

Assessing DSM-IV symptoms of panic attack in the general population: An item response analysis

Author:

Sunderland, Matthew; Hobbs, Megan J.; Andrews, Gavin; Craske, Michelle G.

Publication details:

Journal of Affective Disorders

v. 143

Chapter No. 1-3

pp. 187-195

0165-0327 (ISSN)

Publication Date:

2012

Publisher DOI:

<http://dx.doi.org/10.1016/j.jad.2012.06.001>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/52558> in <https://unsworks.unsw.edu.au> on 2024-03-28

Word Count: 4,598 (excluding abstract and references).

**Assessing DSM-IV symptoms of panic attack in the general population: An item
response analysis**

Matthew Sunderland¹

Megan J. Hobbs¹

Gavin Andrews¹

Michelle G. Craske²

1. Clinical Research Unit for Anxiety and Depression (CRUfAD), School of Psychiatry, University of New South Wales, Sydney, Australia.
2. Department of Psychology, University of California, Los Angeles, CA.

Corresponding Author: Dr Matthew Sunderland, CRUfAD, Level 4 O'Brien Centre, St Vincent's Hospital, 394-404 Victoria Street, Darlinghurst, NSW, 2010, Australia. Ph: +612 8382 1437. Email: matthews@unsw.edu.au

Running title: Assessing DSM-IV symptoms of panic

ABSTRACT

Background: Unexpected panic attacks may represent a non-specific risk factor for future depression and anxiety disorders. The examination of panic symptoms and associated latent severity levels may lead to improvements in the identification, prevention, and treatment of panic attacks and subsequent psychopathology for ‘at risk’ individuals in the general population.

Methods: The current study utilised Item Response Theory to assess the DSM-IV symptoms of panic in relation to the latent severity level of the panic attack construct in a sample of 5,913 respondents from the National Epidemiologic Survey on Alcohol and Related conditions. Additionally, differential item functioning (DIF) was assessed to determine if each symptom of panic targets the same level of latent severity between different sociodemographic groups (male/female, young/old).

Results: Symptoms indexing ‘choking’, ‘fear of dying’, and ‘tingling/numbness’ are some of the more severe symptoms of panic whilst ‘heart racing’, ‘short of breath’, ‘tremble/shake’, ‘dizzy/faint’, and ‘perspire’ are some of the least severe symptoms. Significant levels of DIF were detected in the ‘perspire’ symptom between males and females and the ‘fear of dying’ symptom between young and old respondents.

Limitations: The current study was limited to examining cross-sectional data from respondents who had experienced at least one panic attack across their lifetime.

Conclusions: The findings of the current study provide additional information regarding panic symptoms in the general population that may enable researchers and clinicians to further refine the detection of ‘at-risk’ individuals who experience threshold and sub-threshold levels of panic.

Key words: panic symptoms; anxiety; Item Response Theory; item bias; assessment.

1. INTRODUCTION

A panic attack is a sudden rush of intense fear or distress associated with four or more of a list of thirteen somatic and cognitive symptoms. Panic attacks are associated with nearly all the DSM-IV anxiety disorders but their regular unexpected occurrence is considered a core feature of DSM-IV Panic disorder. Approximately 28% of Americans experience panic attacks at some point in their life with approximately 23% experiencing panic attacks without ever meeting criteria for Panic disorder and/or Agoraphobia (Kessler et al., 2006). Individuals who experience isolated panic attacks without meeting full diagnostic criteria for Panic disorder and/or Agoraphobia have been shown to experience higher levels of impairment, distress, comorbidity and health service utilisation compared to respondents who have never experienced an attack (Kinley et al. , 2009; Klerman et al., 1991). As a result, researchers have concluded that panic attacks may represent a nonspecific risk factor for psychopathology and that the severity and frequency of the attacks may predict the onset of other anxiety and mood disorders (Baillie & Rapee, 2005; Goodwin et al., 2004; Kessler et al., 2006; Kinley et al., 2011). Therefore the accurate examination of panic symptoms and assessment of latent panic severity may lead to improvements in the identification, prevention, and treatment of panic attacks and subsequent psychopathology in high risk individuals of the population (Batelaan et al., 2010), and may reduce the associated burden on the individual and public health care system (Roy-Byrne et al., 2006; Skapinakis et al., in press).

The DSM-IV defines a panic attack with respect to thirteen symptoms. These are: 1) shortness of breath, 2) heart racing, 3) tremble/shake, 4) perspire, 5) choking, 6) dizzy/faint, 7) feel unreal/detached, 8) tingling/numbness, 9) hot flushes/chills, 10) nauseous, 11) chest pain, 12) go crazy/lose control, and 13) fear of dying. The DSM-IV assumes that these symptoms are manifest indicators or markers of severity of an underlying latent category

(panic attack). This categorical approach assumes that all symptoms are equal indicators of latent panic severity and that endorsing four or more symptoms provides sufficient evidence that the disorder is severe enough to warrant a diagnosis. However this polythetic approach ignores the extensive heterogeneity observed amongst individuals in relation to the number and content of symptoms endorsed.

Recently, research has focused on expanding the DSM-IV binary classification of panic by investigating important variations amongst individuals with respect to their symptoms and severity levels in the general population. Ietsugu, Sukigara, and Furukawa (2007) were the first to assess symptoms of panic in relation to latent severity levels using Item Response Theory (IRT) in the general population. The IRT model of psychopathology differs from the categorical DSM-IV model as it does not assume that latent panic severity can be classified into classes based on a set symptom threshold. Instead, IRT assumes that the symptoms of panic are markers of an underlying *dimension* of severity and each symptom indexes a different point of severity. This assumption provides further information regarding the latent severity of panic attacks based on the number and, more importantly, the content of symptoms endorsed.

Ietsugu et al. (2007) demonstrated that ‘tingling/numbness’ and ‘choking’ were good indicators of severe panic attacks whilst ‘short of breath’, ‘chest pain’, ‘dizzy/faint’, and ‘heart racing’ were markers of moderate panic attacks and could discriminate well between non-cases and those with low panic severity. However, in order to apply these results across various populations of the community it first must be assumed that the symptoms of panic display adequate measurement invariance. Measurement invariance is important because it ensures that each symptom targets the same level of severity regardless of any external factors such as age, sex, or ethnicity. If evidence for measurement invariance is not shown then using one set of symptom criteria to define panic across the population may distort the

severity level of panic for individuals belonging to a specific sociodemographic group (Embretson & Reise, 2000; Teresi & Fleishman, 2007).

There are several empirical reasons to suspect that the invariance assumption may not hold particularly across gender and age in the general population. Epidemiological studies have consistently indicated that females are more likely to experience panic across their lifetime in comparison to males (Goodwin et al., 2005; Kessler et al., 2006; Skapinakis et al., in press). Females are also more likely to endorse respiration related symptoms of panic (short of breath, choking) compared to males (Sheikh et al., 2002). However, it is difficult to interpret these findings as representing either a true difference in the severity of panic attacks between males and females or whether differing physiological factors related to sex are contributing to the observed differences in symptom endorsement. Also, differences have been found in panic phenomenology between old and young respondents with researchers suggesting that various psychological and social factors relating to the normal aging process may contribute to these differences (Sheikh et al., 2004; Wolitzky-Taylor et al., 2010). One important benefit of IRT is that it allows for the evaluation of panic symptoms across sex and age while controlling for the underlying severity on the panic attack trait. It achieves this by equating each group on the underlying severity level prior to examining any differences in item responses. This should indicate whether any differences between the sexes or age groups are due to some other explanation related to sex or age instead of actual differences in underlying panic attack severity.

The purpose of the current study is to investigate the DSM-IV panic symptoms and the associated latent levels of severity using IRT. The current study will attempt to replicate the findings by Ietsugu et al. (2007) using another large epidemiological sample of the U.S. general population. Additionally, the current study will aim to investigate the invariance

assumption of panic symptoms across sex and age using the Item Response Theory log-likelihood ratio procedure of differential item functioning (IRTLR-DIF).

2. METHOD

2.1 Sample

Data were from a nationally representative household survey known as the 2001-2002 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC wave 1). The NESARC sampled 43,093 non-institutionalised U.S civilians aged 18 and over with 43% being male and a mean age of 46 (SD=18) in the total population. A multistage stratified sampling design was used to stratify according to certain sociodemographic characteristics of the U.S. general population with over-sampling of non-Hispanic Black households and young adults aged 18-24 (see Grant et al., 2003a for more information on the complex sampling design and methodology of the NESARC). As a result of the skip questions included in the survey, data were collected on a subsample of respondents who skip into the panic disorder module. Therefore the current study examined the subsample of 5,913 respondents.

2.2 Diagnostic Assessment

The lifetime presence of thirteen DSM-IV panic attack symptoms was assessed using the Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS). This fully structured lay-administered interview possesses good to fair psychometric properties for the common mental disorders (Grant et al., 2003b). The AUDADIS contains a stem and leaf system designed to reduce overall administration time/burden. This system uses core symptoms or features of each disorder to initially screen for the overall presence of a diagnosis, e.g. if a respondent fails to endorse the core symptoms/features they skip from the remaining questions in the disorder module and move on to the next module.

The panic disorder module first screens for the potential presence of at least one panic attack that must have occurred out of the blue or unexpectedly across the respondent's lifespan. If the respondent endorses having experienced at least one unexpected panic attack in their life they are then administered thirteen additional questions with each question relating to a specific DSM-IV symptom of panic. Symptoms assessed include: 1) shortness of breath, 2) heart racing, 3) tremble/shake, 4) perspire, 5) choking, 6) dizzy/faint, 7) feel unreal/detached, 8) tingling/numbness, 9) hot flushes/chills, 10) nauseous, 11) chest pain, 12) fear of going crazy/lose control, and 13) fear of dying.

2.3 Statistical Analysis

The analysis of DSM-IV panic symptoms and the measurement invariance assumption proceeded in a series of three phases. The first phase involved investigating the dimensionality of panic symptoms in relation to the unidimensionality assumption of IRT (i.e. all symptoms must load significantly on one underlying latent dimension). The second phase involved modelling each panic symptom in the entire sample using the two-parameter logistic item response model. The final phase involved examining the invariance assumption across sex and age for each panic symptom using IRTLR-DIF (Thissen, 2001).

The unidimensionality assumption of IRT was assessed in the first phase of the analysis using confirmatory factor analysis (CFA) of tetrachoric correlations with a robust weighted least squares mean and variance adjusted estimator (WLSMV), as suggested for the analysis of binary data by Muthén and Muthén (2010). A one factor model with all thirteen symptoms loading on one latent panic attack severity variable was estimated and model fit was assessed using a series of fit indices: Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Fit Index (TLI; Tucker & Lewis, 1973), and the root mean square error of approximation (RMSEA; Steiger, 1990). Good model fit was determined using pre-defined

cut points. Current recommendations suggest that TLI and CFI values ≥ 0.90 indicate acceptable fit and values ≥ 0.95 imply very good fit (Hu & Bentler, 1998; Vandenberg & Lance, 2000). RMSEA values less than 0.05 indicate close model fit, values up to 0.08 indicate reasonable model fit, and values exceeding 0.10 indicate poor fit (Browne & Cudeck, 1993; Jöreskog & Sörbom, 1993). The CFA and model fit statistics were generated using Mplus version 6 (Muthén & Muthén, 2010).

The 2-parameter logistic item response (2-PL) model was utilised in the second phase to examine the performance of the thirteen DSM-IV panic symptoms in the total population of respondents who have experienced at least one panic attack. Briefly, the 2-PL model describes the probability of a response to each symptom in the general population as a monotonically increasing function of a normally distributed latent variable, which was identified in the CFA previously. In this instance the latent variable represents panic attack latent severity levels and therefore the more severe the latent severity level the more likely the respondent will endorse the DSM-IV symptoms. The 2-PL model describes the performance of each symptom through the use of two item parameters, known as the *a*-parameter or discrimination and the *b*-parameter or difficulty. Each parameter describes a certain feature of the Item Characteristic Curve (ICC) that has been described in detail (see Embretson & Reise, 2000). The ICCs for the thirteen DSM-IV panic symptoms were estimated using Mplus version 6 (Muthén & Muthén, 2010).

The third and final phase of the overall analysis involved testing the invariance assumption for the thirteen panic symptoms between male and female and young (≤ 45) and old (> 45) respondents. This was achieved using separate IRTLR-DIF analyses. Briefly, IRTLR-DIF works by matching (or anchoring) two groups (male/female or young/old) on the panic attack latent severity level prior to estimating the *a*-parameters and *b*-parameters for each symptom separately in both groups. If the symptom is invariant across the groups then

the a - and/or b -parameters will be equivalent however if the symptom exhibits non-invariance or DIF then the a - and/or b -parameters will be significantly different. The presence of DIF in the a -parameters indicates that certain symptoms for one group discriminate better between the respondents than for the other group whilst the presence of DIF in the b -parameters indicates that certain symptoms for one group have an increased or decreased probability of response in comparison to the other group despite being match in terms of severity. To test for a significant difference in item parameters across groups a model is first estimated constraining the item parameters to be equal across both groups. A second model is then estimated allowing the item parameters to freely vary between the two groups. The loglikelihoods for both models are evaluated and the significance of the ratio statistic or G^2 statistic is calculated using a χ^2 distribution with 1 degree of freedom. If the G^2 statistic is significant then this provides evidence that the item parameters exhibit DIF between the two groups and the symptom is therefore considered non-invariant (Thissen, 2001).

As mentioned above, the two groups must first be matched on the underlying panic latent severity level. To match the two groups, an anchor set of items that do not exhibit any DIF must be used prior to investigating the potential for DIF in the remaining test items. To select the anchor set, an iterative purification procedure was implemented that first analysed the entire set of symptoms for DIF using every other symptom as a temporary anchor set. Those items that exhibit DIF in either the a - or b -parameter were removed from the analysis and the remaining DIF-free items were re-examined. If the items remain DIF-free then they are assumed to be the anchor set for the following DIF analysis (Edelen et al., 2006). A simulation study found that purification of the anchor items prior to DIF detection significantly improved the accuracy rate and reduced false positive results, particularly in IRT-based methods of DIF detection (Navas-Ara & Gomez-Benito, 2002).

The final step of the DIF analysis involves conducting the analysis using the DIF-free anchor set to match the two groups on the underlying latent variable prior to estimating the item parameters separately for both groups in the remaining test symptoms. If the omnibus G^2 statistic was significant at this point, indicating that at least one of the item parameters significantly differs between the two groups, then follow-up tests were conducted to examine if the DIF resides in the a -parameter, b -parameter or both. It was at this stage that the Benjamini-Hochberg method (Benjamini, & Hochberg, 1995; Thissen et al., 2002) was used to adjust the critical p-value for multiple comparisons to reduce the possibility of type I error. For the final item parameter estimation, item parameters of symptoms that were deemed to have significant DIF after Benjamini-Hochberg adjustment were estimated separately in both groups whilst item parameters that did not exhibit DIF were constrained to be equal in both groups. The final estimation of item parameters was conducted using MULTILOG software package version 7.03. For a more detailed description on the IRTLF-DIF procedure used in the current study see Teresi et al. (2007).

3. RESULTS

3.1 Dimensionality

According to the predetermined cut points, the fit statistics estimated with the one factor CFA model exhibited acceptable model fit with CFI = 0.94 and TLI = 0.93. Likewise, the RMSEA value indicated very close model fit between the observed and expected tetracoric correlations with a value of 0.05. The standardised factor loadings for the one factor model are presented in Figure 1. All factor loadings were significant at the 0.05 level and were moderately to highly correlated with the underlying latent variable of panic attack severity. This justified the use of unidimensional IRT models (such as the 2PL model) in comparison to multi-dimensional models.

3.2 IRT Analyses

The factor loadings and thresholds estimated in the CFA above were re-parametrized to generate IRT item parameters for the 2PL model in the total sample of respondents who have experienced at least one panic attack in their lifetime. Figure 2 provides a visual summary of the item parameters for all thirteen DSM-IV panic symptoms in the form of ICCs. Assuming that 99% of the population lies within three standard deviations of the mean, according to the normal distribution, the figure indicates that the majority of DSM-IV panic symptoms clearly discriminate across the population with moderate to high discrimination levels. Albeit, the majority of panic symptoms appear to target panic attack latent severity levels at the low to moderate end of the spectrum as evidenced by negative or small positive *b*-parameters (i.e. the majority of symptoms situated slightly left or right of the mean in Figure 2). The symptom addressing ‘short of breath’ appears to be the most discriminating. However ‘shortness of breath’ along with ‘heart racing’, ‘tremble/shake’, ‘dizzy/faint’, and ‘perspire’ were some of the least severe panic symptoms. On the other end of the spectrum, the symptoms addressing ‘choking’, ‘fear of dying’, and ‘tingling/numbness’ are some of the more severe symptoms of panic attack indicating that only respondents who experience severe panic attacks will endorse these symptoms. For example, at the mean panic attack severity level, approximately 20% of respondents will endorse the symptom representing ‘choking’.

The Item Information Functions (IIF) presented in Figure 2 provide further evidence that four symptoms, ‘heart racing’, ‘short of breath’, ‘go crazy/lose control’, and ‘choking’, provide the most information/precision about various points on the full panic attack latent severity level. Interpreting information can be better understood by knowing the relationship between information and the standard error of measurement. For example the smaller the standard error the larger the amount of information that item provides. Therefore, knowing if

the respondent endorses one or more of these symptoms will provide valuable information regarding their panic attack latent severity level.

3.3 Measurement Invariance

3.3.1 Sex

During the purification stage, two symptoms (feel unreal/detached, hot flushes/chills) were identified as the DIF-free anchor set and were used to match males and females on panic attack latent severity when examining DIF in the remaining symptoms. There was no indication of DIF related to the a -parameters indicating that all symptoms possessed equivalent discriminating power between males and females. Prior to Benjamini-Hochberg adjustment, six symptoms displayed significant DIF in the b -parameters. However, only three (short of breath, perspire, tingling/numbness) remained significant after adjusting the critical p-value for multiple comparisons. Inspection of the ICCs for the three significant symptoms in Figure 3 reveals that the DIF associated with ‘short of breath’ and ‘tingling/numbness’ appears to be relatively minor in terms of magnitude, whereas the DIF associated with ‘perspire’ appears to be of greater concern. The b DIF associated with ‘perspire’ indicated that overall females tend endorse the symptom at a much lower rate in comparison to males despite being matched in terms of panic attack latent severity levels.

3.3.2 Age

During the purification stage, one symptom (dizzy/faint) displayed no indication of DIF and was therefore selected to form the DIF-free anchor set for the analysis of the remaining symptoms. Much like the previous analysis, there was no indication of a DIF between young and old respondents, indicating that all the symptoms display equivalent discriminating power across the age groups. Only one symptom (fear of dying) displayed

significant levels of *b*DIF after adjusting for multiple comparisons using the Benjamini-Hochberg method. Examining the ‘fear of dying’ ICC for young and old respondents reveals that the old respondents are more likely to endorse the ‘fear of dying’ symptom in comparison to young respondents despite being matched in terms of panic attack latent severity levels. This indicates that the ‘fear of dying’ symptom is not invariant between young and old respondents and ‘fear of dying’ should be interpreted as reflecting a more severe panic attack in young respondents in comparison to older respondents.

4. DISCUSSION

The current study assumes that the DSM-IV symptoms used in the analysis are indicators of one underlying dimension and that these symptoms are the most widely known and used manifest indicators of latent panic attack severity. Having said that, there may be many additional symptoms and even biomarkers that likewise target this dimension, possibly in a better manner, however an investigation of additional symptoms or biomarkers of panic severity was beyond the scope of the current manuscript. The main intention of the current study was to therefore provide additional information regarding the DSM-IV symptoms and how they can be interpreted for the assessment of panic severity in the general population as opposed to simply applying the symptom threshold proposed by DSM-IV.

The findings of the current study replicate those found previously by Ietsugu, et al. (2007), namely that ‘choking’, ‘fear of dying’, and ‘tingling/numbness’ represent severe attacks whilst ‘short of breath’, ‘heart racing’, ‘tremble/shake’, ‘dizzy/faint’, and ‘perspire’ represent attacks of lesser severity. The strong replication of Ietsugu et al., (2007) represents a greater generalizability of the findings and enables researchers and clinicians to confidently use the information reported in both studies to inform their assessment of panic disorder in the general population. Despite this, further replication of the current findings is required,

particularly in population data from non-westernized and developing countries as well as in clinical populations of treatment seeking individuals. These findings are rather timely given the revisions for the DSM-IV and ICD-10 are underway.

The current study further extends the previous findings of Ietsugu et al., (2007) by demonstrating, through the use of DIF analysis, that full measurement invariance of the panic symptoms between sex and age cannot be assumed. Instead, a partial invariance model is more appropriate with ‘perspire’ and to a lesser extent ‘short of breath’ and ‘tingling/numbness’ exhibiting non-invariance between male and female whilst ‘fear of dying’ exhibits non-invariance between young and old respondents. The current study does, however, utilise a large sample size that could indicate minor differences between the targeted groups IRT parameters as statistically significant yet they are not large enough to be clinically meaningful. The idea of whether a difference is clinically meaningful or not is arbitrary but it was decided that a difference of approximately 10% or greater in the endorsement of each symptom at the same severity level would be clinically meaningful. Based on this assumption, the endorsement of ‘perspire’ displayed a clinically meaningful difference between males and females whereas the endorsement of ‘fear of dying’ displayed a clinically meaningful difference between young and old.

The finding of partial invariance has strong implications for the interpretation of panic attack latent severity experienced between the sexes and age groups with respect to the specific symptoms endorsed by each group. For example, take into consideration a male and a female who, based on prior evidence, have the same level of latent panic attack severity however the assessment of DSM-IV panic symptoms reveals that the female respondent has endorsed the same symptoms as the male respondent except for the ‘perspire’ symptom. Assuming full measurement invariance between sexes, a clinician may conclude that the male respondent exhibits attacks at a slightly higher severity level in comparison to the female

respondent. On the other hand, taking into account the results found in the current study a clinician can reasonably conclude that male respondents are more likely to endorse the 'perspire' symptom than females in the general population despite being matched in terms of panic severity, and therefore the two individuals are likely to experience panic attacks of similar severity.

A variety of factors could contribute to the finding of DIF amongst the observed groups. One of the more likely explanations for the DIF found in the 'fear of dying' symptom between young and old respondents may be related to the perception of death in older age. For instance, it seems more likely that death is a possibility during the later years of life; therefore older respondents may be more likely to endorse the fear of dying in comparison to younger respondents despite being matched in terms of severity. Likewise, younger respondents may feel that death, such as the possibility of having a heart attack, is unlikely and therefore do not endorse that they fear dying as much as older respondents.

Similar reasoning could be applied to the DIF observed in the 'perspire' symptom between males and females. It may be more socially acceptable for males to sweat during times of fear in comparison to females. Alternatively, female respondents may feel embarrassed to admit that they perspire in comparison to males and therefore exhibit a bias towards not reporting the 'perspire' symptom. In any case, the suggested reasons for the observed DIF between sex and age in the current study are purely speculative. Further investigations are required to empirically investigate the actual cause of DIF by applying cognitive interviewing techniques to elicit the respondent's cognitions and justification for endorsing or not endorsing particular symptoms (Willis, 2005).

Once DIF has been identified there are several recommendations for dealing with items that exhibit DIF. One of the prevailing suggestions is to remove the offending

symptoms from the assessment battery. The rationale for this is that the symptom that exhibits DIF is not working as intended and therefore the use of these symptoms may lead to erroneous judgements and diagnostic mistakes (Embretson & Reise, 2000). However removing the symptoms entirely also results in a loss of information, particularly about how different individuals may interpret and respond to their symptoms of panic in a different manner. An alternative suggestion is to retain the symptoms but investigate the potential sources of the DIF using the exploratory and qualitative methods suggested above. Once the source of the DIF has been identified, the symptom criteria can then be revised or reformatted to reduce the individual variations in the interpretation of panic symptoms. This option appears to be particularly pertinent and offers a significant opportunity for additional research since the existing DSM-IV diagnostic criteria are currently under revision with a primary goal being to reduce any sex and age differences that occur in the criteria (Narrow et al., 2007). A third option in contrast to the removal or revision of symptoms would be to simply leave the symptoms as they currently are and instead use the results presented here to place more or less weight on symptoms that exhibit DIF depending on the specific sociodemographic features of the individual under assessment.

Regardless of how the items that exhibit DIF are treated, the overall results of the current study indicate that the majority of items assessing panic attack severity do not exhibit DIF and therefore can be considered invariant between males and females and young and old. The implication of this finding is that any observed differences in the prevalence of panic attack between the sociodemographic groups in the population will represent a true difference in panic attack severity experienced by those groups as a whole. Additionally, the results presented in the current study can be used to improve the overall assessment of panic attack severity in the general population. In particular, the IRT analysis indicates that four symptoms, ‘heart racing’, ‘short of breath’, ‘go crazy/lose control’, and ‘choking’, provide

the most information regarding the full continuum of panic attack latent severity in respondents who experience at least one panic attack across their lifespan. Furthermore, these four symptoms show no indication of DIF between the examined sociodemographic groups. Therefore, quick and efficient screening of panic attack severity in the general population could be reliably based on the assessment of these four symptoms.

The results of the current study need to be interpreted with some limitations kept in mind. The first limitation involves the use of the skip questions in the AUDADIS interview to filter out individuals who have never experienced a panic attack. This limits the current study to an investigation of a subsample of 5,913 respondents who endorse the skip questions. Although, this may not be entirely problematic since the current study aimed to assess the symptoms of panic and by definition any individual who experiences symptoms of panic must also have experienced at least one panic attack at some point in their lives. A second limitation involves the self-report and retrospective nature of the assessment of panic symptoms in the NESARC survey. The current study must assume that the survey respondents can adequately recall their experiences of panic no matter how long ago they may have experienced them. Moreover, a recent review of the literature has indicated that there is a high level of variability in the endorsement of panic symptoms across multiple self-monitored panic attacks experienced over time (Craske et al., 2010). A more accurate approach in the future would be to analyse panic symptom data from prospective longitudinal studies to remove any influence of recall bias. However, these studies are rare and often costly to administer, particularly at a general population level that requires the collection of very large samples. A third limitation involves the use of small anchor sets (consisting of only one and two anchor items) in the DIF analysis. A simulation study indicated that, whilst the use of one item as an anchor set provides an acceptable level of power, the pure anchor set used to equate the measurement scale between the two groups should ideally contain four

or more items (Wang & Yeh, 2003). However, the use of small pure anchor sets appears to be favourable over the use of anchor sets that may contain DIF items, which could distort the results if used to equate the measurement scale (Wang & Yeh, 2003; Woods, 2009). As such, it was decided to maintain the use of one and two item anchors instead of increasing the chance of including additional items in the anchor that have some indication of DIF.

In summary, the findings of the current study demonstrate several important factors regarding the symptoms of panic attack. First, the DSM-IV symptoms can be viewed as markers of one underlying dimension of latent panic attack severity. Second, the symptoms associated with ‘choking’, ‘fear of dying’, and ‘tingling/numbness’ represent severe attacks whilst ‘shot of breath’, ‘heart racing’, ‘tremble/shake’, ‘dizzy/faint’, and ‘perspire’ represent attacks of lesser severity. Third, four symptoms, ‘heart racing’, ‘short of breath’, ‘go crazy/lose control’, and ‘choking’, provide the most information across the full continuum of panic attack latent severity and could be used for quick and efficient screening in the population. Finally, the results indicate that a partial invariance model is more appropriate when comparing panic attack severity between age and sex, with ‘perspire’ and to a lesser extent ‘short of breath’ and ‘tingling/numbness’ exhibiting non-invariance between male and female whilst ‘fear of dying’ exhibits non-invariance between young and old. The findings of the current study further contribute to the overall understanding and assessment of panic attacks in general population.

REFERENCES

- Baillie, A. J., Rapee, R. M., 2005. Panic attacks as risk markers for mental disorders. *Soc Psychiatry and Psychiatr Epidemiol.* 40, 240-244.
- Batelaan, N. M., Smit, F., de Graaf, R., van Balkom, A. J. L. M., Vollebergh, W. A. M., Beekman, A. T. F., 2010. Identifying target groups for the prevention of anxiety disorders in the general population. *Acta Psychiatr Scand.* 122, 56-65.
- Benjamini, Y., & Hochberg, Y., 1995. Controlling for false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statistical Society, Series B.* 57, 289-300.
- Bentler, P. M., 1990. Comparative fit indexes in structural models. *Quant Meth Psychol.* 107, 238-246.
- Browne, M. W., Cudeck, R., 1993. Alternative ways of assessing model fit. In K. A. Bollen J. S. Long (Eds.), *Testing Structural Equation Models*. Newbury Park, CA: Sage.
- Craske, M. G., Kircanski, K., Epstein, A., Wittchen, H-U., Pine, D. S., Lewis-Fernandez, R., Hinton, D., Phillips, K., Hofmann, S., Klein, D., Klein, R., Rapee, R., Stein, M., 2010. Panic disorder: a review of DSM-IV panic disorder and proposals for DSM-V. *Depress Anxiety.* 27, 93-112.
- Edelen, M. O., Thissen, D., Teresi, J. A., Kleinman, M., Ocepek-Welikson, K., 2006. Identification of differential item functioning using item response theory and the likelihood-based model comparison approach. *Med Care.* 44, S134-S142.
- Embretson, S. E., Reise, S. P., 2000. *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Earlbaum Associates.
- Goodwin, R. D., Lieb, R., Hoefler, M., Pfister, H., Bittner, A., Beesdo, K., Wittchen, H-U.,

2004. Panic attack as a risk factor for severe psychopathology. *Am J Psychiatry*. 161, 2207-2214.
- Grant, B. F., Dawson, D. A., Stinson, F. S., Chou, P. S., Kay, W., Pickering, R., 2003b. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): Reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. *Drug Alcohol Depend*. 71, 7-16.
- Grant, B. F., Kaplan, K., Shepard, J., & Moore, T., 2003a. Source and Accuracy Statement for Wave 1 of the 2001-2002 National Epidemiologic Survey on Alcohol and Related Conditions. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.
- Hu, L. T., Bentler, P. M., 1998. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychol Meth*. 3, 424-453.
- Ietsugu, T., Sukigara, M., Furukawa, T. A., 2007. Evaluation of diagnostic criteria for panic attack using item response theory: findings from the National Comorbidity Survey in USA. *J Affect Disord*. 104, 197-201.
- Joreskog, K. G., Sorbom, D., 1993. LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language. Lincolnwood, IL: Scientific Software International.
- Kessler, R. C., Chiu, W. T., Jin, R., Ruscio, A. M., Shear, K., Walters, E. E., 2006. The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 63,415-424.
- Kinley, D. J., Cox, B. J., Clara, I., Goodwin, R. D., Sareen, J., 2009. Panic attacks and their relation to psychological and physical functioning in Canadians: results from a nationally representative sample. *Can J Psychiatry*. 54, 113-122.
- Kinley, D. J., Walker, J. R., Enns, M. W., Sareen, J., 2011. Panic attacks as a risk for later

- psychopathology: results from a nationally representative survey. *Depress Anxiety*. 28, 412-429.
- Klerman, G. L., Weissman, W. W., Ouellette, R., Johnson, J., Greenwald, S., 1991. Panic attacks in the community: social morbidity and health care utilization. *JAMA*. 265, 742-746.
- Muthén, L. K., Muthén, B. O., 2010. *Mplus Users' Guide Sixth Edition*. Los Angeles, CA: Muthén & Muthén
- Narrow, W. E., First, M. B., Sirovatka, P. J., Regier, D. A., 2007. *Age and Gender Considerations in Psychiatric Diagnosis: A Research Agenda for DSM-V*. Arlington, VA: American Psychiatric Publishing.
- Navas-Ara, M. J., Gomez-Benito, J., 2002. Effects of ability scale purification on the identification of DIF. *Eur J Psychol Assess*. 18, 9-15.
- Roy-Byrne, P. P., Craske, M. G., Stein, M. B., 2006. Panic disorder. *Lancet*. 368, 1023-1032.
- Sheikh, J. I., Leskin, G. A., Klein, D. F., 2002. Gender differences in panic disorder: findings from the national comorbidity survey. *Am J Psychiatry*. 159, 55-58.
- Sheikh, J. I., Swales, P. J., Carlson, E. B., Lindley, S. E., 2004. Aging and panic disorder: phenomenology, comorbidity, and risk factors. *Am J Geriatr Psychiatry*. 12, 102-109.
- Skapinakis, P., Lewis, G., Davies, S., Brugha, T., Prince, M., Singleton, N., in press. Panic disorder and subthreshold panic in the UK general population: epidemiology, comorbidity, and functional limitation. *Eur Psychiatry*.
- Steiger, J. H., 1990. Structural model evaluation and modification: an interval estimation approach. *Multivar Behav Res*. 25, 173-180.
- Teresi, J. A., Fleishman, J. A., 2007. Differential item functioning and health assessment. *Qual Life Res*. 16, supplement 1, 33-42.

- Teresi, J. A., Ocepek-Welikson, K., Kleinman, M., Cook, K. F., Crane, P. K., Gibbons, L. E., Morales, L. S., Orlando-Edelen, M., Cella, D., 2007. Evaluating measurement equivalence using the item response theory log-likelihood ratio (IRTLR) method to assess difference item functioning (DIF): applications to measures of physical functioning ability and general distress. *Qual Life Res*, 16, 43-68.
- Thissen, D., 2001. IRTLR DIF v2.0: Software for the computation of the statistics involved in item response theory likelihood-ratio tests for differential item functioning. Available from <http://www.unc.edu/~dthissen/dl.html>.
- Thissen, D., Steinberg, L., Kuang, D., 2002. Quick and easy implementation of the Benjamini-Hochberg procedure for controlling the false positive rate in multiple comparisons. *J Educ Behav Stat*. 27, 77-83.
- Tucker, L. R., Lewis, C., 1973. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 38, 1-10.
- Vandenberg, R. J., Lance, C. E., 2000. A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ Res Meth*, 3, 4-70.
- Willis, G. B., 2005. *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. Thousand Oaks, CA: Sage.
- Wang, W-C., Yeh, Y-L., 2003. Effects of anchor item methods on differential item functioning detection with the likelihood ratio test. *Appl Psych Meas*. 27, 479-498.
- Wolitzky-Taylor, K. B., Castriotta, N., Lenze, E. J., Stanley, M. A., Craske, M. G., 2010. Anxiety disorders in older adults: a comprehensive review. *Depress Anxiety*. 27, 190-211.
- Woods, C., 2009. Empirical selection of anchors for tests of differential item functioning. *Appl Psych Meas*. 33, 42-57.

TABLES

Table 1: Item Response parameters and tests of differential item functioning in DSM-IV panic symptoms for female and male respondents.

Panic Item	Group	a (SE)	b (SE)	aDIF(p)	bDIF(p)
Short of Breath	female	1.84 (0.06)	-0.67 (0.03)	0.0 (1.00)	14.1 (<0.01)
	male	1.84 (0.06)	-0.49 (0.04)		
Heart Racing	female	1.74 (0.07)	-1.36 (0.04)	NS	
	male	-	-		
Tremble/Shake	female	1.34 (0.05)	-0.20 (0.03)	NS	
	male	-	-		
Perspire	female	1.47 (0.06)	-0.11 (0.03)	2.7 (0.10)	60.6 (<0.01)
	male	1.47 (0.06)	-0.58 (0.05)		
Choking	female	1.55 (0.06)	0.93 (0.04)	1.0 (0.32)	5.9 (0.02)
	male	-	-		
Dizzy/Faint	female	1.67 (0.06)	-0.19 (0.03)	0.9 (0.34)	4.2 (0.04)
	male	-	-		
Feel unreal/Detached	female	1.14 (0.05)	0.28 (0.04)	Anchor	
	male	-	-		
Tingling/Numbness	female	1.52 (0.05)	0.75 (0.04)	3.3 (0.07)	8.8 (<0.01)
	male	1.52 (0.05)	0.59 (0.06)		
Hot flushes/Chills	female	1.67 (0.06)	0.18 (0.03)	Anchor	
	male	-	-		
Nauseous	female	1.09 (0.05)	0.50 (0.04)	0.1 (0.75)	4.5 (0.03)
	male	-	-		
Pain in chest	female	1.23 (0.05)	0.33 (0.04)	NS	
	male	-	-		
Go crazy/Lose control	female	1.08 (0.05)	0.51 (0.04)	NS	
	male	-	-		
Fear of dying	female	1.05 (0.05)	0.66 (0.05)	NS	
	male	-	-		

Note: Bold indicates the item parameters were significantly different after Benjamini-Hochberg adjustments. a = discrimination parameter, b = difficulty parameter, SE = standard error, aDIF = loglikelihood difference test for DIF in the discrimination parameter, bDIF = loglikelihood difference test for DIF in the difficulty parameter.

Table 2: Item Response parameters and tests of differential item functioning in DSM-IV panic symptoms for old and young respondents.

Panic Item	Group	a (SE)	b (SE)	aDIF(p)	bDIF(p)
Short of Breath	old	1.83 (0.07)	-0.41 (0.03)	NS	
	young	-	-		
Heart Racing	old	1.74 (0.07)	-1.16 (0.04)	NS	
	young	-	-		
Tremble/Shake	old	1.34 (0.05)	0.01 (0.03)	NS	
	young	-	-		
Perspire	old	1.39 (0.05)	-0.05 (0.03)	NS	
	young	-	-		
Choking	old	1.53 (0.06)	1.15 (0.04)	NS	
	young	-	-		
Dizzy/Faint	old	1.65 (0.06)	0.02 (0.03)	Anchor	
	young	-	-		
Feel unreal/Detached	old	1.14 (0.05)	0.49 (0.04)	NS	
	young	-	-		
Tingling/Numbness	old	1.49 (0.06)	0.94 (0.04)	NS	
	young	-	-		
Hot flushes/Chills	old	1.66 (0.06)	0.39 (0.03)	NS	
	young	-	-		
Nauseous	old	1.08 (0.05)	0.71 (0.04)	NS	
	young	-	-		
Pain in chest	old	1.22 (0.05)	0.54 (0.04)	NS	
	young	-	-		
Go crazy/Lose control	old	1.08 (0.05)	0.72 (0.04)	NS	
	young	-	-		
Fear of dying	old	1.12 (0.04)	0.59 (0.05)	1.7 (0.19)	26.3 (<0.01)
	young	1.12 (0.04)	1.06 (0.05)		

Note: Bold indicates the item parameters were significantly different after Benjamini-Hochberg adjustments. a = discrimination parameter, b = difficulty parameter, SE = standard error, aDIF = loglikelihood difference test for DIF in the discrimination parameter, bDIF = loglikelihood difference test for DIF in the difficulty parameter.

FIGURES

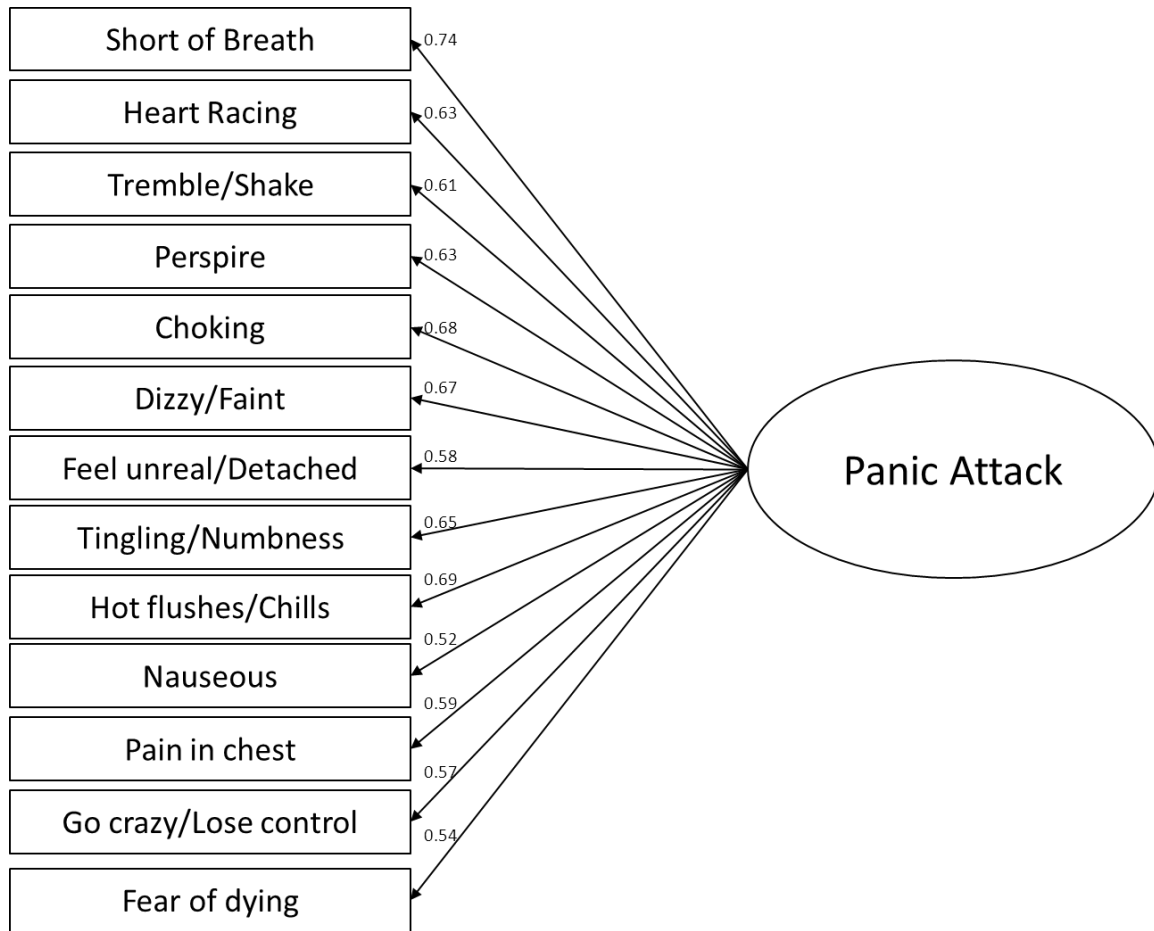


Figure 1: Path diagram and standardised factor loadings for a unidimensional confirmatory factor model of panic symptoms in respondents who have experienced at least one panic attack in their lifetime.

Figure 2: Item Response curves (top) and Item Information curves (bottom) for DSM-IV panic symptoms in respondents who have experienced at least one panic attack in their lifetime.

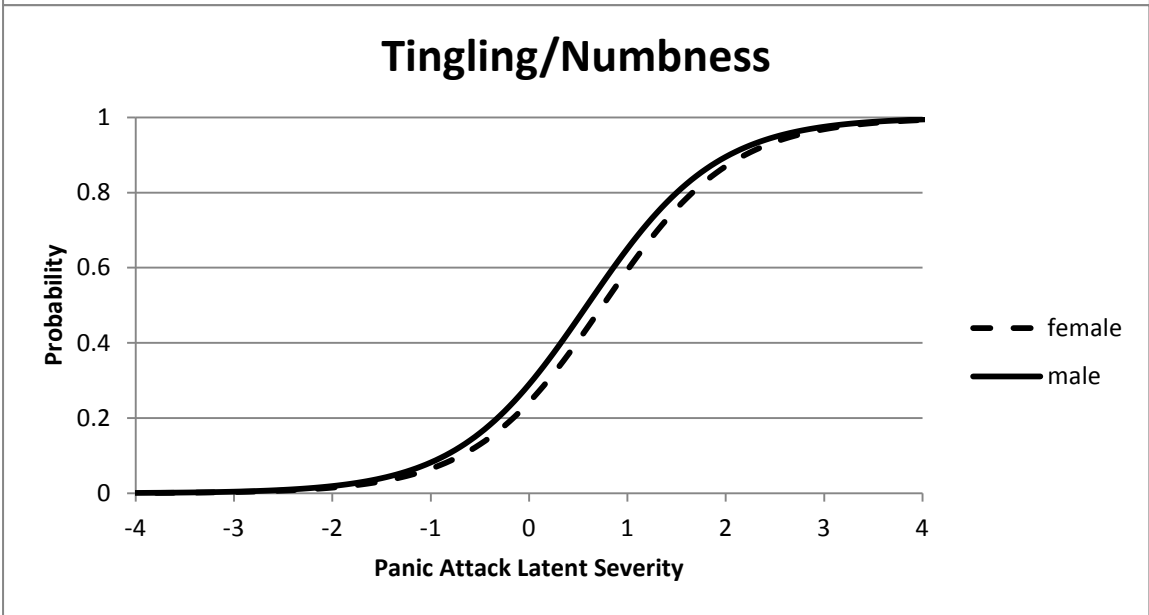
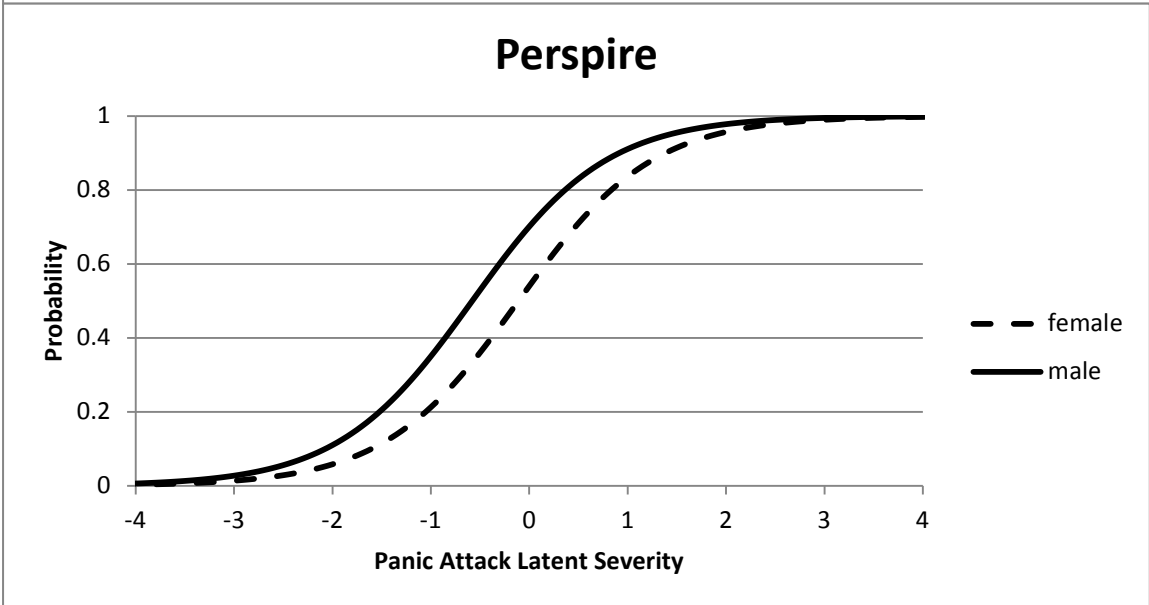
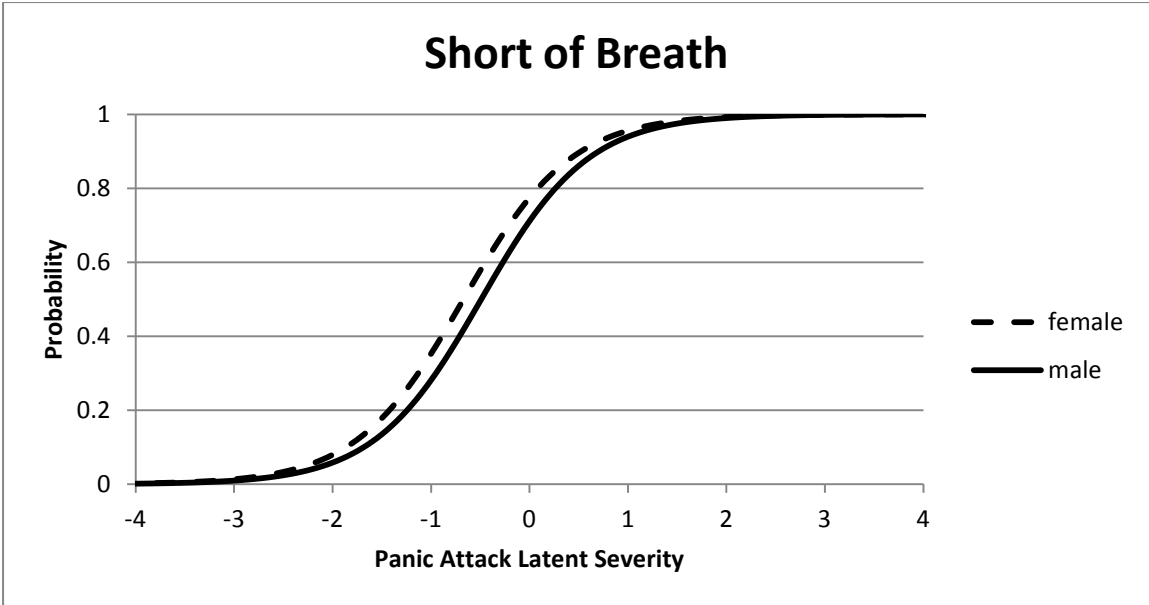


Figure 3: Item Response curves representing differential item functioning between male and female respondents for three DSM-IV panic symptoms (short of breath, perspire, tingling/numbness).

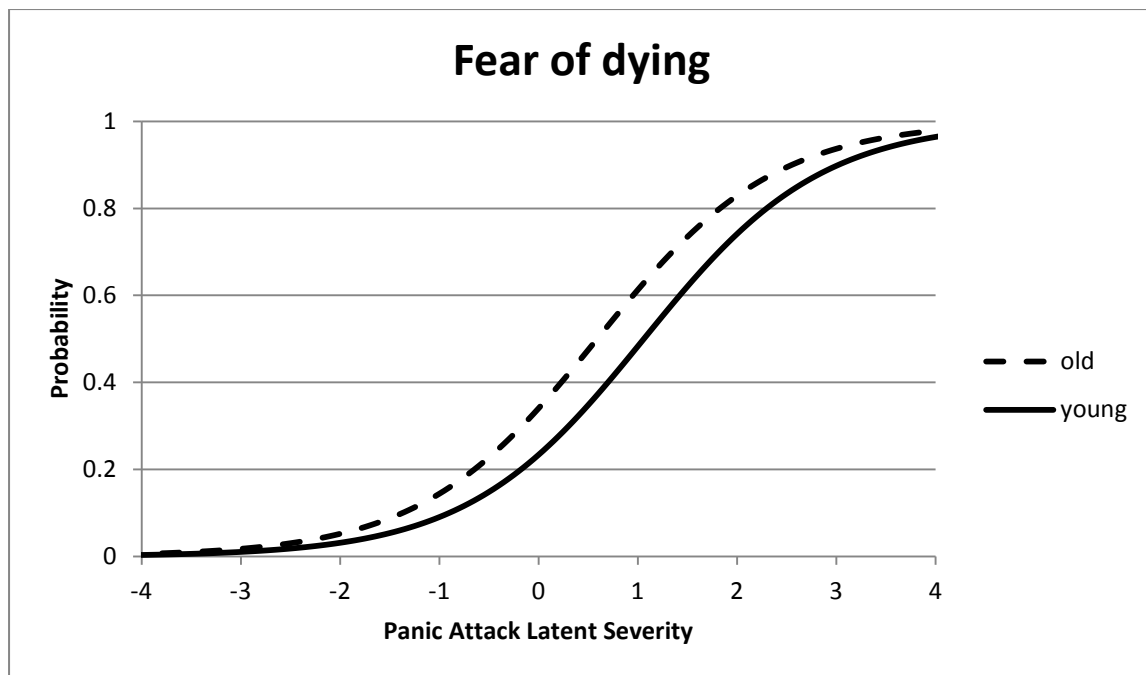


Figure 4: Item Response curves representing differential item functioning between young and old respondents for symptom addressing the fear of dying.