjustment

Table 3.2. *Comparisons within each age group in the number of individuals who made zero versus one or more FP errors across subtypes, based on chi-square tests of independence.*

Comparison	Older			Young		
	χ^2	*p*	*r*	χ^2	*p*	*r*
Prototypical List B vs. Unrelated List B	6.88	<.01*	.25	0.91	.341	.09
Prototypical List B vs. Prototypical Novel	0.15	.702	.04	2.18	.140	.14
Prototypical List B vs. Unrelated Novel	20.65	<.001*	.43	4.84	<.05	.21
Unrelated List B vs. Prototypical Novel	8.94	<.01*	.28	5.60	<.05	.22
Unrelated List B vs. Unrelated Novel	4.68	<.05*	.20	1.88	.170	.13
Prototypical Novel vs. Unrelated Novel	23.76	<.001*	.46	11.73	<.001*	.32

**p* value retains significance following FDR adjustment

Table 3.3. *Descriptive information for older and young adults on all six recognition discriminability (RD) indices.*

RD Index	Older				Young			
	Mean (SD)	25th %ile	50th %ile (*median*)	75th %ile	*Mean (SD)*	25th %ile	50th %ile (*median*)	75th %ile
SoRD	2.96 (0.71)	2.40	3.10	3.70	3.40 (0.45)	3.40	3.70	3.70
SoRD-prototypical	3.00 (0.94)	2.04	3.05	4.00	3.61 (0.55)	3.25	4.00	4.00
SoRD-unrelated	3.34 (0.73)	2.99	3.63	4.00	3.67 (0.53)	3.63	4.00	4.00
NRD	2.94 (0.67)	2.40	3.10	3.70	3.40 (0.40)	3.25	3.40	3.70
NRD-prototypical	2.85 (0.98)	2.21	2.77	4.00	3.50 (0.65)	3.05	3.63	4.00
NRD-unrelated	3.52 (0.56)	2.99	3.63	4.00	3.76 (0.34)	3.63	4.00	4.00

Note: RD = recognition discriminability; SoRD (source recognition discriminability) = List A targets vs. all List B distractors; SoRD-prototypical = List A targets vs. prototypical List B distractors only; SoRD-unrelated = List A targets vs. unrelated List B distractors only; NRD (novel recognition discriminability) = List A targets vs. all novel distractors; NRD-prototypical = List A targets vs. prototypical novel distractors only; NRD-unrelated = List A targets vs. unrelated novel distractors only

Table 3.4. *Age group differences on SoRD, NRD, SoRD-unrelated, and NRD-unrelated indices, with associated statistics based on Mann-Whitney U tests.*

RD Index	Older	Young			
	Mean rank (sum of ranks)	*Mean rank (sum of ranks)*	*U*	*p*	*r*
SoRD	45.90 (2524.50)	66.73 (3803.50)	984.50	<.001*	.33
SoRD-prototypical	46.12 (2536.50)	66.52 (3791.50)	996.50	<.01*	.33
SoRD-unrelated	48.81 (2684.50)	63.92 (3643.50)	1144.50	<.01*	.25
NRD	45.06 (2478.50)	67.54 (3849.50)	938.50	<.001*	.36
NRD-prototypical	45.35 (2494.50)	67.25 (3833.50)	954.50	<.001*	.22
NRD-unrelated	49.70 (2733.50)	63.06 (3594.50)	1193.50	<.05*	.22

**p* value retains significance following FDR adjustment

Note: RD = recognition discriminability; SoRD (source recognition discriminability) = List A targets vs. all List B distractors; SoRD-prototypical = List A targets vs. prototypical List B distractors only; SoRD-unrelated = List A targets vs. unrelated List B distractors only; NRD (novel recognition discriminability) = List A targets vs. all novel distractors; NRD-prototypical = List A targets vs. prototypical novel distractors only; NRD-unrelated = List A targets vs. unrelated novel distractors only

Table 3.5. *Within-group differences on SoRD, NRD, SoRD-unrelated, and NRD-unrelated indices, with associated statistics based on Wilcoxon signed-rank tests.*

Comparison	Older			Young		
	Z	*p*	*r*	*Z*	*p*	*r*
SoRD vs. NRD	0.33	.744	.04	0.43	.670	.06
SoRD-prototypical vs. NRD-prototypical	1.36	.173	.18	1.24	.214	.16
SoRD-unrelated vs. NRD-unrelated	2.16	<.05	.29	1.63	.102	.22
SoRD vs. SoRD-prototypical	0.18	.860	.02	4.16	<.001*	.55
SoRD vs. SoRD-unrelated	5.83	<.001*	.79	5.03	<.001*	.67
SoRD-prototypical vs. SoRD-unrelated	3.65	<.001*	.49	1.19	.234	.16
NRD vs. NRD-prototypical	2.05	<.05	.28	1.06	.291	.14
NRD vs. NRD-unrelated	6.21	<.001*	.84	5.88	<.001*	.78
NRD-prototypical vs. NRD-unrelated	4.40	<.001*	.59	2.89	<.01*	.38

*p value retains significance following FDR adjustment

Note: RD = recognition discriminability; SoRD (source recognition discriminability) = List A targets vs. all List B distractors; SoRD-prototypical = List A targets vs. prototypical List B distractors only; SoRD-unrelated = List A targets vs. unrelated List B distractors only; NRD (novel recognition discriminability) = List A targets vs. all novel distractors; NRD-prototypical = List A targets vs. prototypical novel distractors only; NRD-unrelated = List A targets vs. unrelated novel distractors only

References

Awipi, T., & Davachi, L. (2008). Content-specific source encoding in the human medial temporal lobe. *Journal of Experimental Psychology: Learning Memory and Cognition, 34,* 769–779.

Baddeley, A. D. (1966). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18*(4), 302–309.

Bayer, Z. C., Hernandez, R. J., Morris, A. M., Salomonczyk, D., Pirogovsky, E., & Gilbert, P. E. (2011). Age-related source memory deficits persist despite superior item memory. *Experimental Aging Research*, 37, 473–480.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological), 57*(1), 289–300.

Cansino, S., Maquet, P., Dolan, R. J., & Rugg, M. D. (2002). Brain activity underlying encoding and retrieval of source memory. *Cerebral Cortex, 12,* 1048–1056.

Craik, F. I., & McDowd, J. M. (1987). Age differences in recall and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(3), 474-479.

Craik, F. I., Morris, L. W., Morris, R. G., & Loewen, E. R. (1990). Relations between source amnesia and frontal lobe functioning in older adults. *Psychology and Aging, 5,* 148–151.

Danckert, S. L., & Craik, F. I. (2013). Does aging affect recall more than recognition memory? *Psychology and Aging, 28*(4), 902–909.

Delis, D. C. & Fridlund, A. J. (2000). *CVLT-II Comprehensive Scoring System and Computerized Report.* San Antonio, TX: The Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test.* San Antonio, TX: The Psychological Corporation.

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

Dennis, N. A., Hayes, S. M., Prince, S. E., Madden, D. J., Huettel, S. A., & Cabeza, R. (2008). Effects of aging on the neural correlates of successful item and source memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 791–808.

Ebert, P. L., & Anderson, N. D. (2009). Proactive and retroactive interference in young adults, healthy older adults, and older adults with amnestic mild cognitive impairment. *Journal of the International Neuropsychological Society, 15,* 83-93.

Ekstrom, A. D., & Bookheimer, S. Y. (2007). Spatial and temporal episodic memory retrieval recruit dissociable functional networks in the human brain. *Learning and Memory, 14,* 645–654.

Fan, J., Snodgrass, J., & Bilder, R. M. (2003). Functional magnetic resonance imaging of source versus item memory. *Neuroreport, 14,* 2275–2281.

Glisky, E. L., & Kong, L. L. (2008). Do young and older adults rely on different processes in source memory tasks? A neuropsychological study. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 34,* 809–822.

Glisky, E. L., Polster, M. R., & Routhieaux, B. C. (1995). Double dissociation between item and source memory. *Neuropsychology, 9,* 229–235.

Glisky, E. L., Rubin, S. R., & Davidson, P. S. R. (2001). Source memory in older adults: An encoding or retrieval problems? *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 1131–1146.

Hashtroudi, S., Johnson, M. K., & Chrosniak, L. D. (1989). Aging and source monitoring. *Psychology and Aging, 4,* 106–112.

Henkel, L. A., Johnson, M. K., & De Leonardis, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General, 127,* 251–268.

Janowsky, J. S., Shimamura, A. P., & Squire, L. R. (1989). Source memory impairment in patients with frontal lobe lesions. *Neuropsychologia, 27,* 1043–1056.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–8.

Jurica, P. J., Leitten, S., Mattis, S. (2001). *Dementia Rating Scale-2: Professional manual.* Lutz, FL: Psychological Assessment Resources.

Kausler, D. H. (1994). *Learning and memory in normal aging.* Academic Press: San Diego, CA.

Kirwan, C. B., Wixted, J. T., & Squire, L. R. (2008). Activity in the medial temporal lobe predicts memory strength, whereas activity in the prefrontal cortex predicts recollection. *Journal of Neuroscience, 28,* 10541–10548.

Macmillan, N. A., & Creelman, D. C. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

McIntyre, J. S., & Craik, F. I. M. (1987). Age difference in memory for item and source. *Canadian Journal of Psychology, 41,* 175–192.

Mitchell, K. J., Raye, C. L., Johnson, M. K., & Greene, E. J. (2006). An fMRI investigation of short-term source memory in young and older adults. *Neuroimage, 30,* 627–633.

Naveh-Benjamin, M., & Craik, F. I. (1995). Memory for context and its use in item memory: Comparisons of younger and older persons. *Psychology and Aging, 10,* 284–293.

Peters, J., Koch, B., Schwarz, M., & Daum, I. (2007). Domain-specific impairment of source memory following a right posterior medial temporal lobe lesion. *Hippocampus, 17,* 505–509.

Schacter, D. L., Harbluk, J. L., & McLachlin, D. R. (1984). Retrieval without recollection: An experimental analysis of source amnesia. *Journal of Verbal Learning and Verbal Behavior, 23,* 593–611.

Schacter, D. L., Kaszniak, A. W., Kihlstrom, J. F., & Valdiserri, M. (1991). The relation between source memory and aging. *Psychology and Aging, 6,* 559–568.

Shimamura, A. P., & Squire, L. R. (1987). A neuropsychological study of fact memory and source amnesia. *Journal Experimental Psychology: Learning, Memory, and Cognition, 13,* 464–473.

Spaniol, J., Madden, D. J., & Voss, A. (2006). A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32,* 101–117.

Stark, C. E., & Squire, L. R. (2000). Functional magnetic resonance imaging (fMRI) activity in the hippocampal region during recognition memory. *Journal of Neuroscience, 20,* 7776–7781.

Stark, C. E. & Squire, L. R. (2003). Hippocampal damage equally impairs memory for single items and memory for conjunctions. *Hippocampus, 13,* 281–292.

Trott, C. T., Friedman, D., Ritter, W., Fabiani, M., & Snodgrass, J. G. (1999). Episodic priming and memory for temporal source: Event related potentials reveal age-related differences in prefrontal functioning. *Psychology and Aging, 14,* 390–413.

Turner, M. L., & Pinkston, R. S. (1993). Effects of a memory and aging workshop on negative beliefs of memory loss in the elderly. *Educational Gerontology, 19*(5), 359-373. doi:10.1080/0360127930190501

Van der Linden, M., Philippot, P., & Heinen, P. (1997). Effect of age, education and verbal efficiency on memory performance and memory self-assessment. *Archives de Psychologie, 65*(254), 171-185.

Woodruff-Pak, D. S., & Finkbiner, R. G. (1995). Larger nondeclarative than declarative deficits in learning and memory in human aging. *Psychology and Aging, 10*(3), 416-426.

CHAPTER 4:

Study 3

The content within this section, titled "Chapter 4: Study 3," reflects material from a paper that has been published in the *Journal of the International Neuropsychological Society*. The proper citation is as follows:

Graves, L. V., Holden, H. M., Van Etten, E. J., Delano-Wood, L., Bondi, M. W., Salmon, D. P., Corey-Bloom, J., Delis, D. C., & Gilbert, P. E. (2018). New Yes/No Recognition analysis on the California Verbal Learning Test-3: Clinical utility in Alzheimer's and Huntington's disease. *Journal of the International Neuropsychological Society, 24*(8), 833-841. doi:10.1017/S1355617718000474

Abstract

Objective: The third edition of the California Verbal Learning Test (CVLT-3) includes a new index termed List A vs. Novel/Unrelated recognition discriminability (RD) on the Yes/No Recognition trial. Whereas the Total RD index incorporates false positive (FP) errors associated with all distractors (including List B and semantically-related items), the new List A vs. Novel/Unrelated RD index incorporates only FP errors associated with novel, semantically-unrelated distractors. Thus, in minimizing levels of source and semantic interference, the List A vs. Novel/Unrelated RD index may yield purer assessments of yes/no recognition memory independent of vulnerability to source memory difficulties or semantic confusion, both of which are often seen in individuals with primarily frontal-system dysfunction (e.g., early Huntington's disease [HD]). **Method:** We compared the performance of individuals with Alzheimer's disease (AD) and HD in mild and moderate stages of dementia on CVLT-3 indices of Total RD and List A vs. Novel/Unrelated RD. **Results:** Although AD and HD subgroups exhibited deficits on both RD indices relative to healthy comparison groups, those with HD generally outperformed those with AD, and group differences were more robust on List A vs. Novel/Unrelated RD than on Total RD. **Conclusions:** Our findings highlight the clinical utility of the new CVLT-3 List A vs. Novel/Unrelated RD index, which (a) maximally assesses yes/no recognition memory independent of source and semantic interference; and (b) provides a greater differentiation between individuals whose memory disorder is primarily at the encoding/storage level (e.g., as in AD) versus at the retrieval level (e.g., as in early HD).

Keywords: Alzheimer disease, Huntington disease, memory disorders, recognition, memory and learning tests, neuropsychological tests

New Yes/No Recognition Memory Analysis on the California Verbal Learning Test-3: Clinical Utility in Alzheimer's and Huntington's Disease

Alzheimer's disease (AD) is associated with early medial temporal lobe damage, particularly in the hippocampal formation, subsequent damage to cortical association areas, and relative sparing of most subcortical structures (Braak & Braak, 1991; Hyman, Van Hoesen, Damasio, & Barnes, 1984). Huntington's disease (HD), in contrast, is associated with early damage to basal ganglia structures (Vonsattel, 2000; Vonsattel et al., 1985) that have extensive projections to the frontal lobes (Alexander, Crutcher, & DeLong, 1990; Crosson et al., 2003; Cummings, 1993), followed by more diffuse involvement of other cortical and subcortical regions and networks. Research has shown that the different patterns of neurodegeneration associated with AD and HD yield distinct profiles of memory loss (Delis et al., 1991; Hodges, Salmon, & Butters, 1990; Moss, Albert, Butters, & Payne, 1986; Salmon & Bondi, 2009; Salmon & Filoteo, 2007; Troster et al., 1993). Individuals with AD usually have pervasive memory deficits characterized by poor learning, rapid forgetting, and poor recognition (Budson & Kowall, 2013), a profile of memory loss thought to reflect an encoding/storage deficit. Patients with early stage HD often have significant deficits in recall with less compromised recognition (Butters, Wolfe, Granholm, & Martone, 1986; Butters, Wolfe, Martone, Granholm, & Cermak, 1985; Lundervold, Reinvang, & Lundervold, 1994; Martone, Butters, Payne, Becker, & Sax, 1984; Massman, Delis, Butters, Levin, & Salmon, 1990), a profile of memory loss thought to reflect primarily a retrieval deficit. Although recognition is less impaired than recall in early HD, recognition is still often significantly impaired, particularly in the later stages of disease, raising the possibility that encoding processes

are also compromised to at least some degree (see Montoya et al., 2006 for review). Given that the prefrontal cortex has been shown to be implicated in encoding processes (e.g., Blumenfeld & Ranganath, 2007; Tulving et al., 1994) and that HD is associated with frontal system pathology and dysfunction, the extent to which encoding is affected in HD may at least partly depend on the degree to which prefrontal networks become compromised throughout the disease process. Nonetheless, it is unlikely that encoding deficits in HD would ever reach a level of severity comparable to what is observed in AD, given the disproportionately greater impact of AD on medial temporal regions that play a more integral role in encoding processes. Rather, the pattern of memory dysfunction in HD is likely best characterized as primarily a retrieval deficit, even when accompanied by mild encoding difficulties.

In efforts to characterize profiles of memory loss, the degree to which recognition memory is affected provides insight into whether impaired recall reflects (a) failure to encode/store information at the outset (i.e., an encoding/storage deficit, as in AD), or (b) compromised retrieval processes that warrant prompting or cuing to facilitate recognition of previously encoded information (i.e., a retrieval deficit, as in early HD). Although the extant literature suggests that a major distinction between the memory profiles associated with AD and with HD largely involves the extent to which recognition memory is impaired, the nature and degree to which it is affected in HD in particular is less clear and warrants further exploration.

CVLT Studies of Yes/No Recognition Memory in AD and HD

The California Verbal Learning Test (CVLT) is a list-learning measure that assesses a multitude of verbal learning and memory indices, including immediate recall,

free and cued recall over short and long delays, and Yes/No Recognition. Studies using the original CVLT (Delis, Kramer, Kaplan, & Ober, 1987) consistently demonstrated that among individuals in mild stages of dementia, deficits on the Yes/No Recognition trial are less severe in those with HD than in those with AD (Delis et al., 1991; Kramer et al., 1988; Kramer, Levin, Brandt, & Delis, 1989). This difference was shown with the original CVLT recognition discriminability (RD) index that measures the ability to distinguish List A targets from all distractors on the Yes/No Recognition trial (Delis et al., 1991; Kramer et al., 1988). In contrast, studies that compared individuals with AD or HD on the Yes/No Recognition trial of the second edition of the CVLT (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000) have produced inconsistent results. While one study found that individuals with HD obtained higher standardized scores than those with AD on the CVLT-II Total RD index (Fine et al., 2008), another study with a larger sample found that AD and HD groups performed comparably on this measure (Graves et al., 2017). One implication of this pattern of results is that patients with HD may have worse Yes/No Recognition performance on the CVLT-II than on the original CVLT. This possibility is consistent with the clinical observation that Total RD scores of patients with HD are generally lower on the CVLT-II than on the original CVLT (Dean C. Delis, personal communication, October 26, 2017).

Reasons for differences in the performance of individuals with HD on the recognition components of the two versions of the CVLT may lie in differences in how RD is determined. The Yes/No Recognition trial of the original CVLT included only half (i.e., eight) of the 16 List B items as distractors (Delis et al., 1987). Due to a ceiling effect in cognitively normal individuals, the trial's difficulty was increased in the CVLT-II by

including all 16 List B items as distractors (Dean C. Delis, personal communication, September 26, 2017). Although this had the intended effect of making the Yes/No Recognition trial more difficult, it potentially made the test more sensitive to deficits in source memory. Individuals with frontal-system dysfunction (e.g., HD) may have particular difficulty in identifying the source of each previously-presented item (List A or List B) during the Yes/No Recognition trial when asked whether or not an item had been on List A (Fine et al., 2008). Increasing the number of List B distractors on the CVLT-II Yes/No Recognition trial may have amplified this difficulty.

The CVLT-II Yes/No Recognition trial also had an increased proportion of distractors that are semantically-related to List A target items (8 of 28 distractors for CVLT versus 16 of 32 distractors for CVLT-II; Delis et al., 2000). Research has shown that patients with frontal-system dysfunction are prone to making semantic intrusion or semantic confusion errors due to impaired inhibition of activation within semantic networks (e.g., Baldo et al., 2002). A deficit in inhibition of the semantic network during the CVLT may lead individuals with HD to have greater difficulty in rejecting distractors that share obvious semantic associations with targets than in rejecting distractors that do not (the same deficit could lead to semantically-related intrusion errors during free recall trials). This would have a greater adverse effect on the CVLT-II than the CVLT for individuals with HD due to the increased proportion of semantically-related distractors. Increasing the proportion of semantically-related distractors may not have the same effect on individuals with AD since their severe recognition memory deficits reflect a profound encoding/storage deficit that can be attributed to more extensive neuropathology targeting the medial temporal lobes and cortical association areas. Thus, individuals with AD are

likely to exhibit relatively comparable levels of difficulty in rejecting novel distractors whether or not they share obvious semantic associations with targets.

A Purer Sub-Measure of Novel RD on the CVLT-3

While the CVLT-II included eight novel unrelated distractor items on the Yes/No Recognition trial, it did not provide a separate index that assessed the ability of individuals to endorse List A targets while rejecting those novel unrelated distractors. The second and third editions of the CVLT (CVLT-II and CVLT-3, respectively) contain the same target words on the recall trials and the same targets and distractors on the Yes/No Recognition trial (in fact, the only word-item changes that were made to the CVLT-3 are on the Forced Choice Recognition trial). However, the CVLT-3 (Delis, Kramer, Kaplan, & Ober, 2017) includes a purer RD index, List A vs. Novel/Unrelated RD, that isolates the ability to distinguish List A targets from distractors that were not previously presented during test administration and do not share obvious semantic associations with targets. Thus, the new List A vs. Novel/Unrelated RD index minimizes any potential influences of source and semantic interference, and is therefore thought to provide a more refined assessment of yes/no recognition memory.

The present study sought to elucidate the nature of AD and HD differences in yes/no recognition memory by comparing the performance of individuals with AD and HD in mild and moderate stages of dementia on CVLT-3 indices of Total RD and List A vs. Novel/Unrelated RD. It was hypothesized that although both AD and HD would be associated with deficits on Yes/No Recognition testing, individuals with HD would perform better than those with AD, particularly on the new List A vs. Novel/Unrelated RD index. In other words, the List A vs. Novel/Unrelated RD index, in minimizing any

potential influences of source and semantic interference, was expected to exhibit greater utility than the Total RD index in distinguishing the memory profiles of individuals with AD versus HD.

Method

Participants

Study participants were 52 individuals with AD, 55 individuals with HD, 53 healthy older adults (OA), and 31 healthy middle-age adults (MA); the healthy OA and MA groups were included to serve as AD and HD comparison groups, respectively. The Dementia Rating Scale (DRS) or the Dementia Rating Scale-2 (DRS-2; Jurica, Leitten, & Mattis, 2001) was administered to individuals in the AD and HD subgroups to provide an assessment of global cognitive functioning. Individuals with AD and HD were characterized as mild or moderate in dementia severity based on DRS/DRS-2 scores: 120+ = mild, 100-119 = moderate (mod). Accordingly, the study sample consisted of six total groups, with 25 Alzheimer's disease-mild (AD-mild), 27 Alzheimer's disease-moderate (AD-mod), 39 Huntington's disease-mild (HD-mild), 16 Huntington's disease-moderate (HD-mod), 53 OA, and 31 MA participants. Individuals with AD were recruited from the Shiley-Marcos Alzheimer's Disease Research Center (ADRC) affiliated with the University of California, San Diego (UCSD). Diagnoses of individuals with probable AD were made by a senior staff neurologist at the ADRC and were consistent with the criteria established by the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (ADRDA) workgroups (McKhann et al., 1984; McKhann et al., 2011).

Individuals with HD were recruited from the Huntington's Disease Clinical Research Center (HDCRC) at UCSD and were administered the Unified Huntington's Disease Rating Scale (UHDRS; Huntington Study Group, 1996) by a senior staff neurologist. Individuals with HD were diagnosed with definite HD on the basis of unequivocal motor signs on the UHDRS and a positive family history of HD. Participants in the HD-mild group had an average Total Motor Score (TMS) of 34.89 (*SD* = 14.24), and participants in the HD-mod group had an average TMS of 50.00 (*SD* = 16.94), with higher scores indicating greater severity of motor symptoms. In addition, all HD participants had a CAG repeat length greater than 39, indicating that all carried the fully penetrant genetic mutation for HD. Participants in the HD-mild group had an average of 44.57 (*SD* = 3.48) CAG repeats, and participants in the HD-mod group had an average of 45.88 (*SD* = 4.30) CAG repeats. Portions of the AD and HD groups in the present study overlap with the samples used in previous studies (Delis et al., 2005; Fine et al., 2008; Graves et al., 2017). Exclusionary criteria for AD and HD participants included the following: a diagnosis of any neurological disorder aside from AD or HD, respectively; a diagnosis of any major medical condition (e.g., cancer); a diagnosis of any major psychiatric disorder (with the exception of a mood or anxiety disorder for which any current symptoms must have been well managed); a history of traumatic brain injury; and a history of a substance use disorder. Whether participants with AD or HD met exclusionary criteria was determined based on information gathered via a combination of self-report, informant-report, and medical records.

Healthy OA participants were recruited from the San Diego community by the Center for Healthy Aging and Neurodegenerative Disease Research (CHANDR) at San

Diego State University (SDSU) directed by P. E. G. and the Normal Aging Laboratory at UCSD directed by M. W. B. Healthy MA participants were recruited from the San Diego community by the CHANDR directed by P. E. G. and the HDCRC directed by J. C. B. Exclusionary criteria for all healthy adult participants included the following: a diagnosis of any neurological disorder, a diagnosis of any major medical condition (e.g., cancer), a diagnosis of any major psychiatric disorder (with the exception of a mood or anxiety disorder, for which any current symptoms must have been well managed), a history of traumatic brain injury, and a history of a substance use disorder. Whether OA and MA participants met exclusionary criteria was determined based on information gathered primarily via self-report.

CVLT-II data were extracted from archival databases that included data from a larger battery of neuropsychological tests administered by trained research assistants or psychometrists at the ADRC, HDCRC, CHANDR, and Normal Aging Laboratory. All participants provided informed written consent and the study was approved by the Institutional Review Board of SDSU and/or UCSD.

CVLT-II and Yes/No Recognition Indices

The CVLT-II is a list-learning test that provides a multitude of verbal learning and memory indices, including immediate recall, free and cued recall over short and long delays, and Yes/No Recognition. The CVLT-II was administered using standard procedures outlined by Delis and colleagues (2000). Short- and long-delay tests of recall were separated by an interval of approximately 20 minutes, during which other nonverbal neuropsychological measures were administered.

Given that the CVLT-II and CVLT-3 contain identical target words on the recall trials and identical targets and distractors on the Yes/No Recognition trial, CVLT-3 algorithms were applied to CVLT-II data to generate scores on variables of interest in the present study: Total RD and List A vs. Novel/Unrelated RD. Raw *d'* scores on Total RD and List A vs. Novel/Unrelated RD indices were calculated using methods employed on all three versions of the CVLT (Delis et al., 1987; Delis et al., 2000; Delis et al., 2017) that are based on signal detection theory (see Macmillan & Creelman, 1991). In general, *d'* = *z*(hit rate) – *z*(FP rate), and raw *d'* scores on the Total RD and List A vs. Novel/Unrelated RD indices are therefore generated using the following formulas:

1. Total RD = *z*(Total Hit rate) – *z*(Total FP rate)
2. List A vs. Novel/Unrelated RD = *z*(Total Hit rate) – *z*(Novel/Unrelated FP rate)

The hit rate refers to the proportion of targets endorsed and the FP rate refers to the proportion of distractors endorsed. A *z*-transform is applied to each hit rate and FP rate, and subtracting the latter from the former yields *d'*. Thus, as the CVLT-3 manual (Delis et al., 2017) states, the raw *d'* score reflects the difference in standard deviation (*SD*) units between the examinee's hit rate (signal) and FP rate (noise). For instance, if the hit rate is 84% of the possible targets (approximately one *SD* above the expected mean) and the FP rate is 16% of the possible distractors (approximately one *SD* below the expected mean), the raw *d'* score is approximately +2.0. In contrast, if the hit rate is 16% and the FP rate is 84%, the raw *d'* score is approximately -2.0. If the hit rate and FP rate are both at 50% accuracy, then *d'* is 0. While the range of raw *d'* scores will vary depending on the number of FP errors, Total RD on the CVLT-3 can range from a high of +4.0 (16 hits,

0 FP errors) to a low of -4.0 (0 hits, 32 FP errors). Scaled scores on Total RD and List A vs. Novel/Unrelated RD indices were derived using the CVLT-3 standardization sample norms that adjust for age and gender.

Statistical Analyses

Analyses were conducted in the Statistical Package for the Social Sciences (SPSS) Version 25.

Demographic and preliminary analyses. Prior to conducting primary analyses, one-way analysis of variance (ANOVA) tests (with Tukey's post-hoc pairwise comparisons) and chi-square analyses were conducted to examine group differences on demographic variables, including age, gender, and education, as well as DRS/DRS-2 scores. Additionally, preliminary ANOVA and ANCOVA tests were conducted to determine whether any demographic variables were significant predictors of raw scores on Yes/No Recognition variables of interest.

Primary analyses. ANCOVA tests were conducted to examine the effect of group (AD-mod, AD-mild, HD-mod, HD-mild, OA, MA) on raw scores on Total RD and List A vs. Novel/Unrelated RD indices, while controlling for demographic factors when appropriate. Additionally, ANOVA tests were conducted to examine the effect of group on scaled scores on Total RD and List A vs. Novel/Unrelated RD indices. Post-hoc pairwise comparisons were conducted to examine group differences on raw and scaled scores on Total RD and List A vs. Novel/Unrelated RD indices in the context of significant group effects. The following comparisons were of primary interest and are emphasized in the discussion of results and their implications: 1) AD-mod versus HD-mod, 2) AD-mild versus HD-mild, 3) AD-mod versus AD-mild, 4) HD-mod versus HD-

mild, 5) AD-mod versus HD-mild, and 6) AD-mild versus HD-mod. The Bonferroni adjustment applied to this set of comparisons was: $\alpha = .05/6 = .008$. Cohen's *d* effect size values associated with significant AD and HD group differences were calculated and reported. In addition, the following comparisons were conducted to provide information regarding the level of performance that may be expected among cognitively healthy groups relative to demographically similar but clinically impaired AD and HD groups on the new CVLT-3 List A vs. Novel/Unrelated index: 1) AD-mod versus OA, 2) AD-mild versus OA, 3) HD-mod versus MA, and 4) HD-mild versus MA. The Bonferroni adjustment applied to this set of comparisons was: $\alpha = .05/4 = .013$.

Exploratory analyses. Regression analyses were conducted to explore whether TMS scores and number of CAG repeats were significant predictors of raw scores on Total RD and List A vs. Novel/Unrelated RD indices in participants with HD. Exploratory analyses involving participants with AD could not be conducted, as clinical data were not available on these individuals.

Results

Demographic Analyses

Demographic information on study participants is provided in Table 1. One-way ANOVA tests revealed a significant effect of group on age, $F(5, 185) = 109.05$, $p < .001$, education, $F(5, 185) = 5.04$, $p < .001$, and DRS/DRS-2 scores, $F(3, 103) = 114.35$, $p < .001$. Tukey's post-hoc pairwise comparisons revealed that, as expected, the AD-mild, AD-mod, and OA groups were significantly older than the HD-mild, HD-mod, and MA groups (*p*s < .001). However, there were no differences in age among the AD-mod, AD-mild, and OA groups (*p*s > .05), or among the HD-mod, HD-mild, and MA groups (*p*s >

.05). In addition, Tukey's post-hoc pairwise comparisons revealed that the OA group completed significantly more years of education than the HD-mild, HD-mod, and MA groups (*p*s < .05). However, there were no differences in education among the AD-mild, AD-mod, HD-mild, and HD-mod groups (*p*s > .05); among the AD-mild, AD-mod, and OA groups (*p*s > .05); or among the HD-mild, HD-mod, and MA groups (*p*s > .05). Furthermore, Tukey's post-hoc pairwise comparisons revealed that, as expected, the AD-mod and HD-mod groups had significantly lower DRS/DRS-2 scores than the AD-mild and HD-mild groups (*p*s < .001). However, there were no differences in DRS/DRS-2 scores between the AD-mod and HD-mod groups ($p > .05$), or between the AD-mild and HD-mild groups ($p > .05$). The chi-square analysis revealed no differences in gender distributions across groups, χ^2 (5, $N = 191$) = 9.52, $p = .09$.

Preliminary Analyses

Age was shown to be a significant predictor of raw scores on the Total RD index, $F(1, 189) = 3.74$, $p < .05$, but not the List A vs. Novel/Unrelated RD index, $F(1, 189) = 2.59$, $p = .11$. Given the evidence for significant group differences on age, and for a significant effect of age on aspects of Yes/No Recognition performance, age was included as a covariate in all primary analyses involving raw scores. As scaled scores correct for age, age was not included as a covariate in any primary analyses involving scaled scores.

Gender was shown to be a significant predictor of raw scores on the Total RD index, $F(1, 189) = 3.87$, $p < .05$, but not the List A vs. Novel/Unrelated RD index, $F(1, 189) = 1.29$, $p = .26$. Although gender distributions did not vary significantly across groups, gender was controlled for in all primary analyses involving raw scores given the

evidence for a significant effect of gender on aspects of Yes/No Recognition performance. As scaled scores correct for gender, gender was not controlled for in any primary analyses involving scaled scores.

DRS/DRS-2 scores were shown to be a significant predictor of raw scores on the Total RD index, $F(1, 105) = 25.85$, $p < .001$, and the List A vs. Novel/Unrelated RD index, $F(1, 105) = 18.41$, $p < .001$. However, given that DRS/DRS-2 scores were systematically varied by group (i.e., individuals with AD and HD were characterized as mild or moderate in dementia severity), DRS/DRS-2 scores were not controlled for in primary analyses involving raw or scaled scores. Education was not shown to be a significant predictor of raw scores on Yes/No Recognition variables of interest (*p*s > .05).

Primary Analyses: AD and HD Performances on Total RD and List A vs. Novel/Unrelated RD Indices

ANCOVA tests revealed a significant effect of group on raw scores on the Total RD index, $F(5, 183) = 71.88$, $p < .001$, and the List A vs. Novel/Unrelated RD index, $F(5, 183) = 39.86$, $p < .001$, controlling for age and gender. Post-hoc tests with Bonferroni adjustments for multiple comparisons revealed that, on both indices, the AD-mild and AD-mod groups exhibited significantly lower raw scores than the OA group (*p*s < .001), and the HD-mild and HD-mod groups exhibited significantly lower raw scores than the MA group (*p*s < .001). Additionally, the HD-mild group exhibited significantly higher raw scores than the AD-mild and AD-mod groups on both indices (*ps* < .01). Furthermore, the HD-mod group exhibited significantly higher raw scores than the AD-mod group on the List A vs. Novel/Unrelated RD index ($p = .001$), although this difference was not observed on the Total RD index (after a Bonferroni adjustment). No

other significant group differences on raw scores on the Total RD and List A vs. Novel/Unrelated RD indices were observed.

ANOVA tests also revealed a significant effect of group on scaled scores on the Total RD index, $F(5, 185) = 66.68$, $p < .001$, and the List A vs. Novel/Unrelated RD index, $F(5, 185) = 41.16$, $p < .001$. Post-hoc tests with Bonferroni adjustments for multiple comparisons revealed that, on both indices, the AD-mild and AD-mod groups exhibited significantly lower scaled scores than the OA group (*p*s < .001), and the HD-mild and HD-mod groups exhibited significantly lower scaled scores than the MA group (*p*s < .001). Additionally, the HD-mild group exhibited significantly higher scaled scores than the AD-mod group on the List A vs. Novel/Unrelated RD index ($p = .001$), although this difference was not observed on the Total RD index (after a Bonferroni adjustment). No other significant group differences on scaled scores on the Total RD and List A vs. Novel/Unrelated RD indices were observed.

Descriptive and inferential statistics associated with analyses involving raw and scaled scores on the Total RD and List A vs. Novel/Unrelated RD indices are provided in Tables 2 and 3. Relevant group differences on raw and scaled scores on the two indices are illustrated in Figures 1 and 2.

Exploratory analyses. Regression analyses indicated neither TMS scores nor number of CAG repeats were significant predictors of raw scores on Total RD or List A vs. Novel/Unrelated RD indices in participants with HD (*p*s > .05).

Discussion

The present study compared the performance of individuals with AD and HD in mild and moderate stages of dementia on indices of Total RD and List A vs. Novel/Unrelated RD

that were developed for the CVLT-3. Group differences on RD indices involving the AD-mod, AD-mild, HD-mod, and HD-mild groups were of primary interest; however, OA and MA groups were included as cognitively healthy comparison groups for AD and HD, respectively. Because the CVLT-3 List A vs. Novel/Unrelated index is a new measure, the OA and MA groups were included in analyses of scaled scores, in addition to raw scores, to provide information regarding the level of scaled score performance that might be expected from cognitively healthy individuals demographically similar to clinically impaired patients with AD or HD. Results showed that all AD and HD subgroups performed significantly worse than their respective healthy comparison groups on all Yes/No Recognition RD raw scores and scaled scores.

Analysis of raw scores on Yes/No Recognition indices showed that the HD-mild group performed significantly better than the AD-mod and AD-mild groups on both the Total RD and List A vs. Novel/Unrelated RD indices. Additionally, the HD-mod group performed significantly better than the AD-mod group on the List A vs. Novel/Unrelated RD index; notably, this difference was not observed on the Total RD index. These findings demonstrate that, in the context of raw scores, both the Total RD and List A vs. Novel/Unrelated RD indices are able to reveal less severe yes/no recognition memory deficits in mild HD than in mild AD. Importantly, however, as predicted, the List A vs. Novel/Unrelated RD index, but not the Total RD index, yielded less severe yes/no recognition memory deficits in moderate HD than in moderate AD. The flowchart below outlines the pattern of HD and AD performance that may be expected with raw scores on the CVLT-3 Total RD and List A vs. Novel/Unrelated RD indices, and may serve as a

helpful reference for clinicians and researchers when using the CVLT-3 to assess Yes/No Recognition performance in individuals with HD or AD.

Analysis of scaled scores on Yes/No Recognition indices showed no group differences among the AD and HD subgroups on the Total RD index. This is consistent with previous findings of comparable performance by individuals with AD or HD on the CVLT-II Total RD index (Graves et al., 2017), and extends earlier findings by showing comparable performance on the Total RD index in AD and HD across mild and moderate stages of dementia severity. Importantly, the HD-mild group performed significantly better than the AD-mod group on the scaled score for the List A vs. Novel/Unrelated RD index. Thus, even in the context of scaled scores (albeit to a lesser extent than in the context of raw scores), AD and HD differences on yes/no recognition memory are detectable, but only using a purer index of RD that minimizes potential influences of source and semantic interference (i.e., List A vs. Novel/Unrelated RD).

The discrepancy between findings from analyses involving raw scores and those involving scaled scores warrants discussion. First, we acknowledge that the difference may have been partly due to limited statistical power, given the relatively small number of participants in the HD-mod group in particular. However, we believe the discrepancy more likely reflects an issue in converting raw scores into scaled scores on indices with potential ceiling effects. Given that most cognitively normal individuals are expected to perform well on these RD indices (particularly on List A vs. Novel/Unrelated RD), lower raw scores in cognitively impaired individuals (e.g., the participants with AD or HD in our study) are likely to correspond with significantly reduced scaled scores. In addition, while we treated age as a continuous variable in our analyses of raw scores, the CVLT-3

normative sample was stratified into age groups, and age-corrected scaled scores are derived from these categorical groupings. Moreover, given that individuals with HD are younger, on average, than individuals with AD, raw scores in those with HD may be submitted to a more stringent age correction, which could further result in smaller HD and AD differences in the context of scaled scores relative to raw scores. On that premise, it is worth noting that, when analyzing raw scores, including age as a covariate when age is confounded with group or diagnosis is not the most ideal method for parceling out the effects of age on performance, and this is an inherent issue in many studies comparing AD and HD performance using raw scores. Moreover, we did not possess the statistical power that would be required to account for age using more sophisticated statistical methods (e.g., stratification). We encourage readers to take these issues into consideration in the evaluation of the present findings. Nonetheless, due to the aforementioned reasons, we believe that examining performance using raw scores (controlling for age as a continuous variable albeit its limitations) may yield greater sensitivity and better reflect the utility of the CVLT-3 RD indices in elucidating the degree to which yes/no recognition memory is impaired in HD versus AD.

Taken together, the present results with both raw and scaled scores indicate that the new CVLT-3 List A vs. Novel/Unrelated RD index has a more robust capacity than the Total RD index to detect differences in the recognition memory deficits associated with AD and HD. In particular, the present findings suggest that recognition memory deficits are less severe in HD than in AD and support the notion that the memory profile of HD reflects primarily a retrieval deficit, whereas the memory profile of AD reflects a more profound encoding/storage deficit. The Total RD index incorporates FP errors from

all distractor types (including those from List B and those that are novel but share obvious semantic associations with targets), whereas the List A vs. Novel/Unrelated RD index incorporates only FP errors associated with distractors that are novel and do not share obvious semantic associations with targets. Thus, the present findings provide evidence that individuals with HD may be particularly vulnerable to 1) endorsing List B distractors that are likely confounded by source interference, and 2) endorsing novel distractors that share obvious semantic associations with targets and are therefore likely confounded by semantic interference. Accordingly, those with HD may perform more similarly to individuals with AD on the Total RD index due to source memory deficits and semantic interference sensitivity associated with their frontal-system dysfunction. The present findings also support the hypothesis of Graves et al. (2017) that the higher proportion of List B and semantically related distractors relative to targets on the CVLT-II than on the original CVLT may have increased the difficulty of the CVLT-II Yes/No Recognition trial specifically for individuals with HD, thereby making their performance on the CVLT-II Total RD index similar to that of individuals with AD.

In conclusion, the present findings provide evidence that the endorsement of distractors on Yes/No Recognition testing may be influenced by both 1) their novelty (i.e., whether or not they were previously presented during test administration), and 2) their degree of semantic association with targets. While it is probably not feasible to develop an RD index that is completely free of any influences of source and/or semantic interference, the present findings indicate that the new CVLT-3 List A vs. Novel/Unrelated RD index minimizes these effects compared to the Total RD index. Thus, while the Total RD index provides a global, more sensitive measure of yes/no

recognition memory in general, the new List A vs. Novel/Unrelated RD index likely provides a purer measure of yes/no recognition memory independent of source and semantic interference, and may therefore exhibit greater utility in differentiating levels of yes/no recognition memory impairment in HD versus AD.

Acknowledgements

Dean C. Delis, Ph.D. is an author of the CVLT and receives royalties for the test. There are no other conflicts of interest to be declared by any authors. This work was supported, in part, by National Institutes of Health (NIH) grants R01 AG034202 to P. E. G., K24 AG026431 to M. W. B., and P50 AG05131 to the Shiley-Marcos ADRC, and a Huntington's Disease Society of America Center of Excellence grant to J. C. B.

Dissertation Author's Acknowledgements

I, Lisa Graves (primary investigator and author of this material), would like to thank all coauthors – Drs. Gilbert, Holden, Delano-Wood, Bondi, Salmon, Corey-Bloom, and Delis, as well as Ms. Emily Van Etten – for their contributions to this work. I also would like to thank the *Journal of the International Neuropsychological Society* for accepting this work for publication, and the Cambridge University Press for publishing this work.

Figures

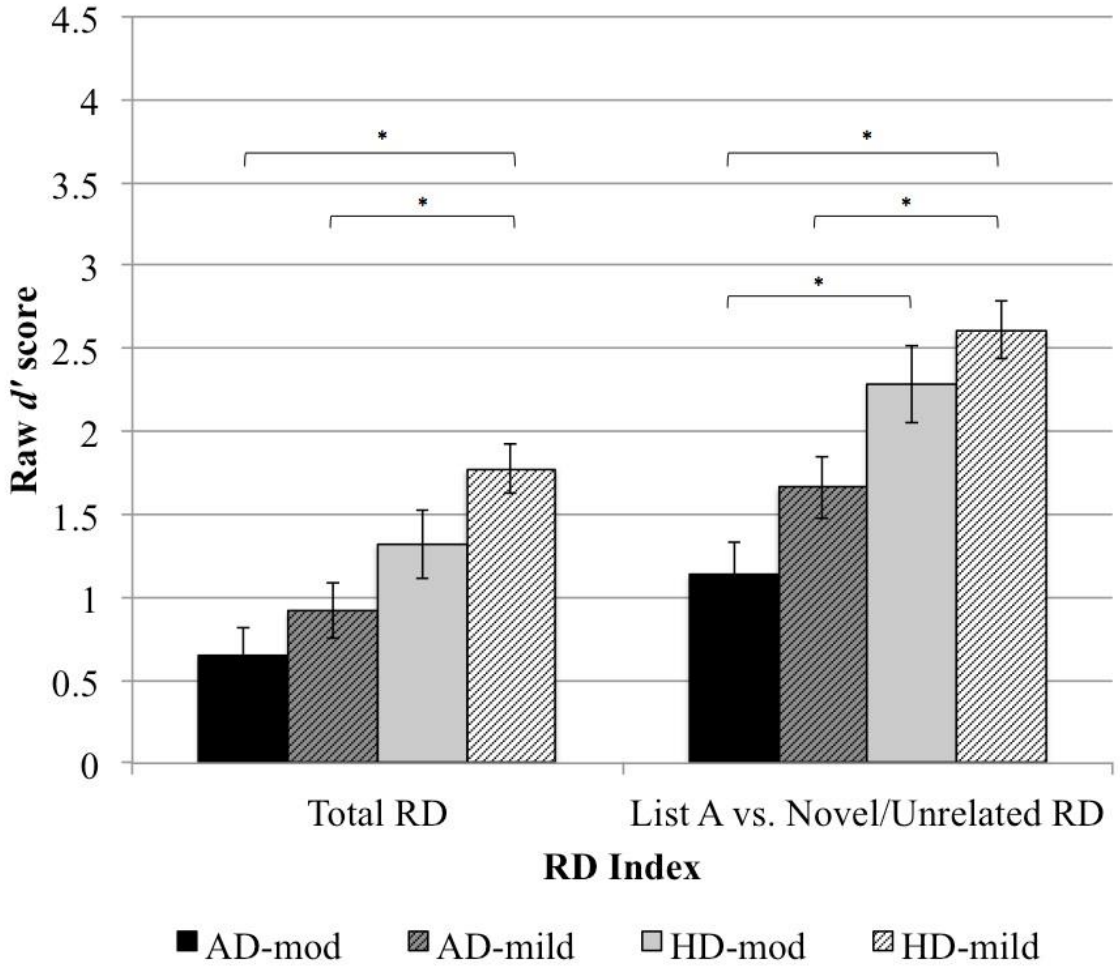

Figure 4.1. *Mean (standard error) raw scores on the Total RD and List A vs. Novel/Unrelated RD indices in the Alzheimer's disease-moderate (AD-mod), Alzheimer's disease-mild (AD-mild), Huntington's disease-moderate (HD-mod), and Huntington's disease-mild (HD-mild) groups.*

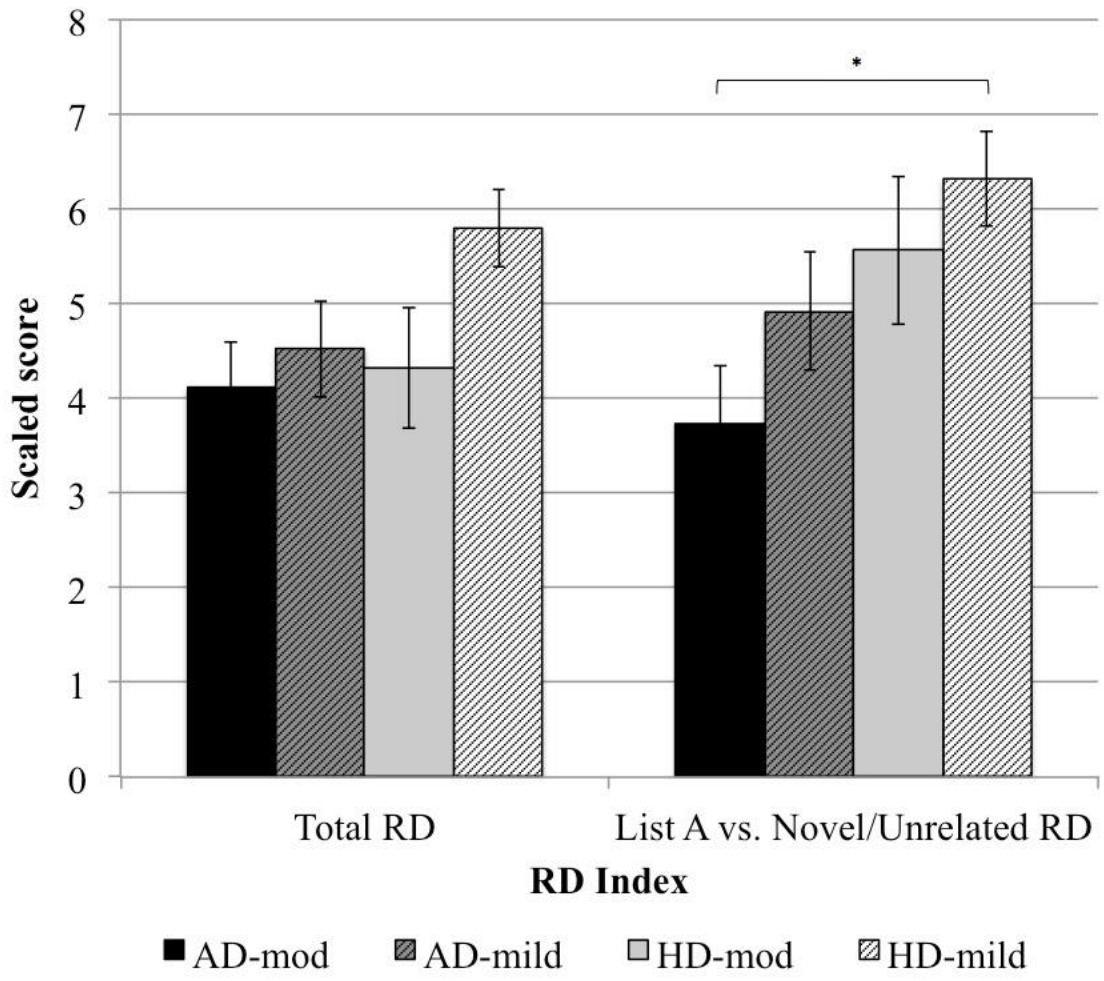

Figure 4.2. *Mean (standard error) scaled scores on the Total RD and List A vs. Novel/Unrelated RD indices in the Alzheimer's disease-moderate (AD-mod), Alzheimer's disease-mild (AD-mild), Huntington's disease-moderate (HD-mod), and Huntington's disease-mild (HD-mild) groups.*

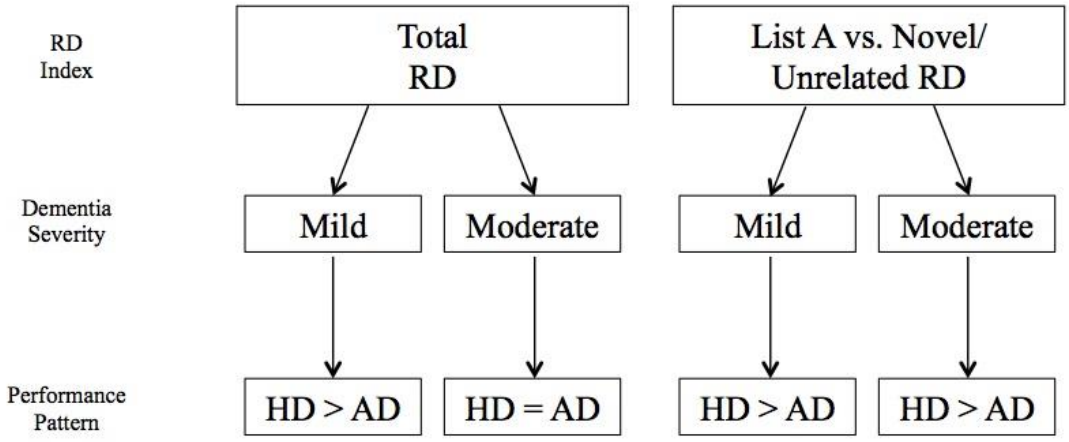

Figure 4.3. *Flowchart outlining the pattern of performance that may be expected with raw scores on the CVLT-3 Total RD and List A vs. Novel/Unrelated RD indices in Huntington's disease and Alzheimer's disease.*

Tables

Table 4.1. *Demographic information on participants in the Alzheimer's disease-moderate (AD-mod), Alzheimer's disease-mild (AD-mild), Huntington's disease-moderate (HD-mod), Huntington's disease-mild (HD-mild), healthy older adult (OA), and healthy middle-aged adult (MA) groups.*

Variable	AD-mod	AD-mild	HD-mod	HD-mild	OA	MA
n	27	25	16	39	53	31
% Female	33.33	40.00	56.25	66.67	47.17	58.06
Age	78.67 (5.02) 67-85	75.28 (4.84) 65-84	49.38 (11.93) 25-73	49.87 (11.67) 34-78	74.57 (6.38) 65-89	48.90 (4.53) 41-55
Education	15.22 (3.39) 6-20	15.44 (2.89) 9-20	14.13 (2.28) 12-20	14.18 (2.33) 8-20	16.30 (2.09) 12-20	14.19 (1.97) 10-18
DRS/DRS-2 Total	112.70 (4.11) 101-119	126.88 (3.79) 120-136	112.38 (6.47) 100-119	129.74 (4.02) 121-138	--	--

Note: For age, education, and DRS/DRS-2 Total variables, first row includes mean (standard deviation) values and second row includes range.

Table 4.2. *Mean (standard deviation) values for the Alzheimer's disease-moderate (AD-mod), Alzheimer's disease-mild (AD-mild), Huntington's disease-moderate (HD-mod), Huntington's disease-mild (HD-mild), healthy older adult (OA), and healthy middle-aged adult (MA) groups on raw and scaled scores on Total RD and List A vs. Novel/Unrelated RD indices, as well as RD index components.*

Variable	AD-mod	AD-mild	HD-mod	HD-mild	OA	MA
RD Indices						
Total RD (Raw)	0.64 (0.65)	0.92 (0.67)	1.30 (0.77)	1.77 (0.82)	3.07 (0.76)	3.21 (0.59)
List A vs. Novel/Unrelated RD (Raw)	1.28 (1.09)	1.78 (1.12)	2.14 (1.10)	2.47 (0.76)	3.50 (0.56)	3.55 (0.52)
Total RD (Scaled)	4.11 (2.21)	4.52 (2.20)	4.31 (2.09)	5.80 (2.57)	12.02 (3.01)	10.42 (2.31)
List A vs. Novel/Unrelated RD (Scaled)	3.74 (3.11)	4.92 (4.02)	5.56 (3.69)	6.31 (3.16)	11.83 (2.83)	11.10 (2.36)
RD Index Components						
Hits	11.11 (2.99)	11.64 (3.09)	10.31 (3.91)	12.41 (2.87)	14.74 (1.46)	14.94 (1.34)
List B FP Errors	8.63 (3.16)	7.88 (3.63)	4.56 (3.93)	3.44 (3.50)	1.68 (1.98)	0.94 (1.26)
Novel/Prototypical FP Errors	4.19 (1.88)	4.00 (2.52)	2.31 (1.85)	3.21 (2.22)	0.66 (1.41)	0.77 (1.06)
Novel/Unrelated FP Errors	2.78 (2.14)	1.80 (2.10)	0.69 (1.62)	0.67 (0.98)	0.08 (0.27)	0.10 (0.30)

Table 4.3. *p values associated with relevant pairwise comparisons on raw scores (controlling for age) and scaled scores on Total RD and List A vs. Novel/Unrelated RD indices. Cohen's d values associated with significant AD and HD group differences are reported.*

Comparison	Total RD (Raw)		List A vs. Novel/Unrelated RD (Raw)		Total RD (Scaled)		List A vs. Novel/Unrelated RD (Scaled)	
	p	*d*	*p*	*d*	*p*	*d*	*p*	*d*
AD-mod vs. HD-mod	.027	--	.001*	0.79	.802	--	.066	--
AD-mild vs. HD-mild	.001*	1.14	.001*	0.72	.051	--	.085	--
AD-mod vs. AD-mild	.176	--	.023	--	.562	--	.175	--
HD-mod vs. HD-mild	.035	--	.190	--	.050	--	.423	--
AD-mod vs. HD-mild	<.001*	1.53	<.001*	1.27	.009	--	.001*	0.82
AD-mild vs. HD-mod	.168	--	.062	--	.799	--	.521	--

**p* value retains signficance following Bonferroni adjustment for multiple comparisons

References

Alexander, G. E., Crutcher, M. D., & DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Progress in Brain Research, 85,* 119–146.

Baldo, J. V., Delis, D., Kramer, J., & Shimamura, A. P. (2002). Memory performance on the California Verbal Learning Test-II: Findings from patients with focal frontal lesions. *Journal of the International Neuropsychological Society, 8,* 539–546.

Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *Neuroscientist, 13*(3), 280–291.

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica, 82,* 239–259.

Budson, A. E., & Kowall, N. W. (2013). *Handbook of Alzheimer's disease and other dementias.* Hoboken, NJ: Wiley-Blackwell.

Butters, N., Wolfe, J., Granholm, E., & Martone, M. (1986). An assessment of verbal recall, recognition and fluency abilities in patients with Huntington's disease. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 22,* 11–32.

Butters, N., Wolfe, J., Martone, M., Granholm, E., & Cermak, L. S. (1985). Memory disorders associated with Huntington's disease: Verbal recall, verbal recognition, and procedural memory. *Neuropsychologia, 23,* 729–743.

Crosson, B., Benefield, H., Cato, M. A., Sadek, J. R., Moore, A. B., Wierenga, C. E., … & Briggs, R. W. (2003). Left and right basal ganglia and frontal activity during language generation: Contributions to lexical, semantic, and phonological processes. *Journal of the International Neuropsychological Society, 9,* 1061–1077.

Cummings, J. L. (1993). Frontal-subcortical circuits and human behavior. *Archives of Neurology, 50,* 873–880.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test.* San Antonio, TX: The Psychological Corporation.

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

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2017). *California Verbal Learning Test–3, Third Edition*. San Antonio, TX: The Psychological Corporation.

Delis, D. C., Massman, P. J., Butters, N., Salmon, D. P., Cermak, L. S., & Kramer, J. H. (1991). Profiles of demented and amnesic patients on the California Verbal Learning Test: Implications for the assessment of memory disorders. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 3,* 19–26.

Delis, D. C., Wetter, S. R., Jacobson, M. W., Peavy, G., Hamilton, J., Gongvatana, A., … & Salmon, D. P. (2005). Recall discriminability: Utility of a new CVLT-II measure in the differential diagnosis of dementia. *Journal of the International Neuropsychological Society, 11*(6), 708–715.

Fine, E. M., Delis, D. C., Wetter, S. R., Jacobson, M. W., Hamilton, J. M., Peavy, G., … & Salmon, D. P. (2008). Identifying the "source" of recognition memory deficits in patients with Huntington's disease or Alzheimer's disease: Evidence from the CVLT-II. *Journal of Clinical and Experimental Neuropsychology, 30*(4), 463–470.

Graves, L. V., Holden, H. M., Woods, S. P., Delano-Wood, L., Bondi, M., Salmon, D. P., … & Gilbert, P. E. (2017). Total recognition discriminability in Huntington's and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 39*(2), 120-130.

Hodges, J. R., Salmon, D. P., & Butters, N. (1990). Differential impairment of semantic and episodic memory in Alzheimer's and Huntington's diseases: A controlled prospective study. *Journal of Neurology, Neurosurgery, and Psychiatry, 53(12),* 1089–1095.

Huntington Study Group. (1996). Unified Huntington's disease rating scale: Reliability and consistency. *Movement Disorders, 11,* 136–142.

Hyman, B. T., Van Hoesen, G. W., Damasio, A. R., & Barnes, C. L. (1984). Alzheimer's disease: Cell-specific pathology isolates the hippocampal formation. *Science, 225,* 1168–1170.

Jurica, P. J., Leitten, S., Mattis, S. (2001). *Dementia Rating Scale-2: Professional manual.* Lutz, FL: Psychological Assessment Resources.

Kramer, J. H., Delis, D. C., Blusewicz, M. J., Brandt, J., Ober, B. A., & Strauss, M. (1988). Verbal memory errors in Alzheimer's and Huntington's dementias. *Developmental Neuropsychology, 4,* 1–15.

Kramer, J. H., Levin, B. E., Brandt, J., & Delis, D. C. (1989). Differentiation of Alzheimer's, Huntington's, and Parkinson's disease patients on the basis of verbal learning characteristics. *Neurology, 3,* 111–120.

Lundervold, A. J., Reinvang, I., & Lundervold, A. (1994). Characteristic patterns of verbal memory function in patients with Huntington's disease. *Scandanavian Journal of Psychology, 35(1),* 38–47.

Macmillan, N. A., & Creelman, D. C. (1991). *Detection theory: A user's guide.* New York: Cambridge University Press.

Martone, M., Butters, N., Payne, M., Becker, J. T., & Sax, D. S. (1984). Dissociations between skill learning and verbal recognition in amnesia and dementia. *Archives of Neurology, 41,* 965–970.

Massman, P. J., Delis, D. C., Butters, N., Levin, B. E., & Salmon, D. P. (1990). Are all subcortical dementias alike? Verbal learning and memory in Parkinson's and Huntington's disease patients. *Journal of Clinical and Experimental Neuropsychology, 12(5),* 729–744.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS–ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. Neurology, 34, 939–944.

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, Jr., C. R., Kawas, C. H., … & Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging and the Alzheimer's Association workgroup. *Alzheimer's & Dementia, 7*(3), 263–269.

Montoya, A., Pelletier, M., Menear, M., Duplessis, E., Richer, F., & Lepage, M. (2006). Episodic memory impairment in Huntington's disease: A meta-analysis. *Neuropsychologia, 44,* 1984–1994.

Moss, M. B., Albert, M. S., Butters, N., & Payne, M. (1986). Differential patterns of memory loss among patients with Alzheimer's disease, Huntington's disease, and alcoholic Korsakoff's syndrome. *Archives of Neurology, 43,* 239–246.

Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annual Review of Psychology, 60,* 257–282.

Salmon, D. P., & Filoteo, J. V. (2007). Neuropsychology of cortical versus subcortical dementia syndromes. *Seminars in Neurology, 27,* 7–21.

Troster, A. I., Butters, N., Salmon, D. P., Cullum, C. M., Jacobs, D., Brandt, J., … & White, R. R. (1993). The diagnostic utility of savings scores: Differentiating Alzheimer's and Huntington's diseases with the logical memory and visual reproduction tests. *Journal of Clinical and Experimental Neuropsychology, 15,* 773–788.

Tulving, E., Kapur, S., Craik, F. I., Moscovitch, M., & Houle, S. (1994). Hemispheric encoding/retrieval asymmetry in episodic memory: Positron emission tomography findings. *Proceedings of the National Academy of Sciences, USA, 91,* 2016–2020.

Vonsattel, J. P. (2000). Neuropathology of Huntington's disease. *NeuroScience News, 3,* 45–53.

Vonsattel, J., Myers, R., Stevens, T. J., Ferrante, R. J., Bird, E.D., & Richardson, E.P. (1985). Neuropathological classification of Huntington's disease. *Journal of Neuropathology and Experimental Neurology, 44,* 559–577.

CHAPTER 5:

Integrated Summary

With the expanding older population in the U.S., the burden of cognitive decline in older age and the prevalence of dementia due to neurodegenerative disease are expected to increase. Given that memory loss is a hallmark feature of cognitive impairment in both healthy aging and dementia, more refined measures of memory are needed to further elucidate the nature of memory changes that may accompany normal aging and neurodegenerative processes. Collectively, the three studies that were proposed and completed in this staple dissertation project utilized innovative psychometric approaches to explore age-related differences on more nuanced aspects of yes/no recognition memory, as well as enhance the characterization of and differentiation between profiles of yes/no recognition memory impairment in AD and HD.

Study 1 examined whether AD and HD differences on Total RD varied across applications of nonparametric and parametric formulas for calculating Total RD, and included an emphasis on exploring the extent to which each formula captures high FP error rates. Key findings indicated that relative to the AD group, the HD group exhibited comparable standardized parametric Total RD scores (despite higher raw nonparametric and parametric Total RD scores), whereas the previous CVLT literature had shown that standardized Total RD scores were higher in HD than in AD. Notably, analyses revealed that FP error rates were more strongly correlated with raw nonparametric Total RD scores than with raw parametric Total RD scores in the AD group, although this was not observed in the HD group. These observations suggested that the nonparametric Total RD formula may more fully capture the contribution of FP errors to a Total RD score

and, as a result, provide important information regarding an examinee's recognition memory that may otherwise be lost in the application of the parametric Total RD formula and subsequent standardization of parametric Total RD scores. Thus, standardized parametric Total RD scores may be somewhat overestimated (i.e., the impact of high FP error rates on Total RD scores may be inadvertently reduced) in individuals with AD, which may have led in part to the observation that the HD and AD groups performed comparably on this index. Furthermore, these findings highlighted the importance and utility of examining other yes/no recognition memory indices (e.g., FP error rates, nonparametric Total RD) in addition to standardized parametric Total RD scores when using the CVLT-II to assess and characterize yes/no recognition memory.

Given that Study 1 findings highlighted the important role of FP errors in the assessment of RD, *Study 2* investigated the utility of refined RD indices that correspond to the four FP error subtypes on the CVLT-II Yes/No Recognition trial (prototypical List B, unrelated List B, prototypical novel, and unrelated novel) in characterizing nuanced aspects of yes/no recognition memory in healthy older and young adults. An emphasis was made on the utility of Source and Novel RD indices that isolated the ability to distinguish List A targets from List B and novel distractors that do not share strong semantic associations with targets, relative to original indices that incorporate FP errors associated with both prototypical and semantically-unrelated distractors, in distinguishing healthy older and young adults. Although older adults performed worse than young adults on all RD indices, key findings indicated that age group differences were smaller in magnitude on refined Source and Novel RD indices that exclude FP errors associated with semantically-related distractors relative to indices that include FP errors associated

with both prototypical and semantically-unrelated distractors. Moreover, older adults, but not young adults, performed disproportionately worse on Source RD than on Novel RD in the context of refined indices, providing further evidence for disproportionate age-related decline in source memory relative to item memory. Study 2 findings suggested that although CVLT-II indices of Source and Novel RD in their current form are useful in characterizing age-related differences on aspects of yes/no recognition memory that pertain to source and item memory, the more refined indices may further elucidate the degree of between- and within-group differences on these constructs. Overall, Study 2 findings highlight the utility of refined RD indices in characterizing yes/no recognition memory changes associated with healthy aging.

The discrepancy of findings from Study 1 and those previously reported in the CVLT and CVLT-II literature on yes/no recognition memory in AD and HD (e.g., Fine et al., 2008) highlighted the need for more refined RD indices to enhance the characterization of and differentiation between profiles of memory loss in AD and HD. Whereas the Total RD index incorporates FP errors associated with all distractor types (including List B and semantically-related items), the new List A vs. Novel/Unrelated RD index on the CVLT-3 incorporates only FP errors associated with novel, semantically-unrelated distractors. Thus, in minimizing levels of source and semantic interference, List A vs. Novel/Unrelated RD may yield a purer assessment of yes/no recognition memory that is less subject to source memory difficulties or semantic confusion, both of which are often seen in individuals with primarily frontal-system dysfunction (e.g., early HD). *Study 3* sought to elucidate the nature and extent of AD and HD differences in yes/no recognition memory by examining performances in individuals with AD and those with

HD on CVLT-3 indices of Total RD and List A vs. Novel/Unrelated RD. Although the AD and HD groups were impaired on both RD indices relative to healthy comparison groups, key findings indicated that those with HD generally outperformed those with AD, and group differences were more robust on the new List A vs. Novel/Unrelated RD index than on the Total RD index. Thus, Study 3 findings highlighted that the new CVLT-3 List A vs. Novel/Unrelated RD index (a) maximally assesses yes/no recognition memory independent of source and semantic interference; and (b) provides greater differentiation between individuals whose memory disorder is primarily at the encoding/storage level (e.g., as in AD) versus at the retrieval level (e.g., as in early HD).

Collectively, findings from the three studies that were proposed and completed in this staple dissertation project highlighted the important role of FP errors in the assessment of yes/no recognition memory, and the utility of more refined RD indices derived from the CVLT in elucidating the nature of yes/no recognition memory changes associated with healthy aging, and dementia in AD and HD. Furthermore, the present findings may inform future research on the utility of refined RD indices in characterizing yes/no recognition memory changes associated with preclinical stages of neurodegenerative disease, as well as with other neurological conditions involving cognitive impairment.

APPENDICES

The following appendices contain supplementary data that were associated with the aims addressed in Studies 1, 2, and 3, but were not included in the final published manuscripts associated with these studies due to efforts to maintain brevity in the production of the individual manuscripts. It is anticipated that these data may be incorporated into future manuscripts.

APPENDIX A:

List A vs. List B Recognition Discriminability (RD) in Huntington's and Alzheimer's Disease

Preliminary Analyses

Preliminary ANCOVA tests revealed that age, $F(1, 105) = 21.85$, $p < .001$, and DRS/DRS-2 scores, $F(1, 105) = 30.67$, $p < .001$, were significant predictors of raw scores on the List A vs. List B RD index. However, given that scaled scores correct for age, age was not included as a covariate in primary analyses involving scaled scores. Moreover, given that DRS/DRS-2 scores were systematically varied by group (i.e., individuals with HD and AD were characterized as mild or moderate in dementia severity), DRS/DRS-2 scores were not controlled for in primary analyses involving scaled scores. Gender and education were not shown to be significant predictors of raw scores on the List A vs. List B RD index (*p*s > .05).

Primary Analyses: HD and AD Performance on the List A vs. List B RD Index

ANOVA tests revealed a significant effect of group on scaled scores on the List A vs. List B RD index, $F(3, 103) = 4.87$, $p < .01$. Post-hoc pairwise comparisons revealed that the HD-mild group outperformed all other HD and AD subgroups (*p*s < .01). Thus, the HD-mild group ($M = 6.31$, $SD = 2.75$; $p < .01$) performed significantly better than the HD-mod group ($M = 4.44$, $SD = 2.25$), whereas the AD-mild ($M = 4.64$, $SD = 2.02$) and AD-mod ($M = 4.41$, $SD = 2.10$) groups exhibited comparable performance ($p > .05$). In addition, the HD-mild group performed significantly better than the AD-mild group ($p < .01$), whereas the HD-mod and AD-mod groups exhibited comparable performance ($p > .05$). Although the comparison of the HD-mild and AD-mod groups was not of primary

interest, post-hoc tests revealed that the former performed significantly better than the latter ($p < .01$). Group differences on scaled scores on the List A vs. List B RD index are illustrated in Figure A.

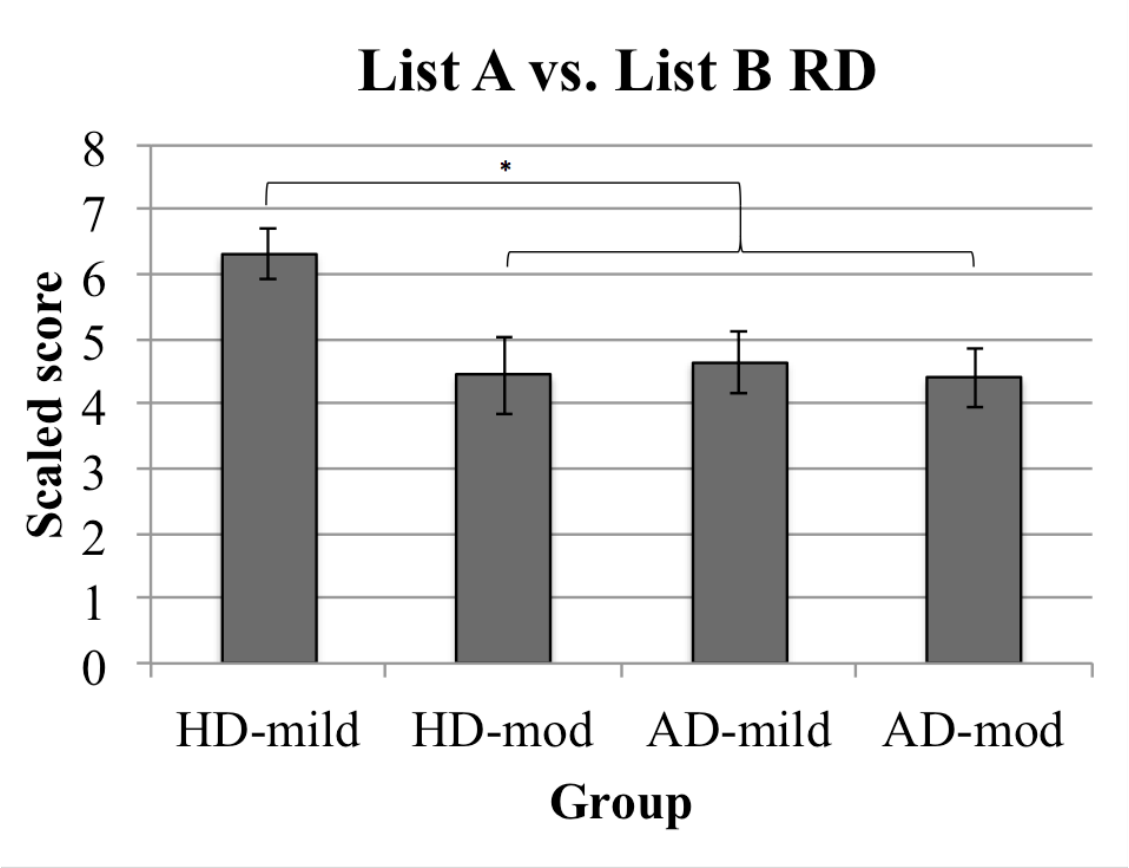

Figure A. *Mean (standard error) scaled scores on the List A vs. List B RD index in the Huntington's disease-mild (HD-mild), Huntington's disease-moderate (HD-mod), Alzheimer's disease-mild (AD-mild), and Alzheimer's disease-moderate (AD-mod) groups.*

APPENDIX B:

Examination of False Positive (FP) Error Subtypes and Corresponding Source and Novel Recognition Discriminability (RD) Indices in Alzheimer's and Huntington's Disease

Table B.1. *Mean (standard deviation) values for the Alzheimer's disease-moderate (AD-mod), Alzheimer's disease-mild (AD-mild), Huntington's disease-moderate (HD-mod), Huntington's disease-mild (HD-mild), healthy older adult (OA), and healthy middle-aged adult (MA) groups on raw scores on Yes/No Recognition variables.*

Variable	AD-mod	AD-mild	HD-mod	HD-mild	OA	MA
List A vs. List B RD	0.64 (0.58)	0.85 (0.58)	1.14 (0.86)	1.82 (0.90)	2.94 (0.71)	3.17 (0.63)
List A vs. Novel/Prototypical RD	0.54 (0.79)	0.71 (1.06)	1.05 (0.85)	1.25 (0.96)	3.15 (0.89)	3.07 (0.75)
List A vs. Novel/Unrelated RD	1.28 (1.09)	1.78 (1.12)	2.19 (1.10)	2.45 (0.76)	3.50 (0.56)	3.55 (0.52)
List B Shared FP	4.52 (1.97)	4.36 (1.87)	2.75 (2.08)	1.79 (1.99)	0.85 (1.17)	0.65 (1.08)
List B Nonshared FP	4.11 (2.06)	3.52 (2.52)	2.00 (2.19)	1.56 (1.79)	0.83 (1.14)	0.29 (0.69)
Novel/Prototypical FP	4.19 (1.88)	4.00 (2.52)	2.56 (1.79)	3.10 (2.27)	0.66 (1.41)	0.77 (1.06)
Novel/Unrelated FP	2.78 (2.14)	1.80 (2.10)	0.69 (1.62)	0.67 (0.98)	0.08 (0.27)	0.10 (0.30)
Total FP	15.59 (6.17)	13.68 (6.81)	8.00 (5.93)	7.13 (6.17)	2.42 (2.87)	1.81 (1.99)
Total Hits	11.11 (2.99)	11.64 (3.09)	10.63 (3.74)	12.28 (3.04)	14.74 (1.46)	14.94 (1.34)

Table B.2. *p values associated with group differences on raw scores on Yes/No Recognition variables in the context of significant group x RD index, and group x FP error subtype, interactions in analysis of covariance (ANCOVA) tests.*

Pairwise Comparison	List A vs. List B RD	List A vs. Novel/ Prototypical RD	List A vs. Novel/ Unrelated RD	List B Shared FP	List B Nonshared FP	Novel/ Prototypical FP	Novel/ Unrelated FP	Total Hits	Total FP
AD-mod vs. HD-mod	.094	.056	<.01*	<.01*	<.05	<.05	<.001*	.909	<.001*
AD-mild vs. HD-mild	<.001*	<.05	<.01*	<.001*	<.05	.097	<.01*	.280	<.001*
AD-mod vs. AD-mild	.292	.427	<.05	.683	.260	.684	<.01	.421	.164
HD-mod vs. HD-mild	<.01*	.459	.281	.053	.397	.327	.947	<.05	.556
AD-mod vs. HD-mild	<.001*	<.01	<.001*	<.001*	<.01*	.060	<.001*	.104	<.001*
AD-mild vs. HD-mod	.314	.151	<.05	<.01	.100	<.05	<.01	.495	<.01*
AD-mod vs. OA	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*
AD-mild vs. OA	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*	<.001*
HD-mod vs. MA	<.001*	<.001*	<.001*	<.001*	<.01*	<.01*	.130	<.001*	<.001*
HD-mild vs. MA	<.001*	<.001*	<.001*	<.01*	<.01*	<.001*	.064	<.001*	<.001*

**p* value retains significance following Bonferroni adjustment

Table B.3. *p values for within-group comparisons on raw scores on Yes/No Recognition variables in the Alzheimer's disease-moderate (AD-mod), Alzheimer's disease-mild (AD-mild), Huntington's disease-moderate (HD-mod), and Huntington's disease-mild (HD-mild) groups.*

Pairwise Comparison	AD-mod	AD-mild	HD-mod	HD-mild
List A vs. List B RD - List A vs. Novel/Prototyical RD	.207	.153	.991	<.01*
List A vs. List B RD - List A vs. Novel/Unrelated RD	<.01*	<.001*	<.001*	<.001*
List A vs. Novel/Prototypial RD - List A vs. Novel/Unrelated RD	<.01*	<.001*	<.001*	<.001*
List B Shared FP - List B Nonshared FP	.065	<.01*	.411	.740
List B Shared FP - Novel/Prototypical FP	.342	.293	.718	<.001*
List B Shared FP - Novel/Unrelated FP	<.001*	<.001*	<.001*	<.001*
List B Nonshared FP - Novel/Prototypical FP	.406	.118	.677	<.01*
List B Nonshared FP - Novel/Unrelated FP	<.05	<.001*	<.001*	<.001*
Novel/Prototypical FP - Novel/Unrelated FP	<.01*	<.001*	<.001*	<.001*

*p value retains significance following Bonferroni adjustment