APPLIED NEUROPSYCHOLOGY: ADULT
http://dx.doi.org/10.1080/23279095.2017.1293671

Routledge
Taylor & Francis Group

Impact of rapport on neuropsychological test performance

Michael D. Barnett, Thomas D. Parsons, Brooke L. Reynolds and Lee A. Bedford

Department of Psychology, University of North Texas, Denton, Texas, USA

ABSTRACT
Guides to neuropsychological assessment emphasize the importance of establishing rapport; however, there has been a minimal amount of empirical investigation of the impact of rapport on neuropsychological test performance. In this experiment, participants ($N = 98$) were randomly assigned to take neuropsychological tests in either a high or low rapport condition. Results showed that we were able to manipulate the level of rapport and that the level of rapport had a significant effect on the Grooved Pegboard Test and the Controlled Oral Word Association Test, with other tests nearing statistical significance. These results suggest that the level of rapport may affect neuropsychological test performance.

KEYWORDS
Neuropsychological assessment; neuropsychological training; rapport; social facilitation

Guides to neuropsychological assessment (Rath et al., 2004; Spreen & Strauss, 1998) as well as manuals for specific psychological and neuropsychological tests (Aiken, 1999; Delis, Kaplan, & Kramer, 2001; Kløve, 1963; Schrank, Mather, & McGrew, 2014; Wechsler, 2008; Wilkonsen & Robertson, 2009) typically emphasize the importance of establishing rapport, with the implication that failure to establish rapport could result in reduced effort from the examinee and thereby affect performance. Accurate assessments are integral for correct diagnosis and application of appropriate treatment (Millis & Volinsky, 2001). Additionally, accurate neuropsychological testing outcomes are predicated on patient participation in a comfortable environment (Gorske, 2007; Harvey et al., 1988; Skuy, Schutte, Fridjhon, & O'Carroll, 2001). There are individual differences among examiners in their ability to establish rapport (Allen, Montgomery, Tubman, Frazier, & Escovar, 2003), which may affect psychological assessment outcomes. The purpose of this exploratory study was to investigate the impact of rapport on neuropsychological test performance in a sample of U.S. college students.

Rapport and the social brain

The human brain is highly attuned to social information (Cacioppo, 2002; Cacioppo & Berntson, 2004; Cozolino, 2014; Easton & Emery, 2005; Geary, 2005; Harmon-Jones & Winkielman, 2007). Historically, neuropsychological testing has implicitly assumed that the examinee responds to test stimuli exclusive of other stimuli, although a large body of research has found that environmental conditions impact neuropsychological test performance (Horowitz & McCaffrey, 2008; Huguet, Galvaing, Monteil, & Dumas, 1999; McCaffrey, Lynch, & Yantz, 2005).

Social facilitation refers to the tendency for an animal to change its behavior due to the presence of another animal (Thomas, Skitka, Christen, & Jurgena, 2002; Uziel, 2007; Yantz & McCaffrey, 2005). Social facilitation in humans has been shown to impact performance negatively when one is problem solving or performing novel tasks and positively when the task has higher habit strength according to social psychological research (Platania & Moran, 2001; Thomas et al., 2002; Uziel, 2007; Yantz & McCaffrey, 2005). Specifically in the area of neuropsychological testing, social facilitation has been found to affect performance on the Stroop task, digit span, the Paced Auditory Serial Addition Test, and the Controlled Oral Word Association Test (Horowitz & McCaffrey, 2008; Huguet et al., 1999; McCaffrey et al., 2005), with memory assessments showing a particular vulnerability (Constantinou, Ashendorf, & McCaffrey, 2005; Gavett, Lynch, & McCaffrey, 2005; Gavett & McCaffrey, 2007). While the mere presence of another individual is a minimally social interaction, this condition is enough to cause changes in behavioral outcomes (Kushnir & Duncan, 1978; Thomas et al., 2002; Zajonc, 1965). With the controlled manipulation of rapport being such an unexplored area of research, it is difficult to gauge the impact rapport has on neuropsychological test performance. In this study, we sought

CONTACT Michael D. Barnett Michael.Barnett@unt.edu Department of Psychology, University of North Texas, 1155 Union Circle #311280, Denton, TX 76203, USA.

to explore the extent to which social stimuli—that is, high and low rapport—might affect the neuropsychological test performance.

Rapport and neuropsychological testing

Rapport refers to the therapeutic alliance that exists between a therapist/examiner and the patient/examinee. Factors such as the warmth and perceived supportiveness of the therapist/examiner are associated with higher levels of rapport (Ackerman & Hilsenroth, 2003). High rapport has been associated with improved therapeutic outcomes (Hilsenroth, Peter, & Ackerman, 2004; Kline, 2013). However, there is a lack of research investigating the effects of rapport on neuropsychological testing. While a minimal amount of research has investigated the relationship between rapport and results of neuropsychological testing, rapport has been studied in other contexts.

In test administration (and/or educational settings), the perceived warmth or aloofness of an evaluator/examiner has been shown to impact task engagement (Downer, Rimm-Kaufman, & Pianta, 2007). Likewise, in counseling, rapport is often referred to as the therapeutic alliance and is documented as an important factor in achieving favorable outcomes (Bordin, 1979; Meyerbroker & Emmelkamp, 2008; Sharpley, Fairnie, Tabary-Collins, Bates, & Lee, 2000; Sharpley, Jeffrey, & Mcmah, 2006) as well as accurate clinical assessments (Patterson, Anderson, & Wei, 2014). This concept indirectly supports the need for rapport between examiners and examinees in neuropsychological assessments (Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000; Roberts & Penn, 2009). Because of the establishment of a therapeutic alliance, patients are more likely to tolerate more involved and time-consuming assessment (Hilsenroth, Kivlinghan, & Slavin-Mulford, 2015). Psychologists are trained in specific skills to facilitate rapport building (Joe, Simpson, Dansereau, & Rowan-Szal, 2001; Sharpley & McNally, 1997). Although assessments of rapport using therapeutic alliance measures work well in counseling and clinical settings, they are not easily applied to research and testing. These measures assess rapport after multiple therapy sessions. With this study, we were able to identify whether the initial level of rapport a participant establishes with the examiner after only a few minutes impacted performance on neuropsychological tests.

Current study

The present study addresses the effect of rapport on several neuropsychological testing outcomes. The examiners who administered the tests in the low rapport condition were cold and withdrawn, while the examiners in the high rapport condition were friendly and supportive (see the following section). With no extant research investigating the impact of rapport on neuropsychological test performance, this study was exploratory in nature. It was hypothesized that individuals in the low rapport condition would perform worse across neuropsychological domains.

Method

Participants

Participants consisted of undergraduate students enrolled in a psychology course ($N = 98$) at a large public university in the southern United States. Participants were recruited from the department research website and received course credit upon completion on the study. Participants were screened to ensure that they were neurotypical (i.e., did not have a history of neurological illness, psychiatric, or physical disorder that would impair test performance), and all participants were neurotypical; however, two participants were excluded from the Stroop Test due to color-blindness. Demographic characteristics of the participants are displayed in Table 1.

Examiners

Examiners ($n = 15$) were all undergraduate students between the age of 18 and 29. Other examiner demographics are listed in Table 2.

Measures

We selected measures with the following aims in mind. First, given the exploratory nature of this experiment,

Table 1. Demographics ($N = 98$).

	n	%
Gender		
Male	40	40.4
Female	58	58.6
Age		
18-19	25	25.3
20-21	32	32.3
22-23	16	16.1
24-25	8	8.1
26-27	6	6.0
28-29	6	6.0
30+	5	5.0
Ethnicity		
White/Caucasian	44	44.4
Black/African-American	16	16.2
Hispanic	16	16.2
Asian/Pacific Islander	4	4.0
Native American	4	4.0
Other	7	7.1
Not Reported	7	7.1

Table 2. Examiner demographics ($N = 15$).

	n	%
Gender		
Male	4	26.7
Female	11	73.3
Ethnicity		
White/Caucasian	8	53.3
Black/African American	5	33.3
Asian/Pacific Islander	1	6.7
Hispanic	1	6.7
Age		
18–24	15	100.0

we wanted to test a variety of neuropsychological domains. Second, we wanted measures that were easy to administer and score. Third, we wanted the battery to be fairly brief so as to avoid the effects of fatigue. Finally, we included measures that previous research has shown may be impacted by social facilitation (e.g., the COWAT; Horowitz & McCaffrey, 2008) as well as measures of domains such as memory (e.g., the RAVLT) that have shown vulnerability to social facilitation (Constantinou et al., 2005; Gavett et al., 2005; Gavett & McCaffrey, 2007).

The *Controlled Oral Word Association Test* (COWAT; Benton, Sivan, Hamsher, Varney, & Spreen, 1983): is a measure of verbal fluency designed to assess the ability of the participant to spontaneously produce words within a certain amount of time. The test asks participants to verbally list as many words as possible that begin with "F," "A," and "S," as well as the category "Animals," excluding proper nouns and the same root word with different endings (e.g., fast, faster). The examiner counted the number of correct responses for each condition as well as a total sum.

The *Stroop Color and Word Test* (Golden, 1978) is used to evaluate an individual's ability to inhibit an over-learned behavior or response. Colorblind participants were excluded from this test ($n = 2$). The test contained three conditions; in the first condition participants read the words "red," "green," and "blue" as quickly as possible without skipping or making any mistakes. In the second condition, participants identified the ink color in which a string of Xs is printed in. The third condition had participants name the color of the ink, not read the name of the color (e.g., the word "green" is in blue ink). Participants had 45 seconds to complete as many items as possible. The examiner counted the number of correct responses as well as the number of errors.

The *Grooved Pegboard Test* (GPT; Kløve, 1963) is an instrument designed to assess motor speed, manual dexterity, and hand-eye coordination. The GPT consists of small key-holes arranged five across by five down for a total of 25 key-holes. There are small key-shaped pegs that must be inserted in each hole in a particular way. In the first condition, the participants were instructed to use their dominant hand to insert a peg into each hole, moving horizontally across the pegboard, without skipping a hole and only picking up one peg at a time. The examiner recorded both the completion time and the number of times that the participants dropped a peg. The second condition was similar, but with the participants using their nondominant hand.

The *Rey Auditory Verbal Learning Test* (RAVLT; Rey, 1964) is a test designed to assess verbal memory. The participants were read a list of fifteen words and asked to recall them immediately. There were five trials administered with the same list of words (RAVLT 1-5). Next, a list of fifteen new words was read to the participants and they were asked to recall them back to the examiner (RAVLT B1). Then, the participants were asked to recall the original list of fifteen words without including any of the new words (RAVLT A6). After a 20-minute delay, the participant was asked to recall the original fifteen (RAVLT A7). Finally, the examiner read a list of 50 words that included the original fifteen words and the participants were instructed to respond "yes" when they recognized a word from the original list. After each trial, the examiner totaled the correct number of words recalled.

Manipulation Check: After the neuropsychological tests were completed, the participants were instructed that they would now evaluate the examiner. Participants were given the manipulation check and an envelope. The examiner instructed the participant to complete the manipulation check as honestly as possible and then seal it in the provided envelope. The examiner assured the participant that the evaluation was anonymous and confidential and that the examiner would not see the evaluation. The manipulation check, which was created for this experiment, asked the participant to indicate, on a 5-point Likert-like scale, in which $1 =$ *strongly disagree* and $5 =$ *strongly agree*, to what extent 27 different adjectives described the examiner. Questions were both positively and negatively worded (e.g., nice, competent, strict, mean), and negatively worded items were reverse scored. The rapport of the examiner was determined by a sum score from the manipulation check, with higher scores equating to a warmer, more positive examiner. The measure was found to have a high degree of internal consistency (Cronbach's $\alpha = .94$). The manipulation check can be found in Appendix 1.

Procedure

This study was approved by the university committee for the protection of human subjects. Informed consent was obtained from all participants. Participants were

rewarded with course credit after completing the study. Upon signing up for the study, participants were randomly assigned to one of two conditions: a high rapport or a low rapport condition. Both conditions were scripted, and examiners were trained in and performed in both conditions. We made the deliberate choice not to blind examiners to the nature of the study. Although not blinding the examiners may have introduced examiner effects (DiPietro & Larson, 1989; Lasky, Felice, Moyer, Buddington, & Elliot, 1973), we felt that there were three primary advantages in having the examiners aware of the two conditions. First, we were unsure of whether an attempt to fully blind the examiners would be successful over time (Fleiss, Slakter, Fischman, Park, & Chilton, 1979; Shrout & Fleiss, 1979). Second, without the knowledge that they were supposed to establish rapport in one condition and not establish rapport in the other, we believed that examiners in a blinded low rapport condition would naturally begin to attempt to establish rapport. Third, and related to the previous point, we believed that examiners would best be able to offer the assessments in a high or low rapport condition if they were trained in and administered both as this would then allow them to make contrasts between one condition and the other. In other words, we believed that making the nature of the experiment explicit to the examiners was essential to establishing the experimental manipulation; establishing rapport, or lack thereof, is different than providing a placebo since the interaction itself is social and more complex. Each participant completed the battery of neuropsychological tests in only one of the conditions, and the tests were administered in the same order in both conditions.

High rapport vs. low rapport

In the high rapport condition, the examiner would enthusiastically greet the participant with a handshake, introduce themselves, and ask, "How is your semester going?" and engage in a moment or two of small talk. The examiner was warm and upbeat. After general directions, the examiner would state, "Just relax and try your best." Throughout the experiment the examiner was directed to make eye contact and respond with a smile and say, "Good job" or "Right." In contrast, for those in the low rapport condition, the examiner excluded all the social aspects and simply acknowledged the participant, confirmed that they were there to participate in the study, and did not shake the participant's hand or engage in any small talk. The examiner was cold and distant. In this condition, the examiner made minimal eye contact and if asked any questions responded, "Please continue" or, "Please complete the test to the best of your ability." The high rapport condition was made to feel like a warm and positive human interaction while the low rapport condition felt like a cold and clinical evaluative experience.

Results

Manipulation check

Participants in the high rapport condition rated their examiner as being more positive than those in the low rapport condition, $t(87) = 10.58$, $p < .001$, $d = 2.27$. This indicated rapport was successfully manipulated between the conditions.

Examiner effects

Rapport conditions did not differ significantly by examiner gender, $\chi^2(1) = .004$, $p = .94$, or examiner ethnicity, $\chi^2(3) = 4.68$, $p = .19$.

Differences in neuropsychological test performance by rapport condition

Raw test scores were used in analyses without converting to demographically adjusted norms or standardized scores. We decided to use raw scores because of the homogeneous nature of the population we were sampling: neurotypical college students, with the majority being between the ages of 18 and 25.

To assess whether neuropsychological test performance was affected by the level of rapport, an initial multivariate analysis of variance (MANOVA) was conducted for high rapport versus low rapport for the total scores of the COWAT battery, the RAVLT battery, and the individual sets of scores for both the grooved pegboard tests and the Stroop Color and Word Tests, Wilks' $\lambda = .80$, $F(12, 76) = 1.58$, $p = .11$. Although the overall MANOVA was not statistically significant, the results revealed that 20% of the variances in the composite dependent variable were explained by group differences. Thus, a series of one-way analyses of variance (ANOVAs) were conducted to explore differences between the high and low rapport conditions.

The letter "F" analysis from the COWAT ($F(1, 95) = 3.96$, $p = .049$, $\eta^2 = .04$) and the Grooved Pegboard Test with the dominant hand ($F(1, 94) = 8.79$, $p = .004$, $\eta^2 = .09$) revealed statistically significant differences between the high and low rapport conditions, while several other tests had moderate effect sizes that were approaching statistical significance. Specifically, participants in the high rapport condition scored significantly higher, revealing medium-large effect sizes, on the "F" section from the COWAT and on the

Table 3. Descriptive statistics and assessment scores, high and low rapport groups.

	High Rapport			Low Rapport					
	n	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	η^2
COWAT F	45	14.51	3.22	53	13.04	3.96	3.96	.05	**.04**
COWAT A	45	11.87	3.81	53	11.38	3.41	0.45	.50	.00
COWAT S	45	14.87	3.85	53	14.15	3.85	0.84	.36	.01
COWAT Animals	45	21.29	4.60	53	20.11	4.62	1.59	.21	**.02**
COWAT FAS Total	45	41.24	9.46	53	38.53	8.91	2.14	.15	**.02**
COWAT Total	45	62.53	1.84	53	58.70	1.56	2.57	.11	**.03**
Rey Auditory Verbal Learning Test (RAVLT)	45	7.38	2.15	53	7.58	2.22	0.22	.64	.00
RAVLT 2	45	10.31	2.13	53	10.36	2.50	0.01	.92	.00
RAVLT 3	45	11.69	2.10	53	11.89	2.08	0.22	.64	.00
RAVLT 4	45	12.71	1.88	53	12.72	1.94	0.00	.99	.00
RAVLT 5	45	13.13	1.80	53	12.94	1.57	0.31	.58	.00
RAVLT B1	45	7.02	2.68	51	7.31	2.11	0.35	.55	.00
RAVLT A6	44	11.59	2.25	52	12.29	2.17	2.38	.13	**.03**
RAVLT A7	45	11.69	2.16	52	11.96	2.05	0.41	.53	.00
Grooved Pegboard (GP) dominant hand time	45	61.41	9.16	51	67.25	9.65	8.79*	.01	**.09**
GP dominant hand drops	45	1.00	1.15	51	1.50	1.48	3.24	.08	**.03**
GP non dominant hand time	45	70.77	11.71	51	72.30	11.19	0.43	.52	.01
GP non dominant hand drops	45	1.33	1.43	51	1.67	1.53	1.20	.28	.01
Stroop 1 Raw	42	104.45	12.53	49	98.18	21.00	2.87	.09	**.03**
Stroop 1 Errors	42	0.40	0.73	49	0.37	0.88	0.05	.83	.00
Stroop 2 Raw	41	76.83	11.63	49	77.78	19.81	0.07	.79	.00
Stroop 2 Errors	41	1.34	1.54	49	1.00	0.98	1.62	.21	**.02**
Stroop 3 Raw	41	51.95	12.59	49	51.88	21.56	0.00	.99	.00
Stroop 3 Errors	41	1.10	1.38	49	1.31	1.31	0.54	.46	.01

Notes. Effect sizes greater than or equal to .02 are bolded. Two participants were excluded from the Stroop due to color blindness. All other variance in sample sizes are due to examiner errors.

Grooved Pegboard Test when using their dominant hand than participants in the low rapport group. All descriptive statistics and significance tests are displayed in Table 3.

Discussion

The manipulation check indicated that we were able to manipulate the level of rapport in an observable, measureable manner. Results of this investigation suggest that the level of rapport established between the examiner and the participant may affect the outcome of some neuropsychological tests. There was a significant effect in testing outcomes found between high rapport and low rapport conditions on the GPT with the dominant hand and on the initial trial of the COWAT.

It is possible that affect is the mechanism by which rapport impacts performance on certain neuropsychological tests. Positive and negative affect impact cognition in numerous ways (for review, see Lench, Flores, & Bench, 2011). Bennett and Hacker (2005) assert that short term affective disturbances can temporarily inhibit motivated action. This may explain the initial deficit in testing outcomes for the GPT and COWAT and the rebound in testing outcomes as the tests continue. It is also possible that differences in rapport may impact the effort that examinees put forth, thereby impacting test performance. Lateralization may also explain these results. In this study, most participants were right-handed, meaning that the performance on both the GPT (dominant hand) and COWAT would be lateralized to the left hemisphere (Simos et al., 2005). Thus, it is possible that the low rapport condition renders the left hemisphere particularly vulnerable when attempting to respond to novel circumstances such as the GPT and COWAT require. However, if the results were solely a function of hemispheric lateralization, then differences would be expected on other measures, such as the RAVLT.

Rapport may also be a factor in assessment. Charny (1966) found that in therapy, body motion behavior and lexical content contributed to the relatedness of the patient-therapist relationship. This can be seen in our study as well. When the administrator was cold, their body language was more rigid and uninviting, whereas in the high rapport group, the administrator had a more laid back and approachable stature. Leach (2005) also discusses the importance of rapport in the therapeutic alliance; stating that clients who have a strong rapport with their therapist are more likely to have a better and more positive outcome. This was also evident in our study. Participants who were in the high rapport group felt more positively toward their examiner, as indicated by the manipulation check, and produced better scores on the dominant hand of the GPT and the initial trial of the COWAT than their low rapport counterparts. This has implications for a number of neuropsychological assessment applications, such as court-ordered evaluations.

A limitation of this study was the limited age representation of participants as well as the lack of measures of effort. Future research should include a wider age representation, which would make the results more generalizable to a wider population, and include measures of effort, and perhaps adapted versions of measures of the therapeutic alliance. Another limitation of this study was that the neuropsychological tests were administered in the same order, meaning that order effects may have influenced the results. Future studies should utilize different tests and randomize or counterbalance the sequence of tests administered. It would also be useful to apply more varied experimental scripts so that optimal testing outcomes may be more closely linked to specific examiner/participant interactions. Additionally, future studies with neuropsychological tests that focus on specific cortical areas may further inform the impact of rapport on testing outcomes.

Being able to manipulate rapport, as shown in this study, allows examiners to become an uncontrolled variable on the outcome of certain neuropsychology assessment scores. Because of this, it is important to determine what the optimum rapport condition looks like in order to give all participants a fair and standardized administration. Also, by addressing the tests which are most influenced by rapport, we may be able to determine the extent to which rapport influences test performance and the ways in which performance on some of these tests may be more sensitive to these social emotional cues/stimuli. After further research into what tests are being affected and what rapport conditions allow for the greatest outcome, more structured rapport guidelines may need to be imposed on neuropsychological tests.

References

Ackerman, S. J., & Hilsenroth, M. J. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clinical Psychology Review*, *23*, 1–33. doi:10.1016/S0272-7358(02)00146-0

Aiken, L. R. (1999). *Instructor's manual with test questions for personality assessment methods and practices*. Kirkland, WA: Hogrefe & Huber.

Allen, A., Montgomery, M., Tubman, J., Frazier, L., & Escovar, L. (2003). The effects of assessment feedback on rapport-building and self-enhancement processes. *Journal of Mental Health Counseling*, *25*, 165–181. doi:10.17744/mehc.25.3.lw7h84q861dw6ytj

Bennett, M. R., & Hacker, P. M. S. (2005). Emotion and cortical-subcortical function: Conceptual developments. *Progress in Neurobiology*, *75*, 29–52. doi:10.1016/j.pneurobio.2004.11.002

Benton, A., Sivan, A., Hamsher, K., Varney, N., & Spreen, O. (1983). *Neuropsychological assessment: A clinical manual*. New York, NY: Oxford University Press.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, *16*, 252. doi:10.1037/h0085885

Cacioppo, J. T. (2002). Social neuroscience: Understanding the pieces fosters understanding the whole and vice versa. *American Psychologist*, *57*, 819. doi:10.1037/0003-066x.57.11.819

Cacioppo, J. T., & Berntson, G. G. (2004). *Social neuroscience: Key readings*. New York, NY: Ohio State University Psychology Press.

Charny, J. E. (1966). Psychosomatic manifestations of rapport in psychotherapy. *Psychosomatic Medicine*, *28*, 305–315. doi:10.1097/00006842-196607000-00002

Constantinou, M., Ashendorf, L., & McCaffrey, R. J. (2005). Effects of a third party observer during neuropsychological assessment: When the observer is a video camera. *Journal of Forensic Neuropsychology*, *4*, 39–47. doi:10.1300/j151v04n02_04

Cozolino, L. J. (2014). *The neuroscience of human relationships: Attachment and the developing social brain*. New York, NY: W.W. Norton.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive functions system*. San Antonio, TX: The Psychological Corporation.

DiPietro, J. A., & Larson, S. K. (1989). Examiner effects in the administration of the NBAS: The illusion of reliability. *Infant Behavior and Development*, *12*, 119–123. doi:10.1016/0163-6383(89)90057-x

Downer, J. T., Rimm-Kaufman, S. E., & Pianta, R. C. (2007). How do classroom conditions and children's risk for school problems contribute to children's behavioral engagement in learning? *School Psychology Review*, *36*, 413.

Easton, A., & Emery, N. (2005). *The cognitive neuroscience of social behaviour*. Hove, NY: Psychology Press. doi:10.4324/9780203311875

Fleiss, J. L., Slakter, M. J., Fischman, S. L., Park, M. H., & Chilton, N. W. (1979). Inter-examiner reliability in caries trials. *Journal of Dental Research*, *58*, 604–609. doi:10.1037/0033-2909.86.2.420

Gavett, B. E., Lynch, J. K., & McCaffrey, R. J. (2005). Third party observers: The effect size is greater than you might think. *Journal of Forensic Neuropsychology*, *4*, 49–64. doi:10.1300/j151v04n02_05

Gavett, B. E., & McCaffrey, R. J. (2007). The influence of an adaptation period in reducing the third party observer effect during a neuropsychological evaluation. *Archives of Clinical Neuropsychology*, *22*, 699–710. doi:10.1016/j.acn.2007.05.002

Geary, D. C. (2005). *The origin of mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.

Golden, C. (1978). Stroop Colour and Word Test. *Age*, *15*, 90. doi:10.4135/9781412952644.n441

Gorske, T. T. (2007). Therapeutic neuropsychological assessment: A humanistic model and case example. *Journal of Humanistic Psychology*, *48*, 320–339. doi:10.1177/0022167807303735

Harmon-Jones, E., & Winkielman, P. (2007). *Social neuroscience: Integrating biological and psychological explanations of social behavior*. New York, NY: Guilford Press. doi:10.5860/choice.45-0561

Harvey, P., Hamlin, M., Kumar, R., Morgan, G., Spurgeon, A., & Delves, H. (1988). Relationships between blood lead, behaviour, psychometric and neuropsychological test performance in young children. *British Journal of Developmental Psychology*, *6*, 145–156. doi:10.1111/j.2044-835x.1988.tb01089.x

Hilsenroth, M. J., Kivlinghan, Jr. D. M., & Slavin-Mulford, J. (2015). Structures supervisions of graduate clinicians on psychodynamic psychotherapy: Alliance and technique. *Journal of Counseling Psychology*, *62*, 173. doi:10.1037/cou0000058

Hilsenroth, M. J., Peters, E. J., & Ackerman, S. J. (2004). The development of therapeutic alliance during psychological assessment: Patient and therapist perspectives across treatment. *Journal of Personality Assessment*, *83*, 332–344. doi:10.1207/s15327752jpa8303_14

Horowitz, J. E., & McCaffrey, R. J. (2008). Effects of a third party observer and anxiety tests on executive function. *Archives of Clinical Neuropsychology*, *23*, 409–417. doi:10.1016/j.acn.2008.02.002

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, *38*, 139. doi:10.1037/0022-0167.38.2.139

Huguet, P., Galavaing, M. P., Monteil, J. M., & Dumas, F. (1999). Social presence effects in the stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology*, *77*, 1011–1025. doi:10.1037/0022-3514.77.5.1011

Joe, G. W., Simpson, D. D., Dansereau, D. F., & Rowan-Szal, G. A. (2001). Relationships between counseling rapport and drug abuse treatment outcomes. *Psychiatric Services*, *52*, 1223–1229. doi:10.1176/appi.ps.52.9.1223

Kline, P. (2013). *The handbook of psychological testing*. 2nd ed. Florence, KY: Taylor & Francis/Routledge.

Kløve, H. (1963). *Grooved pegboard*. Lafayette, IN: Lafayette Instruments.

Kushnir, T., & Duncan, K. (1978). An analysis of social facilitation effects in terms of signal detection theory. *The Psychological Record*, *28*, 535.

Lasky, D. I., Felice, A., Moyer, R. C., Buddington, J. F., & Elliot, E. S. (1973). Examiner effects with the Peabody Picture Vocabulary Test. *Journal of Clinical Psychology*, *29*(4), 456–457. doi:10.1002/1097-4679(197310)29:4<456::aid-jclp2270290416>3.0.co;2-z

Lench, H. C., Flores, S. A., & Bench, S. W. (2011). Discrete emotions predict changes in cognition, judgement, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, *137*, 834–855. doi:10.1037/a0024244

Leach, M. J. (2005). Rapport: A key to treatment success. *Complementary Therapies in Clinical Practice*, *11*, 262–265. doi:10.1016/j.ctcp.2005.05.005

Martin, D. J., Garske, J. P., & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *Journal of Consulting and Clinical Psychology*, *68*, 438. doi:10.1037/0022-006x.68.3.438

McCaffrey, R. J., Lynch, J. K., & Yantz, C. L. (2005). Third party observers: Why all the fuss? *Journal of Forensic Neuropsychology*, *4*, 1–15. doi:10.1300/j151v04n02_01

Meyerbroker, K., & Emmelkamp, P. M. (2008). Therapeutic processes in virtual reality exposure therapy: The role of cognitions and the therapeutic alliance. *Journal of Cyber Therapy and Rehabilitation*, *1*, 247–258.

Millis, S. R., & Volinsky, C. T. (2001). Assessment of response bias in mild head injury: Beyond malingering tests. *Journal of Clinical and Experimental Neuropsychology*, *23*, 809–828. doi:10.1076/jcen.23.6.809.1017

Patterson, C. L., Anderson, T., & Wei, C. (2014). Clients' pretreatment role expectations, the therapeutic alliance, and clinical outcomes in outpatient therapy. *Journal of Clinical Psychology*, *70*, 673–680. doi:10.1002/jclp.22054

Platania, J., & Moran, G. P. (2001). Social facilitation as a function of the mere presence of others. *The Journal of Social Psychology*, *141*, 190–197. doi:10.1080/00224540109600546

Rath, J. F., Langenbahn, D. M., Simon, D., Sherr, R. L., Fletcher, J., & Diller, L. (2004). The construct of problem solving in higher level neuropsychological assessment and rehabilitation. *Archives of Clinical Neuropsychology*, *19*, 613–635. doi:10.1016/j.acn.2003.08.006

Rey, A. (1964). *L'examen clinique en psychologie*. Paris, France: Presses Universitaires de France.

Roberts, D. L., & Penn, D. L. (2009). The effects of task engagement and interpersonal rapport on WCST performance in schizophrenia. *American Journal of Psychiatric Rehabilitation*, *12*, 57–72. doi:10.1080/15487760802615798

Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV tests of achievement*. Rolling Meadows, IL: Riverside.

Sharpley, C. F., Fairnie, E., Tabary-Collins, E., Bates, R., & Lee, P. (2000). The use of counsellor verbal response modes and client-perceived rapport. *Counseling Psychology Quarterly*, *13*, 99–116. doi:10.1080/09515070050011097

Sharpley, C. F., Jeffrey, A. M., & Mcmah, T. (2006). Counselor facial expression and client-perceived rapport. *Counseling Psychology Quarterly*, *19*, 343–356. doi:10.1080/09515070601058706

Sharpley, C. F., & McNally, J. (1997). Effects of level of academic training on client-perceived rapport and use of verbal response modes in counseling dyads. *Counseling Psychology Quarterly*, *10*, 449–460. doi:10.1080/09515079708254191

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*, 420. doi:10.1037/0033-2909.86.2.420

Simos, P. G., Sarkari, S., Castillo, E. M., Billingsley-Marshall, R. L., Pataraia, E., Clear, T., & Papanicolaou, A. C. (2005). Reproducibility of measures of neurophysiological activity in Wernicke's area: A magnetic source imaging study. *Clinical Neurophysiology*, *116*, 2381–2391. doi:10.1016/j.clinph.2005.06.019

Skuy, M., Schutte, E., Fridjhon, P., & O'Carroll, S. (2001). Suitability of published neuropsychological test norms for urban African secondary school students in South Africa. *Personality and Individual Differences*, *30*, 1413–1425. doi:10.1016/s0191-8869(00)00123-9

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York, NY: Oxford University Press.

Thomas, S. L., Skitka, L. J., Christen, S., & Jurgena, M. (2002). Social facilitation and impression formation. *Basic and Applied Social Psychology*, *24*, 67–70. doi:10.1207/s15324834basp2401_6

Uziel, L. (2007). Individual differences in the social facilitation effect: A review and meta-analysis. *Journal of Research in Personality*, *41*, 579–601. doi:10.1016/j.jrp.2006.06.008

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4th ed.). San Antonio, TX: Pearson.

Wilkonsen, G. S., & Robertson, G. J. (2009). *Wide Range Achievement Test* (4th ed.). (WRAT-IV). Lutz, FL: Psychological Assessment Resources.

Yantz, C. L., & McCaffrey, R. J. (2005). Effects of a supervisor's observation in memory test performance of the examinee: Third party observer effect confirmed. *Journal of Forensic Neuropsychology*, *4*, 27–38. doi:10.1300/j151v04n02_03

Zajonc, R. B. (1965). Social facilitation. *Science*, *149*, 269–274. doi:10.1126/science.149.3681.269

Appendix 1. Manipulation check

1. For each word below, indicate to what extent you believe that it describes your examiner (the person who just administered the tests to you). This feedback is totally anonymous. The examiner will not be told and will not see your responses.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Nice	○	○	○	○	○	○	○
Competent	○	○	○	○	○	○	○
Strict	○	○	○	○	○	○	○
Mean	○	○	○	○	○	○	○
Friendly	○	○	○	○	○	○	○
Warm	○	○	○	○	○	○	○
Professional	○	○	○	○	○	○	○
Incompetent	○	○	○	○	○	○	○
Lenient	○	○	○	○	○	○	○
Cold	○	○	○	○	○	○	○
Unprofessional	○	○	○	○	○	○	○
Wanted me to do well	○	○	○	○	○	○	○
Didn't want me to do well	○	○	○	○	○	○	○
Inviting	○	○	○	○	○	○	○
Aloof	○	○	○	○	○	○	○
Irritated	○	○	○	○	○	○	○
Stiff	○	○	○	○	○	○	○
Easy-going	○	○	○	○	○	○	○
Helpful	○	○	○	○	○	○	○

2. For each word below, indicate to what extent you believe that it describes the testing environment and overall experience. This feedback is totally anonymous. The examiner will not be told and will not see your responses.

	Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
Pleasant	○	○	○	○	○	○	○
Boring	○	○	○	○	○	○	○
Comfortable	○	○	○	○	○	○	○
Intriguing	○	○	○	○	○	○	○
Uncomfortable	○	○	○	○	○	○	○
Awkward	○	○	○	○	○	○	○
Fun	○	○	○	○	○	○	○
Unpleasant	○	○	○	○	○	○	○