5-7-2018

Using Text Comprehension Analyses to Enhance the Readability of Self-Report Questionnaires of Gambling Disorder

Samuel Cody Peter

USING TEXT COMPREHENSION ANALYSES TO ENHANCE THE READABILITY
OF SELF-REPORT QUESTIONNAIRES OF GAMBLING DISORDER

by

Samuel C. Peter

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Major: Psychology

The University of Memphis

May 2018

Abstract

The purpose of this study was to use advanced text-comprehension tools to develop a questionnaire of gambling disorder symptoms, the Memphis Gambling Measure (MGM), then experimentally compare rates of accurate comprehension and symptom identification as compared to the NODS, an often used and theoretically less readable questionnaire of the same symptoms. Eighty-five volunteers identified symptoms in a clinical vignette by completing either the MGM or NODS in a between-subjects experimental design. Participants who completed the MGM correctly identified more symptoms of gambling disorder than participants who completed the NODS. Participants with more education more accurately responded to the questionnaire items, but this did not moderate the effect of questionnaire assignment on item comprehension. We concluded that item comprehension can be accurately predicted using the present text-analysis assessment methods, and that rates at which individuals accurately report on symptoms of psychopathology is related to the readability of the questionnaire items themselves.

Table of Contents

List of Tables

Introduction

The degree of reliability of responses obtained via a self-report questionnaire depends upon the respondent's comprehension of the items that comprise it (Lenzner, 2012). The items of self-report questionnaires of psychopathology symptoms have been subjected to readability formulas that primarily focus on sentence- and word-length to predict respondent-comprehension of these items for decades, despite advancements in text-comprehension assessment methods (Ash & Edgell, 1975; McHugh et al., 2014). Peter, Whelan, Pfund, and Meyers (2017) assessed self-report questionnaires of gambling disorder symptoms using contemporary text-comprehension assessment software and found that that significant problems existed with commonly used questionnaires, calling to question their ability to accurately identify symptoms of gambling disorder, especially among individuals of relatively lower reading abilities. The purpose of the present study was two-fold: (1) to develop a self-report questionnaire of gambling disorder symptoms using contemporary text-comprehension assessment methods to maximize the comprehensibility of the questionnaire's items; and (2) to test if participants could accurately interpret the items and thereby accurately identify the presence or absence of diagnostic symptoms in an individual depicted in a clinical vignette, in comparison to a commonly used self-report questionnaire of gambling disorder. We compared rates of accurate symptom identification between participants who were assigned to complete either this newly developed self-report questionnaire of gambling disorder symptoms or the National Opinion Research Center Screen for Gambling Problems (NODS; Gerstein et al., 1999), a widely used diagnostic measure of gambling disorder.

In his 1948 paper, Flesch proposed that greater sentence length and words length would result in greater difficulty in comprehension. This definition led to the proliferation of readability

formulas that primarily focused on these two variables, including the Flesch Reading Ease formula (FRE; Flesch, 1948) and Simple Measure of Gobbledygook (SMOG; McLaughlin, 1969), among others. These formulas have been used for decades to review psychopathology questionnaire readability. For example, both Ash and Edgell (1975) and McHugh and colleagues (2014) used the FRE and SMOG to assess the readability of items on self-report questionnaires of psychological constructs.

The validity of relying upon traditional readability formulas to predict question comprehension has been evaluated. Lenzner (2014) identified pairs of survey questions from journal articles and textbooks on questionnaire design where one question was a problematic version and the other a non-problematic version of the same question (i.e., more difficult to comprehend versus easier to comprehend). He then tested whether commonly used readability formulas that focused on sentence- and word-length, such as the FRE, correctly differentiated the problematic survey questions from the improved questions. In more than half of the trials readability formulas identified the problematic version as the more comprehensible option illustrating the limits of traditional readability assessments.

Lenzner (2014)'s study was not the first to question the validity of relying on traditional readability formulas to predict question comprehension. In another study, Holbrook, Cho, & Johnson (2006) found that the readability of a question, operationalized as the Flesch-Kinkaid Grade Level (FKG; Kincaid, Fishburne Jr, Rogers, & Chissom, 1975), was not linearly related to comprehension difficulties. In fact, questions that scored highest on the FKG, which would theoretically be the hardest to understand, were better understood than questions of a "medium" reading level. These studies, among others (Lenzner, 2014), challenge the assumption that

readability formulas that primarily focus on sentence and word length can be reliably used to predict comprehension difficulties on self-report questionnaires.

Gambling disorder affects individuals of all education levels, perhaps disproportionally affecting those with relatively lower education attainment (Bastiani et al., 2013). Unsurprisingly, results of a large-scale adult literacy survey (Kirsch, 1993) indicated that the best predictor of reading abilities was amount of educational attainment. Thus, a significant portion of individuals with gambling problems likely have limited reading abilities. The presence of individuals with relatively lower reading abilities among populations of disordered gamblers has key implications for the emphasis psychologists place on the readability of items on diagnostic questionnaires of gambling disorder. The readability of each individual item may substantially impact the decision to classify an individual as a disordered gambler.

Peter and colleagues (2017) reviewed self-report questionnaires of gambling disorder severity using two modern text comprehension tools. The first was Coh-Metrix (McNamara, Graesser, McCarthy, & Cai, 2014; Graesser, McNamara, & Kulikowich, 2011), an instrument used to analyze a body of text to produce composite scores of numerous variables that more reliably predict comprehension difficulties. For example, Peter and colleagues (2017) utilized the "syntactic simplicity" composite score, which considers sentence characteristics such as the number of words before the main verb, the number of embedded clauses within a sentence, and active versus passive voice, among other characteristics of sentence-syntax in order to gauge the simplicity of the syntax of the sentence, which in turn predicts the ease of reader-comprehension. The second tool used in this review was QUAID, which stands for "question understanding aid" (Graesser, Cai, Louwerse, & Daniel, 2006). QUAID has also been used in at least two other reviews of questionnaire readability (McHugh, Rasmussen, & Otto, 2011; Richards et al., 2013).

3

QUAID uses similar technology to Coh-Metrix, but identifies specific problems within questions and makes suggestions for how these problems could be best ameliorated. For example, if a question is likely to overload on a reader's working memory, QUAID may suggest that this question be broken up into multiple sentences.

Based on their review, Peter et al. (2017) concluded that individuals of below-average reading abilities likely misunderstand a significant number of items on commonly used self-report questionnaires of gambling disorder. One measure that was analyzed was the NODS, which has been used in large prevalence studies of gambling disorder (Gerstein et al., 1999), and continues to be used within empirical investigations (Weinstock, April, & Kallmi, 2017). The NODS assesses for the ten Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV)' criteria for pathological gambling. When mapped onto the diagnostic criteria that they represent, 8 of the 10 groups of questions contained items written at a below-average level of syntactic simplicity (i.e., less readable). Furthermore, eight items were likely to overload the working memory of the respondent. Peter and colleagues (2017) therefore concluded that individuals completing the NODS would likely misunderstand many of its items, negatively impacting the diagnostic accuracy of the questionnaire.

The Present Study

The present study was designed to accomplish two goals: (1) to develop a self-report questionnaire of gambling disorder symptomatology with the specific goal of optimizing the readability of its items; and (2) to test the rate at which participants accurately comprehend the items of this self-report questionnaire, relative to rates obtained using an alternative, theoretically less readable questionnaire, the NODS. Regarding the first goal, our research team primarily utilized Coh-Metrix and QUAID to guide the development of the Memphis Gambling Measure

(MGM). Regarding the second goal, a between-subjects, experimental design was utilized. Academic achievement tests often have individuals answer questions about a story in order to assess their comprehension of both the story and questions about it (e.g., Wechsler, 2009). Additionally, many psychological studies utilize vignettes as a standardized stimulus that all individuals can respond to under variable conditions (e.g., Hing, Russell, Gainsbury, & Nuske, 2016). A combination of these two techniques was used to assess how well individuals comprehend questions about an individual struggling with a mental health problem. In this study, we asked participants to read a vignette about an individual with gambling problems, then assessed their comprehension of questions about the individual depicted in the vignette. The questionnaire that individuals completed was experimentally manipulated, such that participants were randomly assigned to complete either the MGM or the NODS in relation to the individual depicted in the study vignette. We recruited individuals of different educational attainment in order to assess whether the effect of the readability of the questionnaires was moderated by the educational level of the respondent.

We hypothesized that individuals who were assigned to respond to the MGM would interpret the items more accurately and correctly identify more diagnostic criteria than individuals who were assigned to read the NODS. We also hypothesized that this effect would be stronger among individuals of the lower educational attainment group than individuals in the higher educational attainment group.

Method

Participants

Participants were adults aged 18 years or older who were proficient in English. They were required to have either: (1) never attended college or, or (2) graduated from college with a

four-year degree. This inclusion criteria was employed to test for the moderating effect of education on the relationship between the primarily independent variable of questionnaire assignment and dependent variables of item comprehension and diagnostic accuracy. Participants were recruited at local health fairs, community talks on problem gambling awareness, social media pages, job training facilities, and resource centers for economically disadvantaged individuals. Participation was voluntarily. In total, 105 were eligible and agreed to participate.

Of these 105 individuals, 20 were excluded from the final database based on a failure to follow study procedures, identified from reading checks (e.g., "If you are reading this question, please do not select either answer"). The characteristics of the resulting sample ($N = 85$) can be found in Tables 1 and 2.

Materials

Demographics. This questionnaire asked individuals about their age, gender, race, ethnicity, income, and level of education.

National Opinion Research Center DSM-IV screen for gambling problems (NODS; Gerstein et al., 1999). The NODS was originally developed as a structured interview assessing for DSM-IV criteria for pathological gambling. It has demonstrated fair psychometric properties when used as a structured interview, such as fair internal consistency and moderate correlation with gambling behavior, but poor agreement with other measures of gambling disorder severity (Hodgins, 2004). Because the DSM-5 symptoms of gambling disorder are all assessed in DSM-IV measures of pathological gambling, it continues to be used to assess for DSM-5 symptoms of gambling disorder, often times as a self-report questionnaire despite its original development as a structured interview (e.g., Nehlin, Nyberg, & Jess, 2016; Ledgerwood & Arfken, 2017). The items of the NODS can be found in Table 3, as well as the corresponding DSM criteria the items

are intended to assess. Readability indices of the NODS can be found in Table 4. The average

syntactic simplicity of the items of the NODS is $z = -0.30$. Eleven items of the NODS are written

at a below-average level of syntactic simplicity, as measured by Coh-Metrix. Eight of the items

on the NODS are likely to overload the working memory of the respondent, as judged by

QUAID (for further details regarding the readability of the NODS items, see review by Peter et

al., 2017).

     Memphis Gambling Measure (MGM; Table 3). Three methods were used to develop

the MGM: (1) expert judgment; (2) Coh-Metrix and QUAID; and (3) pilot testing with

community volunteers. First, a research team of gambling researchers and clinicians, including

both licensed Ph.D. level clinical psychologists and clinical psychology graduate students,

developed an initial draft of the MGM with the intent to create an assessment tool that was both

readable and adequately assessed the DSM-5 criteria for gambling disorder. This first draft was

then systematically edited until all items met predefined objective readability criteria, namely

that: (1) no item was written at a below average level of syntactic simplicity (assessed with Coh-

Metrix); and (2) QUAID did not identify any items as placing a working memory overload on

respondents. This penultimate draft was then examined by the research team to ensure that the

original integrity of the questions was retained throughout the editing process.

     Next, this penultimate draft underwent several rounds of pilot testing with individuals

from the community, primarily at community health fairs and problem gambling awareness

presentations. Individuals volunteered to read the questions and provide qualitative feedback on

the items' clarity. Several modifications were made to the MGM based on this feedback. For

example, while one item originally read, "Have you ever tried to limit your gambling?" feedback

from community volunteers suggested that this question was abstract and lacked clarity, so the

question was changed to, "Have you ever tried to spend less time or money gambling?" Despite the readability tools' resulting indices, which would suggest that this item was now less readable, the team allowed the pilot data to drive the wording of the questions from this point forward, until all questions were appraised as being acceptably clear by the community volunteers, and still met the previously specified readability criteria.

The resulting readability indices of the MGM can be found in Table 4. The average syntactic simplicity of the MGM items was $z = 1.91$. No item is written at a below-average level of syntactic simplicity or was identified as placing an excessive burden on the respondents' working memory, as judged by QUAID.

Study Vignette (Appendix A). This vignette describes a 28-year-old white male named Michael struggling with a gambling problem. Readability analyses indicate a narrativity of $z = 1.71$, indicating it is more readable than 96% of the texts that typical student would encounter in their kindergarten-12th grade educational career (for more details, see Graesser et al., 2011). The narrativity score was selected given that it is a general measure of the ease of comprehending the narrative, or story, or the body of text. A detailed explanation of the measures that make up this composite score can be found in Graesser et al. (2011).

Two licensed Ph.D.-level clinical psychologists independent of the research team were recruited to aid with initial scoring of the MGM and NODS in relation to the study vignette. Both clinical psychologists independently read the study vignette and completed the NODS and MGM. Their answers were then compared. As expected, the two psychologists responded identically (i.e., the two psychologists gave precisely the same answers to both questionnaires). Therefore, these answers were used as the standard for comparison (i.e., judged as the correct answers) for subsequent scoring of participants answers as correct or incorrect.

8

Procedure

      All study procedures were approved by the University's Institutional Review Board prior to data collection. Consenting participants completed the demographics questionnaire and were instructed to read the vignette about Michael struggling with his gambling problems. Next, participants were asked to respond to questions as if they were the individual in the vignette (i.e., from Michael's perspective). A sample item was provided that asked whether participants understood that they were to answer the following items from Michael's perspective and not their own. If participants answered incorrectly (indicating that they were to respond to the questions from their own perspective), they were provided with corrective feedback, reminding them to complete the questions as if they were Michael, the subject of the vignette. If participants responded correctly (that they were supposed to respond as if they were Michael), they were provided with confirmatory feedback (i.e., "correct! Please respond to the following questions as if you were Michael, the subject of the vignette). Following this instruction, participants were randomized to complete one of the two possible questionnaires described above, the NODS or MGM. Participants were allowed to look back at the vignette while responding to the questions.

<p align="center">Results</p>

Data Cleaning

      An initial inspection of the data revealed substantial negative skewness in the proportional variable of percent of items answered correctly. Therefore, we applied a reflection and logarithmic transformation to this variable (Tabachnick & Fidell, 2007). This transformation significantly improved the normality of the data. Therefore, this transformed variable was used in all subsequent analyses.

Sample Demographic Characteristics

Participants did not significantly differ on any demographic variables based on the

questionnaire that they were randomly assigned to (Table 1; all $ps > .05$). Alternatively, groups

based on education level did significantly differ on multiple demographic variables. Specifically,

individuals in the high education group were significantly older [$t(83) = 4.77$, $p < .0005$], more

likely to be female [χ^2 ($N = 39$, 1) = 19.23, $p < .0005$] and more likely to be white [χ^2 ($N = 25$, 3)

= 16.51, $p < .001$] (Table 2).

To determine whether these variables should be included as covariates in the primary

analysis, the relations between demographic variables and the primary dependent variable of

correctly answered questions were explored. Four separate analyses were conducted. Because we

did not hypothesize that these variables would be significantly related to the primary dependent

variable, a bonferroni correction was applied, dividing the standard alpha of $p = .05$ by four,

resulting in a more conservative alpha of $p = .01$ being used to indicate significance. The

correlation between participants' age and the percent of questions answered correctly was weak

and nonsignificant, $r = .12$, $p = .29$. The relationship between the percent of questions answered

correctly and participants' race [$F(3, 81) = 2.85$, $p = .04$], ethnicity [$F(2, 82) = .06$, $p = .94$], and

gender [$F(1, 83) = 5.19$, $p = .03$] were all non-significant as well. Interactions between these

variables were also explored, and none were significantly related to the percent of questions

answered correctly. Thus, no demographic variables were included in the primary analyses.

Item Comprehension

A 2x2 ANOVA was conducted to compare the percentage of questions answered

correctly based on two factors: the randomly assigned questionnaire (NODS or MGM) and

participant education group (never attended college or 4-year college completers). There were

significant univariate main effects for both the questionnaire factor and education group factor.

With regards to the questionnaire main effect, individuals correctly answered approximately 80% ($SD = 17.49$) of the MGM questions compared to 69% ($SD = 23.26$) of the NODS questions, $F(1, 81) = 7.64$, $p = .007$, partial $\eta^2 = .09$. With regards to the education group factor, individuals in the high education group correctly answered 82% ($SD = 16.47$) of the MGM questions compared to 66% ($SD = 23.77$) of the NODS questions, $F(1, 81) = 14.99$, $p < .0005$, partial $\eta^2 = .16$. The interaction effect was not significant, $F(1, 81) = .47$, $p = .40$, partial $\eta^2 = .01$. The mean percent of questions answered correctly for these four groups are displayed in Figure 1.

Symptom Identification

In order to better understand the effect of readability on diagnostic accuracy, including both sensitivity and specificity, the questions of the MGM and NODS were mapped onto the DSM criteria they are meant to assess. Then, the readability of these questions was averaged to represent the average readability of the questions used to detect the presence or absence of the diagnostic criteria for gambling disorder in each measure. These metrics are displayed in Table 5. Also in Table 5 are the percentages of individuals who correctly identified the presence or absence of the diagnostic criteria in the clinical vignette, grouped by the questionnaire that they completed.

These data are also displayed in Figure 2, which orders the DSM criteria based on the difference in readability between the MGM and NODS. For example, the DSM criteria with the smallest difference in readability between MGM and NODS criteria, criteria A7, is placed on the far left of the x-axis, and the rest in ascending order towards the right. This analysis was conducted in order to observe whether the effect of individuals more accurately identifying diagnostic criteria using the MGM as oppose to the NODS became stronger as the difference

11

between the items' readability increased. However, there did not appear to be a clear trend in this direction.

Figure 2 also conveys the general improvement in both sensitivity and specificity gained by using the MGM instead of the NODS to judge the clinical vignette. In all but two of the nine diagnostic criteria, individuals who completed the MGM successfully identified the criteria approximately 5-10% more often than individuals who completed the NODS. For one diagnostic criterion, A4, no difference was observed. For criteria A1, individuals who completed the NODS more accurately assessed this criterion than individuals who completed the MGM.

Discussion

The present study was designed to assess whether altering the readability of a self-report questionnaire of gambling disorder would yield more accurate responses and thereby increase the likelihood of correctly identifying the presence or absence of diagnostic criteria for gambling criteria by community volunteers. These goals were achieved by creating a clinical vignette of an individual with gambling disorder, then having community members answer questions about the individual in the vignette. Using a between-subjects experimental design, individuals who completed a more readable measure of gambling disorder, the MGM, more accurately responded to questions and identified the presence or absence of more diagnostic criteria than individuals who completed the NODS, a less readable questionnaire.

We also explored whether the effect of questionnaire comprehensibility would be stronger among individuals with lower reading abilities. To test this potential moderator, we recruited two groups of participants – individuals who either had a four-year college degree, and individuals who had never attended college. A main effect of education level was present, such that individuals of higher education level answered more questions correctly than individuals of

12

lower-educational attainment. However, the interaction effect was non-significant. Thus, the results of the present study did not support the hypothesis that the effect of questionnaire readability would be stronger among individuals of a lower educational attainment. One explanation is that this effect is equally strong regardless of the respondent's educational level, which is supported by the small effect size of the interaction. Perhaps convoluted questions are equally difficult to understand, regardless of how much education an individual has obtained.

Importantly, the present study demonstrated the effect that self-report questionnaire readability has on diagnostic accuracy. The NODS has been used to estimate the prevalence of gambling disorder in the general population (Gerstein et al., 1999). Although the NODS is typically used as a structured interview, research has not supported the assumption that questions read out-loud by someone else are easier to comprehend that questions read silently to oneself (McCallum, Sharp, Bell, & George, 2004). Thus, the present findings may call to the question the accuracy of these prevalence studies, especially among populations of relatively lower reading abilities.

Limitations to the present study center on the reliance on a clinical vignette to provide participants with an individual's gambling experience to reflect on as oppose to a true self-report process. It may be that individuals experience the questions differently when answering them from the perspective of the individual in the vignette as oppose to themselves. A second limitation is that this experimental design relied on the assumption that participants could accurately comprehend the clinical vignette. This assumption was made safer by subjecting the clinical vignette to readability analyses that indicated participants should be able to comprehend it. A third limitation is that we assumed that the clinical vignette provided sufficient information for participants to complete the questionnaires accurately. This assumption was made safer by

having licensed clinical psychologists independent of the research team independently code the presence of diagnostic symptoms using both questionnaires, and observing 100% agreement between the two.

Future directions may focus on further validating the MGM, which may include reconsidering the wording of the items that assess for gambling disorder criteria A1 and A5 (Table 3). Individuals that completed the NODS identified the presence of symptom A1 at a higher rate than individuals assigned to complete the MGM, despite the MGM item being more readable than the NODS item. It could be that although the MGM item was designed to be easier to understand, the specific phrasing itself no longer accurately captures the construct of interest. The MGM asks about gambling more to continue to *enjoy* gambling, whereas the NODS asks about gambling more to attain a feeling of *excitement.* The MGM's departure from the wording of the diagnostic criteria itself may have inadvertently sacrificed diagnostic accuracy for the sake of item comprehension; thus questionnaire developers should consider if enhancing item readability detracts from other aspects of validity and reliability in an iterative process.

It is also worth noting the overall poor identification of criteria A5 across questionnaires. This is the only criteria that more than half of participants were unable to correctly identify. Although our research team and the independent psychologists believed that the vignette did not depict the subject as gambling to alleviate a negative mood state, most participants did. It could be that lay individuals simply do not interpret this phenomenon in the same way that psychologists do. Future directions may examine how respondents conceptualize this specific symptom, and if self-report questionnaires adequately assess its presence.

We would also hope that future questionnaire development processes utilize methods used in the present study, such as assessing the readability of items with contemporary text-

comprehension tools (Coh-Metrix and QUAID), community pilot testing, and direct comprehension tests. Limiting the assessment of a self-report questionnaire's readability to a subjection to classic readability formulas is insufficient and has become obsolete in light of advancements in cognitive psychology. Moving forward, best-practice regarding questionnaire validation would make use of these advancements.

References

Ash, R. A., & Edgell, S. L. (1975). A note on the readability on the Position Analysis

Questionnaire (PAQ). *Journal of Applied Psychology*, *60*, 765.

Bastiani, L., Gori, M., Colasante, E., Siciliano, V., Capitanucci, D., Jarre, P., & Molinaro, S.

(2013). Complex factors and behaviors in the gambling population of Italy. *Journal of*

Gambling Studies, *29*, 1-13.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221.

Flesch, R. F. (1979). *How to write plain English: A book for lawyers and consumers*.

City/location: Harpercollins.

Gerstein, D., Murphy, S., Toce, M., Hoffman, J., Palmer, A., Johnson, R., Larison, C., Chuchro,

L., Bard, A., Engelman, L., Hill, M. A., Buie, T., Volberg, R., Harwood, H., Tucker, A.,

Christiansen, E., Cummings, W., & Sinclair, S. (1999). *Gambling impact and behavior*

study: Report to the national gambling impact study commission. Chicago, IL: National

Opinion Research Center.

Graesser, A. C., Cai, Z., Louwerse, M. M., & Daniel, F. (2006). Question Understanding Aid

(QUAID) A web facility that tests question comprehensibility. *Public Opinion Quarterly*,

70, 3–22.

Graesser, A. C., McNamara, D. S., Cai, Z., Conley, M., Li, H., & Pennebaker, J. (2014). Coh-

Metrix measures text characteristics at multiple levels of language and discourse. *The*

Elementary School Journal, *115*, 210-229.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix providing

multilevel analyses of text characteristics. *Educational Researcher*, *40*, 223-234.

Graesser, A. C., Wiemer-Hastings, K., Kreuz, R., Wiemer-Hastings, P., & Marquis, K. (2000).

    QUAID: A questionnaire evaluation aid for survey methodologists. *Behavior Research*

    Methods, Instruments, & Computers, *32*, 254-262.

Hing, N., Russell, A. T., Gainsbury, S. M., & Nuske, E. (2016). The public stigma of problem

    gambling: Its nature and relative intensity compared to other health conditions. *Journal of*

    Gambling Studies, 32(3), 847-864.

Hodgins, D. C. (2004). Using the NORC DSM Screen for Gambling Problems as an outcome

    measure for pathological gambling: psychometric evaluation. *Addictive Behaviors*, *29*(8),

    1685-1690.

Holbrook, A., Cho, Y. I., & Johnson, T. (2006). The impact of question and respondent

    characteristics on comprehension and mapping difficulties. *Public Opinion Quarterly*, *70*,

    565-595.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new*

    readability formulas (automated readability index, fog count and flesch reading ease

    formula) for navy enlisted personnel (No. RBR-8-75). Naval Technical Training

    Command Millington TN Research Branch.

Kirsch, I. S. (1993). *Adult literacy in America: A first look at the results of the National Adult*

    Literacy Survey. US Government Printing Office, Superintendent of Documents,

    Washington, DC 20402 (Stock No. 065-000-00588-3).

Ledgerwood, D. M., & Arfken, C. L. (2017). Assessing the Need for Higher Levels of Care

    Among Problem Gambling Outpatients. *Journal of Gambling Studies*, 1-13.

Lenzner, T. (2012). Effects of survey question comprehensibility on response quality. *Field*

    Methods, *24*, 409–428.

Lenzner, T. (2014). Are readability formulas valid tools for assessing survey question

difficulty? *Sociological Methods & Research*, *43*(4), 677-698.

McCallum, R. S., Sharp, S., Bell, S. M., & George, T. (2004). Silent versus oral reading

comprehension and efficiency. *Psychology in the Schools*, *41*(2), 241-246.

McHugh, R. K., Sugarman, D. E., Kaufman, J. S., Park, S., Weiss, R. D., & Greenfield, S. F.

(2014). Readability of self-report alcohol misuse measures. *Journal of Studies on Alcohol
and Drugs*, *75*, 328-334.

McLaughlin, G. H. (1969). SMOG grading-a new readability formula. *Journal of reading*, *12*(8),

639-646.

McNamara, D.S., Graesser, A.C., McCarthy, P.M., Cai, Z. (2014). *Automated evaluation of text

and discourse with Coh-Metrix.* Cambridge, MA: Cambridge University Press.

McNamara, D.S., Louwerse, M.M., Cai, Z., & Graesser, A. (2005, January). Coh-Metrix version

1.4. Retrieved August 2016, from http//:cohmetrix.memphis.edu.

Nehlin, C., Nyberg, F., & Jess, K. (2016). Brief Intervention Within Primary Care for At-Risk

Gambling: A Pilot Study. *Journal of gambling studies*, *32*(4), 1327-1335.

Peter, S. C., Whelan, J. P., Pfund, R. A., & Meyers, A. W. (2017). *A Text Comprehension

Approach to Questionnaire Readability: An Example Using Gambling Disorder

Measures.* Manuscript submitted for publication.

Richards, L. K., McHugh, R. K., Pratt, E. M., & Thompson-Brenner, H. (2013). Readability and

comprehension of self-report binge eating measures. *Eating Behaviors*, *14*, 167-170.

Tabachnick, B. L., & Fidell, L. S. (2007). *Using multivariate statistics.* Boston, MA: Pearson

Education, Inc.

Wechsler, D. (2009). *Wechsler Individual Achievement Test–Third Edition*. San Antonio, TX: Psychological Corporation.

Weinstock, J., April, L. M., & Kallmi, S. (2017). Is subclinical gambling really subclinical?. *Addictive behaviors*, *73*, 185.

Table 1

Sample Characteristics by Experimental Group

Sample characteristic	Experimental Group		Statistic (df)	*p*-value
	NODS	MGM		
N	41	44		
Age, *M* (*SD*)	42.37 (20.40)	39.52 (16.48)	$t(83) = -.709$	.48
Gender, *n* (%)			$\chi^2 (1) = .476$	.49
Female	26 (63)	31 (70)		
Male	15 (37)	13 (30)		
Ethnicity, *n* (%)			$\chi^2 (2) = 4.03$	.13
Hispanic	2 (5)	0 (0)		
Non-Hispanic	34 (83)	42 (95)		
Race, *n* (%)			$\chi^2 (3) = 4.01$	.26
Black or African-American	20 (49)	29 (66)		
Multiple Races	2 (5)	1 (2)		
White	19 (28)	13 (30)		
None of these[1]	0 (0)	1 (2)		
Education level			$\chi^2 (4) = .789$	.94
Less than High School	5 (12)	7 (16)		
High school or equivalent	15 (37)	14 (32)		
Four-year degree	8 (20)	11 (25)		
Masters' degree	10 (24)	9 (20)		
Doctoral degree	3 (7)	3 (9)		

Note. Options that no participants selected are not displayed, such as some racial categories. Cumulative percentages may not equal 100% due to missing data.

Table 2

Sample Characteristics by Education Group

Sample characteristic	Education Group		Statistic (df)	*p*-value
	Low	High		
N	41	44		
Age, *M* (*SD*)	32.10 (16.30)	49.09 (16.52)	$t(83) = -4.77$	$< .0005$
Gender, *n* (%)			$\chi^2 (1) = 19.23$	$< .0005$
Female	18 (44)	39 (89)		
Male	23 (56)	5 (11)		
Ethnicity, *n* (%)			$\chi^2 (2) = 4.03$	.13
Hispanic	0 (0)	2 (6)		
Non-Hispanic	37 (90)	39 (89)		
Race, *n* (%)			$\chi^2 (3) = 16.51$	.001
Black or African-American	30 (73)	19 (43)		
Multiple Races	3 (7)	0 (0)		
White	7 (17)	25 (57)		
Other	1 (2)	0 (0)		
Education level			$\chi^2 (4) = 85.00$	$< .0005$
Less than High School	12 (29)	0 (0)		
High school or equivalent	29 (70)	0 (0)		
Four-year degree	0 (0)	19 (43)		
Masters' degree	0 (0)	19 (43)		
Doctoral degree	0 (0)	6 (14)		

Note. Options that no participants selected are not displayed, such as some racial categories. Cumulative percentages may not equal 100% due to missing data.

Table 3

DSM Criteria for Gambling Disorder and Corresponding Items of the NODS and MGM

DSM Criteria	NODS Item(s)	MGM Item(s)
A1. Needs to gamble with increasing amounts of money in order to achieve the desired excitement.	In the past year, have there been any periods when you needed to gamble with increasing amounts of money or with larger bets than before in order to get the same feeling of excitement?	Have you needed to gamble with more money than before to continue to enjoy gambling?
A2. Is restless or irritable when attempting to cut down or stop gambling.	In the past year, have you tried to stop, cut down, or control your gambling? *(if yes)* In the past year, on one or more of the times you tried to stop, cut down, or control your gambling, were you restless or irritable?	Have you tried to spend less time or money gambling? *(if yes)* Did this make you feel restless or irritable?
A3. Has made repeated unsuccessful attempts to control, cut back, or stop gambling.	In the past year, have you tried to stop, cut down, or control your gambling? *(if yes)* In the past year, have you tried but not succeeded in stopping, cutting down, or controlling your gambling? *(if yes)* In the past year, has this happened three or more times?	Have you tried to spend less time or money gambling? *(if yes)* Have you tried this more than once? *(if yes)* Have you been able to spend less time or money gambling?
A4. Is often preoccupied with gambling (e.g., having persistent thoughts of reliving past gambling experiences, handicapping or planning the next venture, thinking of ways to get money with which to gamble).	In the past year, have there been any periods lasting two weeks or longer when you spent a lot of time thinking about your gambling experiences or planning future gambling ventures or bets? *(And/or)* In the past year, have there been any periods lasting two weeks or longer when you spent a lot of time thinking about ways of getting money to gamble with?	Do you spend a lot of time thinking about gambling? *(And/or)* Do you spend a lot of time thinking about ways of getting money to gamble with?

Table 3 (Continued)

A5. Often gambles when feeling distressed (e.g., helpless, guilty, anxious, depressed).	In the past year, have you gambled as a way to escape from personal problems? *(And/or)* In the past year, have you gambled to relieve uncomfortable feelings such as guilt, anxiety, helplessness, or depression?	Do you often gamble when you are distressed?
A6. After losing money gambling, often returns another day to get even ("chasing" one's losses).	In the past year, has there ever been a period when, if you lost money gambling on one day, you would often return another day to get even?	Do you gamble to win back money you have lost gambling?
A7. Lies to conceal the extent of involvement with gambling.	In the past year, have you more than once lied to family members, friends, or others about how much you gamble or how much you lost on gambling? *(And)* *Has this happened three or more times?*	Do you lie about how much you gamble?
A8. Has jeopardized or lost a significant relationship, job, or educational or career opportunity because of gambling.	In the past year, has your gambling caused you serious or repeated problems in your relationships with any of your family members or a friend? *(Or)* In the past year, has your gambling caused you any problems in school, such as missing classes or days of school or getting worse grades? *(Or)* In the past year, has your gambling caused you to lose a job, have trouble with your job, or miss out on an important job or career opportunity?	Has your gambling caused problems with people that you care about? *(Or)* Has your gambling caused problems at school? *(Or)* Has your gambling caused problems at work?

Table 3 (Continued)

| A9. Relies on others to provide money to relieve desperate financial situations caused by gambling. | In the past year, have you needed to ask family members or anyone else to loan you money or otherwise bail you out of a desperate money situation that was largely caused by your gambling? | Has gambling caused you to have money problems? *(And)* Have you received money from other people to help with these problems? |
| Has committed illegal acts such as forgery, fraud, theft, or embezzlement to finance gambling.[1] | In the past year, have you written a bad check or taken money that didn't belong to you from family members or anyone else in order to pay for your gambling? | Not assessed. |

Note. 1. The "illegal acts" criterion was present in the DSM-IV criteria for pathological gambling, but not in the DSM-V criteria for gambling disorder.

Table 4 – Readability of MGM and NODS

Measure	Number of items	Average Syntactic Simplicity	Number of difficult items (SS[1])	Number of difficult items (WMO[2])
MGM	15	1.91	0	0
NODS	17	-0.30	11	8

Note. SS = Syntactic Simplicity; WMO = Working Memory Overload. 1. An item was
considered difficult based on its syntactic simplicity if Coh-Metrix identified the item as
having a negative syntactic simplicity z-score, which indicated that it was more
syntactically complex than the average sentence a student would encounter during their
k-12th grade educational career. 2. An item was considered difficult based on its working
memory load if QUAID identified it as likely to overload the reader's working memory.

Table 5. *Differences between NODS and MGM Readability and the Percentage of Diagnostic*

Criteria Correctly Identified

DSM Criteria	MGM Readability	NODS Readability	Readability Difference	MGM Percent Correctly Identified	NODS Percent Correctly Identified	Difference Percent Correctly Identified
A1	0.37	-1.635	2.006	82	90	-8
A2	2.7995	-0.4525	3.252	75	66	9
A3	1.517667	0.926	0.591667	75	66	9
A4	0.867	-1.2245	2.0915	93	93	0
A5	1.393	0.134	1.259	32	24	8
A6	1.298	-1.832	3.13	91	83	8
A7	1.658	1.2525	0.4055	61	54	7
A8	3.592	-0.66833	4.260333	93	85	8
A9	1.921	0.163	1.758	86	80	6

Note. The items of each measure were mapped onto the DSM criteria they assess for, then the

average readability of these items, operationalized using Coh-Metrix's Syntactic

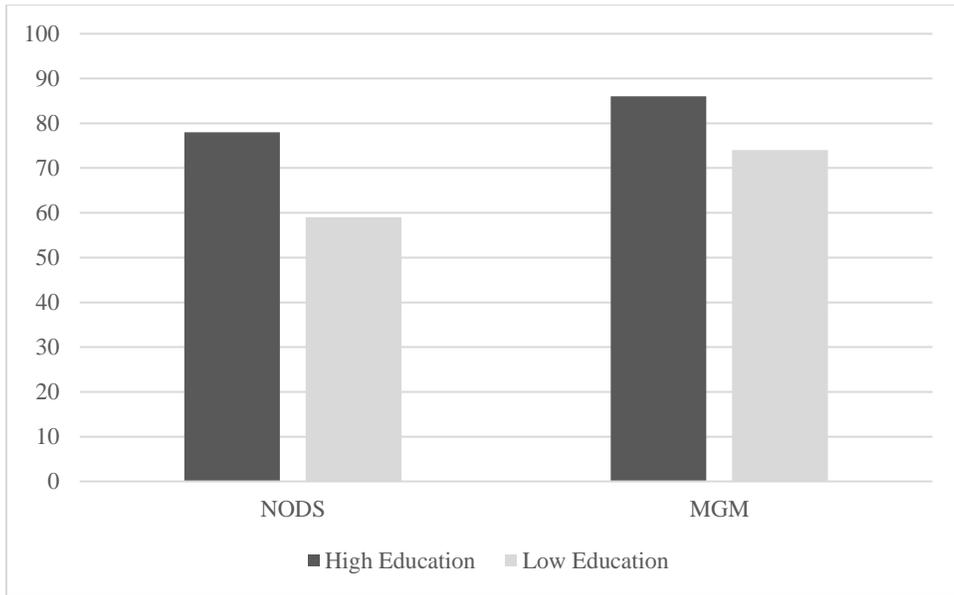Simplicity composite score, was calculated.

Figure 1

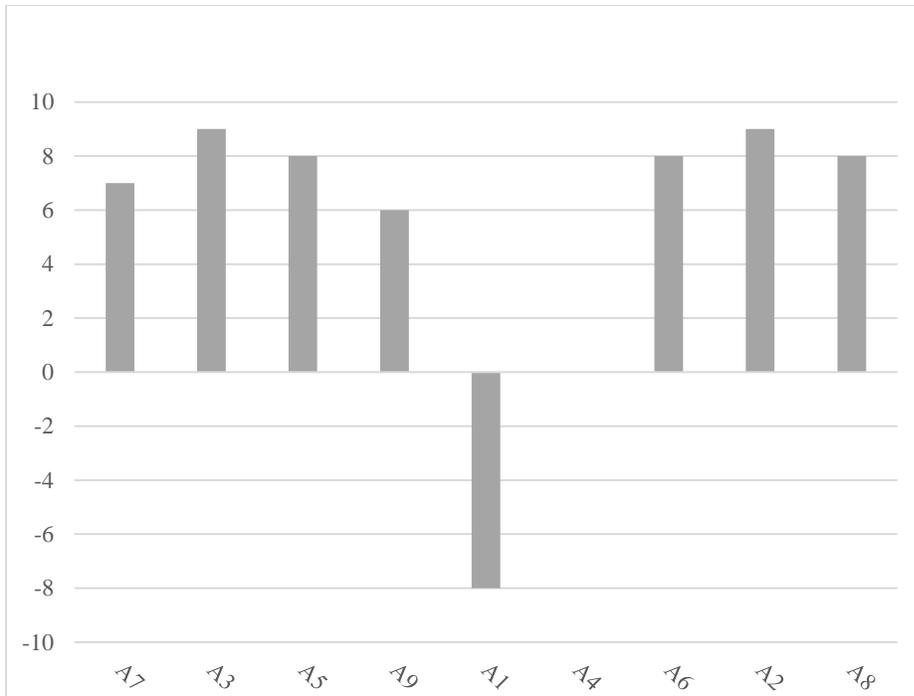*Percent of Items Answered Correctly by Experimental and Education Groups*

Figure 2

Size of Questionnaire Effect ordered by Difference in Readability of Questions Mapped onto

DSM Criteria

Note. Information for this figure was taken from Table 5, columns 4 ("Readability Difference"), and 7 ("Difference in Percent Correctly Identified"). The x-axis was ordered by the magnitude of the difference (Table 5, column 4), and the y-axis represents the difference in the percent of individuals who were able to correctly identify the presence or absence of the DSM criteria. A value > 0 indicates that individuals who completed the MGM correctly identified the presence or absence of the diagnostic criteria at a higher rate than individuals who completed the NODS. The strength of the questionnaire effect did not appear to increase linearly as a product of the difference in the readability of the items mapped onto DSM criteria.

Appendix A

Study Vignette

Michael went to the casino for the first time two years ago. He quickly discovered his love for card games like Texas Hold'em and Blackjack. At first, he only gambled about once a month, bringing $20 to the casino with him each time. In the past year, however, he has been gambling with his friends every Saturday night. He has had to bring $100 with him in order to get the same enjoyment out of the experience. He has so much fun gambling; he gambles every Saturday night because he doesn't want to miss the opportunity to see his friends. He spends weeks looking forward to these gambling ventures and thinking of ways of getting money to gamble with. Michael has never attempted to limit his gambling.

Sometimes after a big loss, he has fun returning the next day to try to win the money back that he had lost. Since he started gambling more regularly he has been having trouble paying his bills and is distracted at work and school. His boss and teachers have noticed that he hasn't been performing as well as he used to. His family is aware of this; Michael has always been very honest with them about his gambling. They give him money to help him with his bills as much as they can. Some of his family members think he should stop gambling, which leads to arguments at home. Recently, Michael had to drop out of school because he cannot afford tuition, which is largely due to the amount of money he has lost gambling. Now he is unable to get the job he wants because he did not get his degree. Michael is starting to think about stopping gambling.

Coh-Metrix Narrativity Z-Score (Percentile): 1.75 (95.99)