2008

Factor structure of the life orientation test and life orientation test-revised: The influence of item framing

Jamie Moore
Edith Cowan University

Edith Cowan University

Copyright Warning

USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.

COPYRIGHT AND ACCESS DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

(i) Incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution or higher education;

(ii) Contain any material previously published or written by another person except where due reference is made in the text; or

(iii) Contain any defamatory material

Signed _

Dated __

Factor Structure of the Life Orientation Test and Life Orientation Test – Revised: The Influence of Item Framing

Jamie Moore

A report submitted in Partial Fulfilment of the Requirements

for the Award of Bachelor of Arts (Psychology) Honours,

Faculty of Computing, Health and Science,

Edith Cowan University.

Submitted (October, 2008)

I declare that this written assignment is my own work and does not include:

(i) material from published sources used without proper acknowledgement; or
(ii) material copied from the work of

Declaration

I certify that this literature review and research project does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any institution of higher education and that, to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where due reference is made in the text.

Acknowledgements

I would like to take this opportunity to thank all of those who assisted me in the completion of this research. I would like to give special thanks to Dr Ricks Allan (Primary Supervisor) for the excellent suggestions and criticisms she gave on several drafts of this Literature Review and Research Report. Also I would like to extend special thanks Mr Craig Harms for his invaluable assistant with the data analysis involved with the following Research Report.

Table of Contents

Using Balanced Scales to Control for Acquiescence: A Review of the Effects on Factor Structure and Validity of such Scales

Jamie Moore

Using Balanced Scales to Control for Acquiescence: A Review of the Effects on Factor Structure and Validity of such Scales

Abstract

Historically psychological scales have used a mix of positively keyed and negatively keyed items (balanced scales) to control for the effects of response sets. While it has been established that the use of balanced scales does effectively control for response sets such as acquiescence, issues relating to the psychometric properties of these scales emerge. The following review investigated issues surrounding the reliability, validity and factor structure of balanced scales by considering whether these issues were caused by positively and negatively keyed items measuring different aspects of a construct or whether they emerged simply due to measurement error. Both these positions are supported by research with various balanced scales, though it is necessary for future research to consider the effect that negative item framing, rather than negative item keying, has on the psychometric properties of balanced scales.

Author: Jamie Moore
Supervisors: Ricks Allan
and Craig Harms

Submitted: October, 2008

Using Balanced Scales to Control for Acquiescence: A Review of the Effects on Factor Structure and Validity of such Scales

Psychological scales are used to determine an individual's position on a range of psychological, emotional or personality constructs. Balanced scales, that is scales with half the items worded in a positive direction and half in a negative direction, are often used. However, various issues have been identified concerning the reliability and construct validity of balanced scales. The following review outlines the history of why the balanced scale technique was developed to provide a context for understanding the problems it created. The main focus of the review is the emerging issues of balanced scales in regards to their reliability, validity in general and factorial structure specifically. These issues have been researched in a number of balanced scales. The emphasis in this review is on the Rosenberg Self Esteem Scale (Rosenberg, 1965), the Quantitative Attitude Questionnaire (Chang, 1995a), the Leader Behaviour Description Questionnaire (Schriesheim & Hill, 1981), the Computer Anxiety Scale (Greenberger, Chen, Dmitrieva, & Farruggia, 2003), and the Life Orientation Test (LOT; Scheier & Carver, 1985) and Life Orientation Test- Revised (LOT-R, Scheier, Carver, & Bridges, 1994). The majority of the review is on the LOT and LOT-R as a number of studies have investigated what item properties influence its factor structure by making changes to the meaning and framing of items.

Response Sets

When psychological measurement scales were first developed it was thought that responses to items on the scale were exact, unbiased estimates of how respondents actually felt or considered the statement or question (Smith, 1967). However, it began to emerge that psychological tests were not *pure* measures of intended constructs and could not predict human behaviour with high accuracy (Cloud & Vaughan, 1970). It was suggested that response sets of respondents, such as acquiescence, were responsible for this observation.

Response sets refer to a personal tendency to respond in a specified way within a testing or interview situation that is independent of the content of the item or question presented (Smith, 1967). The endorsement of a certain response to an item does therefore not reflect the respondent's position on the construct but instead reflects their specific response set. The response set of most interest in this review is acquiescence which reflects the tendency to agree or disagree with an item irrespective of its content (Knowles & Nathan, 1997).

Acquiescence Response Set

Acquiescence has been referred to as *yea-saying* versus *nay-saying*, reflecting the tendency to agree or disagree respectively (Smith, 1967). An example of yea-saying would be when a respondent endorses the question "I am very happy", and later endorses its opposite "I am very sad". Importantly if acquiescence is uncontrolled within a psychological scale responses to items lose their meaning and the respondent's answers are uninterpretable (Knowles & Nathan, 1997). Knowles and Nathan (1997) investigated whether acquiescent responding was a general characteristic of respondents that was stable over a questionnaire. They had 65 college undergraduates complete the Jackson Personality Inventory, which consists of 320 statements, divided into 15 personality subscales where respondents answer *True* or *False* as a description of themselves. They observed acquiescent responding when a respondent answered *True* or *False* many more times than expected, consistently across the 15 scales, indicating a tendency to agree or disagree more than expected. Their results provided evidence of a general acquiescence trait with a relatively equal amount of yea-sayers and nay-sayers. The generalisation of these results to other scales is limited though as all acquiescence scores were based only on true-false choices, not a range of scale answers and extracted from the same personality scale, administered at the same time. Therefore the study did not allow for variations in time, scale, or format that may affect acquiescence responding. However, it

still provided evidence that respondents do show a tendency to agree or disagree with items irrespective of their content when responding to a measurement scale.

Controlling the Effect of Acquiescence by Developing Balanced Scales

As acquiescence was considered a stable trait that has the ability to considerably influence responses to scale items it was suggested that if researchers wanted to investigate a construct they must take care to avoid or correct for the effects of acquiescence during scale construction (Smith, 1967). It was first suggested that instead of using fixed *true-false*, *agree-disagree* response formats, respondents should be provided with contentful alternatives (Smith, 1967). For example instead of using the item "Most people you meet for the first time cannot be trusted, Strongly Agree/Agree, Strongly Disagree/Disagree", the item would instead be written as: "When meeting someone for the first time, should you": (a) Trust them until they prove unworthy of your trust, (b) Be cautious about trusting them until you know them better, or (c) Not trust them because they may take advantage of you. Using this alternative does not allow a respondent to simply respond on the basis of other questions but forces them to consider each response option carefully (Smith, 1967).

The contentful alternative technique was not favoured though as it made item construction time consuming and complicated, instead it was suggested that acquiescence could be controlled by using a balanced item set, where the trait under measurement is indicated by *yes*, *true*, or *agree* for half the items and *no*, *false*, or *disagree* for the other half (Cloud & Vaughan, 1970; Nunnally & Bernstein, 1994). While this technique does not eliminate acquiescence it does distribute it equally across the scale's items so that the trait scores are relatively free of its effects (Rundquist, 1966). Using this technique, when a measurement scale is constructed half of the items are keyed positively (e.g., "I am happy"), and the other half are keyed negatively (e.g., "I am sad"). In terms of the construct being measured, positively keyed items thus have a positive meaning and negatively keyed items a

negative meaning. When interpreting respondents overall scores, negatively keyed items are reversed scored so that endorsing *strongly agree* or *yes* on a positively keyed item is equal to endorsing *strongly disagree* or *no* on a negatively keyed item. This technique is thought to not only balance out the effect of acquiescence but also force respondents to consider the content of each item carefully and respond accordingly, instead of just responding according to their general feeling about what they perceive is the intended construct (Barnette, 2000).

Cloud and Vaughan (1970) investigated the efficacy of the balanced item technique in controlling for acquiescence. In their study they aimed to measure acquiescence in an attitude scale to see to what extent it was controlled by balanced keying. They had 496 college undergraduates and high school students complete the Wilson and Patterson Conservatism Scale, which consists of 50 items of controversial issues responded to on a *yes*, *no*, *don't know* format, depending on a respondent's belief in the issue. There are an equal number of positively and negatively keyed items on the scale, from which a score of conservatism-liberalism is produced. They constructed a formula that measured response style, dependent on expected responses to items based on keying direction, to determine yea-saying versus nay-saying. They found that the strategy of balancing item-keying was successful in eliminating the distorting influence of acquiescent responding. The correlation of conservatism-liberalism and response style was very low, leading them to recommend balanced keying as a standard element of test construction.

The technique of using both positively and negatively keyed items to control response bias was accepted under the assumptions that response biases were threats to scale validity, that negatively keyed items could be used without serious consequences and most importantly that there were no major psychometric differences between positively and negatively keyed items (Schriesheim & Eisenbach, 1995). As a result of these findings and assumptions many psychological measurement scales have adopted the balanced item

technique including the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1940), the RSES (Rosenberg, 1965), the UCLA Loneliness Scale (Russell, 1978), the Meyer and Allen Affective and Continuance Scale (Meyer & Allen, 1984), and the LOT (Scheier & Carver, 1985), to name a few.

Problems Associated with Balanced Scales

The last assumption regarding balanced scales suggests that negatively keyed items measure the same intended construct as their positively keyed counterparts (Woods, 2006). However, this assumption has consistently not been met, leading some to highlight that the recommendation of using both positively and negatively keyed items has received mixed empirical support (Woods, 2006).

Reliability and Validity in General

Schriesheim and Hill (1981) hypothesised that negatively keyed items may in fact elicit response bias or measure unintended aspects of the construct under investigation. By investigating the effects of item keying on the accuracy, and therefore the validity, of results obtained on standard questionnaires, they suggested that the inclusion of negatively keyed items could result in less accurate responses. They had 150 undergraduates read a fictitious account of a supervisor's behaviour, and then rate the behaviour on the Leader Behaviour Description Questionnaire (LBDQ). Participants read an account of a supervisor who *always* or *never* elicited desirable managerial behaviours, then rated this behaviour on one of three forms of the LBDQ. The Initiating Structure and Consideration subscales of the LBDQ were used to create three 20-item questionnaires that rated leadership behaviour using either all positively keyed, all negatively keyed, or mixed items. Participant's responses were analysed to determine how accurate they were in describing the supervisor's actual behaviour. Results indicated that the positively keyed questionnaire yielded significantly greater accuracy than the mixed or negatively keyed questionnaire. They reasoned that negatively keyed items

caused inaccuracy in responding, which actually slightly increased when they were mixed with positive items. These findings therefore challenged the assumption that item reversals are not without consequences (Schriesheim & Eisenbach, 1995).

Holden, Fekken, and Jackson (1985) criticised Schriesheim and Hill by highlighting that they did not distinguish between negative item framing and negative item keying, therefore it was unknown what aspect of the items caused inaccuracy. They defined items that were reverse-scored as negatively keyed and distinguished between three types of negative framing including clear negatives (i.e., use of word not or never), negative prefixes (i.e., such as im- or un-), and negative qualifiers (i.e., seldom or rarely). Schriesheim, Eisenbach, and Hill (1991) took this methodology on board and examined the effects of item keying and item framing on measurement scale validity. In their study they compared four different types of items: regular items that had a positive meaning and positive framing (e.g., "I am happy"), polar opposites items that had a negative meaning but positive framing (e.g., "I am sad"), negated polar opposites items that had a positive meaning but negative framing (e.g., "I am not sad"), and negated regular items that had a negative meaning and negative framing (e.g., "I am not happy"). These four types of items were an improvement on the comparisons made by Schreisheim and Hill (1981) as they successfully distinguished between item framing and item keying. Using a similar procedure to Schriesheim and Hill (1981), 250 undergraduates rated one of two supervisors on one of four versions of the Initiating Structure and Consideration subscales of the LBDQ. Each version had four regular items, then another four items that were either regular, polar opposite, negated regular, or negated polar opposite items. Results indicated that the two types of positively framed items (regular and polar opposite), had the highest internal consistency reliability. Furthermore both forms of reverse scored items (polar opposite and negated regular), had lower internal consistency reliability than regular items. They also found that items that were negatively

framed (negated opposite and negated regular), had lower internal consistency than positively framed items, irrespective of whether they were positively or negatively keyed. It was suggested that negatively framed items may be inappropriately understood by respondents. They went on to suggest that including both negatively keyed and negatively framed items can significantly decrease the reliability of a measurement scale. These studies by Schriesheim and Hill (1981), and Schriesheim, Eisenbach, and Hill (1991) cast doubt on the assumption that positively and negatively worded item stems measure the same aspect of a construct and further indicate that negatively keyed and negatively framed items are often unreliable.

Factor Structure

Beyond the effects on accuracy and scale reliability it is suggested that the use of negative items can also have effects on the factor structure of a measurement scale. Schmitt and Schults (1985) suggested that wording changes in an effort to create a balanced scale may cause significant changes in the intended factor structure of a scale due to questionable item validities. This is often the case when factor analysis reports a two-dimensional scale structure, when a one-dimensional structure is favoured. They looked at how *careless respondents* could affect the factor structure of a balanced scale. They defined careless respondents as those who have either a positive or negative view of the intended construct as they understand it and proceed to respond to all items in a similar manner that reflects this view, even though items may have been negatively keyed. In this case reverse-scoring these items becomes inappropriate and the respondent's scores become a systematic source of variance not a random one (Schmitt & Stults, 1985). Woods (2006) followed this line of argument by creating an artificial balanced item scale with an intended one factor structure. Woods suggested that when a scale undergoes factor analysis, negative items would form a separate *method factor* that is independent of the construct under investigation. His artificial

scale was made up of 10 negatively keyed and 13 positively keyed items that were created on the basis of a one-dimensional logistic, with possible responses being *1* and *0*. A simulation study was carried out where 0, 5, 10, 20 or 30% of respondents were simulated as careless responders on the artificial scale across sample sizes of 250, 500 and 1000. He then used confirmatory factor analysis (CFA) to test the fit of the intended one-factor structure and a possible two-factor structure across conditions. When 0% of respondents were simulated as careless the intended one-factor model was a perfect fit to the data across all sample sizes. However with even 10% of careless respondents the fit of the one-factor model became unacceptable and the two-factor model comprised of positively keyed items on one factor and negatively keyed items on the other factor provided a better fit to the data. With 20% and 30% of "careless" respondents this two-factor fit was excellent across all three sample sizes, while the intended one-factor fit was poor. They concluded that when negatively keyed items are used, even 10% of careless respondents can artificially affect CFA results and make the obtained factor structure of the scale questionable. In this study though the response options were limited, therefore it is easy to imagine alternative types of responding showing less artificial effects. However the study does support the idea that a small amount of careless responding can form a separate method factor comprised entirely of negatively keyed items. Whether this obtained factor structure is actually of concern to how the scale measures the intended construct, or simply *method variance*, must be considered (Schmitt & Stults, 1985). If the obtained factor structure is a result of method variance, this is a problem because it implies that the way items measure the intended construct elicits some form of systematic response bias.

Other ways individuals respond to items that vary in direction can also result in artifactual factor structures comprised of item keying direction. Campostrini and McQueen (1993) conducted a study using 90 items from a lifestyle and health survey. They analysed

responses from 15,221 interviews in which items were presented positively keyed, negatively keyed and then positively keyed again over an 8-month period. For example the item "it is highly unlikely that AIDS will spread in the general population", was also presented as "it is highly likely that AIDS will spread into the general population". They found that respondent's responses to the two forms of the item were not equal; in that simply reverse scoring the negative item did not correspond to the same response on the positively keyed item. Respondents tended to endorse a negative item rather than reject a positive item. They also suggested that those who were less educated possibly did not perceive the subtle differences in the semantics of the positive versus the negative items when responding. Spector, Van Katwyk, Brannick, and Chen (1997) then suggested if individuals respond differently to oppositely keyed items, then item correlations with the overall scale score become unequal, leading to one subset having a higher or lower correlation than the other. If this occurs a two factor structure will emerge when the scale is factor analysed, even if the items assess a single construct (Spector et al., 1997). Ibrahim (2001) refers to these emerging negative factors as *method artifacts* that affect the obtained dimensionality of scales in a systematic instead of a random way. In his study only one item out of 23 was negatively keyed and it still loaded separately on its own factor when exploratory factor analysis was performed. Ford, MacCullum, and Trait (1986) have suggested though that exploratory factor analysis is not as powerful as CFA as it takes advantage of chance variance in a sample, resulting in factors being extracted when none actually exist, therefore it is possible this occurred in Ibrahim's study.

From the studies that investigated the effect of item keying on the factor structure of scales, it is clear that item keying can have dramatic consequences for the factor structure, thus violating the assumption that negatively keyed items can be used without serious consequences. It seems that by using negatively keyed items to create a balanced scale to

guard against response sets, specifically acquiescence, these items actually create further item wording effects that result in unintended factor structures being obtained (Barnette, 2000). Whether this occurs due to negative items being more difficult to interpret (Cordery & Sevastos, 1993), careless responding (Schmitt & Stults, 1985), or socially desirable responding (Reiser, Wallace, & Schuessler, 1986), it is a significant problem regarding the psychometric properties of scales.

Connotatively Consistent and Connotatively Inconsistent Items

Chang (1995a) redefined this concept of negatively keyed items in balanced scales by referring to all items as either connotatively consistent (CC) or connotatively inconsistent (CI). He suggested the connotation of items depends on whether an item agrees or disagrees with the majority of items that make up a scale. For example items are CC on a scale where the items are all positively or all negatively keyed, where as the items on a balanced scale are CI. Chang (1995a) argues that the assumption that reverse scoring negatively keyed items makes the items on the scale CC in regards to the intended construct is empirically unverified. He examined whether positively keyed and negatively keyed items measured the same intended construct conducting a study based on generalisability theory. A sample of 102 masters students were administered eight items taken from the Quantitative Attitude Questionnaire (QAQ), on two separate occasions, one week apart. The first week, four of the eight items were inconsistent with the connotation of the QAQ and the second week all eight items were rewritten to be the opposite of their connotation during the first administration. In both versions, the eight items were mixed with other items from the QAQ and presented using a 6-point Likert scale. Thus, two observations for each item were recorded, one with negatively keyed wording (e.g., "I'm bad with math") and one with positively keyed wording (e.g., "I'm good with math"). Chang found that reverse scored negatively keyed items were not fully equivalent to their positively keyed counterparts. For example participants may have

given a rating of *6* for the positively keyed item and a rating of *4* for its negatively keyed counterpart. He concluded that reverse scoring negatively keyed items prior to analysis was a questionable procedure and suggested that the use of CI items should be avoided when possible, in favour of scales where all items are consistently keyed.

These findings are supported by other researchers who question the validity of CI items. Barnette (2000) suggested that unless there was an important reason for not doing so, it is best that all items be positively or negatively keyed (i.e., CC) and not mixed (i.e., CI). Woods (2006) added that employing factor analysis on scales with CI items, often results in a two-factor structure, even when the construct is expected to be one dimensional. It has been argued that this separate factor emerges due to the items being keyed in opposite directions (i.e., CI) rather than reflecting differences in item content (Woods, 2006). Similar to the QAQ, two factor structures due to item-keying have been reported on other scales that employ CI items and have an intended one-factor structure such as the UCLA Loneliness Scale (Austin, 1983), the Social Interaction Anxiety Scale (Rodebaugh, Woods, & Heimberg, 2007), Meyer and Allen's Affective and Continuance Commitment Scales (Magazine, Williams, & Williams, 1996) and subscales of the MMPI (Messick & Jackson, 1961) to name a few.

Two studies in particular have looked at current applied scales and tested whether a one-dimensional or multidimensional structure is most suited, dependent on the use of CC or CI items. Greenberger, Chen, Dmitrieva, and Farruggia (2003) investigated the factor structure of the Rosenberg Self-Esteem Scale (RSES), a single-factor scale with scores ranging along a continuum from low self-esteem to high self-esteem obtained from ten CI items. Several studies have found the 10 items on the scale actually split into two factors, defined by item-keying-direction (Goldsmith, 1986; Owens, 1993). These authors suggested that these two factors actually measured "positive self-esteem" and "negative self-esteem"

separately, instead of along a continuum. Greenberger, Chen, Dmitrieva, and Farruggia (2003) investigated this by administering three versions (original, revised negative, revised positive) of the RSES to three groups of undergraduates. The original RSES ($n = 257$), comprised of five positively keyed and five negatively keyed items. In the revised negative version ($n = 244$), the five positively keyed items were rewritten to reflect negative keying and in the revised positive version ($n = 240$), the five negatively keyed items were rewritten to reflect positive keying. The items in the original version were CI where as the items in both revised versions were CC. All versions were responded to on a 6-point Likert scale, ranging from 1 (*strongly disagree)* to 6 (*strongly agree).* Confirmatory factor analysis was conducted on all three versions comparing the fit of a one-factor model and a two-factor model to the data. The results indicated that a two-factor model fit the data from the original RSES significantly better than the one-factor model. However for the two CC revised RSES versions the one-factor model fit improved significantly and was not worse than the two-factor model fit. While the one-factor model fit was not ideal for the CC versions it was a significant improvement from the CI version. They suggested that the two-factor structure of the original RSES was a result of item-keying rather than positively and negatively keyed items measuring separate aspects of self-esteem.

Pilotte and Gable (1990) came to a different conclusion when they studied the impact of positively-keyed and negatively-keyed items on the Computer Anxiety Scale (CAS), a scale with an intended one-dimensional structure. Using a similar procedure to Greenberger, et al. (2003) they administered three versions of the CAS to (n = 271), high school students in grades nine to twelve. One sample ($n = 94$) completed the original CAS that is a one-dimensional scale with nine positively keyed items measuring computer anxiety (CC). They also created a negative version ($n = 90$), by negating each original item then reverse scoring it (CC) and also a mixed version ($n = 87$), containing five items from the original version and

four items from the negated version (CI). Using CFA they compared the fit of a one-factor model and two-factor model on the mixed version. They found that the two-factor model corresponding to positively keyed and negatively keyed items on separate factors was a significantly better fit than the one-factor model. They also used multiple groups analysis to test how similar responses were on the two CC versions. Results indicated that the positive and all negative versions of the CAS elicited significantly different responses. While they did not test the factor structure of these two CC versions they did find that negating an item on the CAS affected a student's response to that item. Therefore while Greenberger, et al. (2003) suggest positive and negative items do not measure separate constructs and support the notion that the two-factor structure simply reflects item-keying-direction, Pilotte and Gable's study suggests positive and negative items do measure separate constructs, supporting the notion that the two-factor structure reflects differences in item content. The results indicated that responses to items on a CC positive scale were not equivalent to responses to items on a CC negative scale, indicating that even when a scale is CC, item framing and item keying do affect responses. However as the sample sizes of the two CC versions were not particularly high it is possible a Type I error was made and the significant result found was due to respondent characteristics not item characteristics. To remove these effects each respondent would have to complete both CC versions, which would lead to issues of fatigue and practice, therefore it is difficult to determine whether respondent not item characteristics influenced results. Also most studies that investigated the effect of negative-item-keying have used college students or adults, whereas Pilotte and Gable's study employed high school students. It is therefore possible that these younger less educated respondents may have had more difficultly responding to the more semantically challenging negatively-keyed-items (Cordery & Sevastos, 1993). It is feasible that a sample of older, more educated adults or students may elicit different results.

Life Orientation Test and Life Orientation Test-Revised

Studies that dispute whether the obtained two-factor structure of balanced scales is due to item-keying-direction or meaningful differences in item content leads to a review of the LOT. The LOT is an 8-item scale comprised of four items positively keyed reflecting optimism and four items negatively keyed reflecting pessimism, which individuals respond to on a 5-point Likert scale (Scheier & Carver, 1985). The LOT was based upon a one-dimensional representation of optimism/pessimism existing along a single continuum. It was considered that an individual could not be optimistic and pessimistic, but that their level of optimism/pessimism instead existed along this continuum.

Factor Structure of the Life Orientation Test

Scheier and Carver (1985) created the LOT scale and administered it to 624 undergraduates to assess its psychometric properties. Using principal-axis factor analysis they extracted two factors, the first factor defined by items keyed negatively and the second factor by items keyed positively. They subsequently tested the data using CFA to compare a one-factor model with a two-factor model defined by item-keying-direction. While both the one-factor and two-factor models yielded an acceptable fit to the data, it was found that the two-factor model was significantly better. While they did suggest there was justification for examining the two halves of the scale separately, they still suggested it should be treated as one-dimensional for most purposes as the two factors extracted had a high positive correlation, (r = .64) and the factors reflected item-keying-direction not differences in meaningful content. Several subsequent studies have also supported the two-factor model of the LOT over the one-factor model in a range of samples (Bryant & Cvengros, 2004; Lai, Cheung, Lee, & Yu, 1998; Lai & Yue, 2000; Marshall, Wortman, Kusulas, Hervig, & Vickers, 1992; Robinson-Whelan, Kim, MacCallum, & Kiecolt-Glaser, 1997; Steed, 2002). Other studies also found high correlations between the two factors, (r = .69) (Steed, 2002; Marshall & Lang, 1990), as

well as moderate negative correlations between positive and negative items, indicating a tendency towards a one-dimensional view (Hjelle, Belongia, & Nesser, 1996). The two factors also comprised of negatively-keyed and positively-keyed items respectively, which reflects *method bias* rather than differences in meaningful content (Marshall & Lang, 1990).

Factor Structure of the Life Orientation Test - Revised

Similar results have also been reported for the LOT-R, a revised version of the LOT that contains 6-items, half positively keyed and half negatively keyed, as well as four fillers. The original authors Scheier, Carver, and Bridges (1994) extracted two factors using exploratory factor analysis, with a two-factor model also being favoured over a one-factor model with subsequent CFA. Both analyses reported the two-factor model comprising of negatively and positively keyed items respectively. They did suggest that the scale be treated as one-dimensional however, as the one factor model was favoured when correlated errors among positive items were allowed and the factors were defined by item-keying. However, Vautier, Raufaste, and Cariou (2003) argued that if correlated errors are allowed in CFA these errors tell us that one-dimensionality has been lost. Other studies that suggest a one-dimensional LOT-R factor structure are plagued by issues regarding small sample sizes, inappropriate correlational techniques and reliance on unconservative goodness-of-fit indexes (Mehrabian & Ljunggren, 1997; Rauch, Schweizer, & Moosbrugger, 2007). Similar to the LOT, studies with acceptable sample sizes found data from the LOT-R yielded a poor to moderate one-factor model fit and an acceptable to high two-factor model fit (Vautier & Raufaste, 2006; Creed, Patton, & Bortrum, 2002; Mehrabian & Ljunggren, 1997; Herzberg, Glaesmer, & Hoyer, 2006). Therefore while the original authors still support one-dimensionality, they and others suggest that the responses to the LOT-R can be scored along two subscales reflecting trait optimism and trait pessimism (Scheier, Carver, & Bridges, 1994; Burke, Joyner, Czeck, & Wilson, 2000). Once again some argue though that these two factors emerge due to

measurement error as a result of item-keying-direction and not meaningful differences in item content (Nakano, 2004).

Different Explanations for Two-Factor Structure

While Scheier and Carver (1985) found the one-factor model yielded an acceptable fit to their data, other studies suggest the one-factor model fit is moderate to poor (Marshall et al., 1992; Bryant & Cvengros, 2004; Marshall & Lang, 1990; Robinson-Whelan, Kim, MacCallum, & Kiecolt-Glaser, 1997). In addition it has been suggested that the two-factor model emerges due to differences in the content of positive and negative items, rather than reflecting measurement error due to their opposite keying directions (Lai, 1994). Mook, Kleijn, and Van Der Ploeg (1992) administered the Dutch version of the LOT to 166 undergraduates in the Netherlands. Using exploratory factor analysis, two factors were extracted comprised of negatively and positively keyed items respectively. They reported a factor correlation of $r = .31$, and labelled the two factors optimism and pessimism. They argued that these results supported a two-dimensional view of the LOT, not a one-dimensional view, reflecting differences in item content along the independent dimensions of positive and negative affect. They suggested that rather than measuring optimism, negatively keyed items measured a lack of pessimism. They argued that agreeing that one rarely feels pessimistic is not equivalent to saying one feels optimistic nearly all the time. However, their sample size was much smaller than in the original assessment of the LOT and they did not conduct the more stringent CFA on their data. Therefore Mook, et al's. (1992) results may be misleading.

The argument that the two factors emerge due to measurement error and not differences in item content comes from studies that investigated the LOT with respect to CC and CI items (Chang, 1995a). Chang (1995b) argued that in relation to the LOT a rating of "0" for a negatively keyed item, would not truly be equal to a "5" if the item was reworded to

be positively keyed. He suggested that the two-factor model of the LOT may be caused by the use of CI items that do not measure the same aspects of the construct as CC items due to their inconsistent connotation, therefore endorsing an item reflecting optimism would not be equivalent to disagreeing with an item reflecting pessimism. Chang and McBride-Chang (1996) continued this line of research by investigating the factor structure of the LOT in its original form and when the items were reworded to reflect consistent pessimism or consistent optimism. The consistent optimism version was created by rewriting the four pessimism items to reflect optimism. For instance, the item "Things never work out the way I want them to" was changed into "Things always work out the way I want them to". The consistent pessimism version was created by rewriting the four optimism items to reflect pessimism. For example, "In uncertain times, I usually expect the best" became "In uncertain times, I usually expect the worst". The two rewritten versions therefore contained CC items, whereas the original contained CI items. They had undergraduates complete one version of consistent pessimism (n = 149), or consistent optimism (n = 129) on a 4-point scale ranging from 1 = *strongly disagree* to 4 = *strongly agree*. Another sample completed the original LOT (n = 108) using both a 4-point and 6-point scale. Confirmatory factor analysis indicated an overwhelming superiority of the two-factor model on the original; however the two CC versions supported the one-factor model. They suggested that the two-factor structure of the LOT was therefore a result of response bias due to the use of oppositely-keyed items, rather than reflecting substantial differences in item content. However, they did suggest it was possible optimism and pessimism were not bipolar constructs along a single continuum but instead represented separate but related traits. It could not be determined whether this factor structure was driven by meaning or framing though, as both were changed in the rewritten versions. Therefore as meaning and framing were not split it is unknown whether respondents responded to the content of the item, the framing, or both. Also it is unknown why a 4-point

and 6-point scale were used instead of the original 5-point scale and why the response options were reversed, to indicate strongly disagree first. It is possible these subtle changes of the LOT may have confounded the obtained results.

Kubzansky, Kubzansky, and Maselko (2004) distinguished between framing and meaning in their study that aimed to investigate whether the two factor structure of the LOT was determined by response bias due to item-keying-direction or meaningful content. They stated that a positively framed item uses words with positive connotations (e.g., always or right), whereas a negative framed item uses words with negative connotations (e.g., never or wrong). They suggested that an item could be negatively framed while still maintaining its positive meaning and vice versa. For example the item "I'm never pessimistic about my future" is negatively framed but has a positive meaning. In their study they teased apart the meaning and framing of items, by administering three versions of the LOT to undergraduates. The original version was administered (n = 146), as well as the two reworded versions. The first (n = 141) was derived by changing the framing of half of the items on each subscale but preserving the meaning. This resulted in each subscale having two negatively-framed and two-positively-framed items while still maintaining consistent optimistic or pessimistic meaning. The other (n = 142) was derived by changing the framing of all items but preserving the meaning so that framing and meaning were inconsistent. They tested three models with CFA; the bipolar model, reflecting a one-dimensional view of the LOT, a method artifact model, that suggested items cluster together because they are similarly framed and a bivariate model that suggests items cluster together because they have similar meaning. The bipolar model had an unacceptable fit for all three versions. However for the two derived versions the bivariate model was a better fit than the method artifact model. Items with optimistic content loaded on one factor and items with pessimistic content loaded on another factor regardless of which direction they were framed. They suggested this

strongly implies the factor structure of the LOT is driven by item meaning and not measurement error due to item keying-direction and that optimism as measured by the LOT does not exist along a single continuum with pessimism. One problem with this study is that all versions administered were CI, not CC, as has been recommended by Chang (1995a) and Chang (1995b). It is assumed that the LOT-R would perform similarly under these conditions, due to similar results found for the LOT and LOT-R in other studies that assessed factor structure.

Allan and Giles (2008) addressed the issue further by investigating the factor structure of the LOT in a sample of West Australian prisoners to determine whether it was affected by measurement error. They read items of the LOT to (n = 453) participants, after an hour long interview intended to build rapport, which the participants responded to on a 5-point Likert scale. The results of CFA indicated a two-factor model corresponding to item-keying-direction was a better fit to the data than a one-factor model, once again demonstrating that the structure of the LOT is not in line with the one-dimensional construct it was intended to measure. Following on though, when they removed participants who demonstrated a tendency to consistently agree or disagree, the data was reanalysed and the fit of the one-factor model improved while the two-factor model fit decreased. They suggested this indicated that the two factor structure of the LOT is a result of measurement error due to item-keying-direction rather than reflecting substantial differences in the content of positively and negatively keyed items. Several methodological issues regarding this study exist though, due to the use of prisoners and the implementation of the LOT. The LOT is a scale test that is meant to be completed individually, rather than read out to a participant, therefore in the following study it is possible that socially desirable responding was observed as the participant attempted to present themselves in a positive light to the interviewer. Also, while

care was taken to control for this, it is possible prisoners may have extra incentive to present themselves in a positive light in order to improve their circumstances in prison.

Overview of Balanced Scales and Future Research Possibilities

The current review has outlined issues surrounding the use of balanced scales that contain a mix of positively and negatively keyed items, in an effort to control for the effect of acquiescence. When negatively keyed items are written often both the meaning and framing of the items are negative, which presents difficulty when individuals respond to such items. By examining scales that use this technique such as the RSES, CAS, QAQ and LOT it is apparent that negatively-keyed items affect the validity and reliability of the scale, specifically in regards to the obtained versus the theoretical factor structure of the scale. The important issue is whether the artifactual factor structures obtained on balanced scales are a result of *method bias* due to item-keying-direction or whether they are due to positively-keyed and negatively-keyed items actually measuring separate, unrelated aspects of the construct under investigation. Several studies have looked at this issue by analysing scales with items that have consistent positive or negative keying, some of which suggest the artifactual factor structures emerge due to differences in item content (Chang, 1995a; Pilotte & Gable, 1990; Kubzansky, Kubzansky, & Maselko, 2004), whereas others suggest it is simply a result of item-keying-direction (Greenberger, Chen, Dmitrieva, & Farruggia, 2003; Chang & McBride-Chang, 1996). Nonetheless, it is necessary for future research to focus on scales where the items have consistent positive or negative framing to establish whether this affects the obtained factor structure in a similar manner.

There is need for a study that will further attempt to tease apart the concept of framing and meaning in understanding the factor structure of the LOT and LOT-R by testing what happens to the factor structure of the scales when items have consistent framing but balanced keying. This could be done by using two modified versions of both the LOT and LOT-R,

each with either consistent positive or negative item framing. These versions would contain four items that reflect optimism and four items that reflect pessimism but all the items would be consistently framed. The positively framed version could be created by taking the four original optimism items and combining them with the four reverse framed pessimism items created by Kubzansky, et al. (2004). The negatively framed version could be created by combining the original four pessimism items with the four reverse framed optimism items created by Kubzansky, et al. (2004). A similar procedure could be undertaken for the LOT-R with three original items being used on each scale, along with three reverse framed versions that should be created in line with suggestions by Kubzansky, et al. (2004). Using CFA, if a two-factor model is found to be a better fit to the data for both versions despite all items having consistent framing it would suggest that the two-factor structure of the LOT does emerge due to substantial differences in the content of optimism and pessimism items. If a one-factor model provides a better fit to the data it would suggest that the two factor structure of the LOT is obtained due to response bias due to items being oppositely framed, rather than reflecting substantial differences in item content.

References

Allan, M. M., & Giles, M. (2008). Psychometric properties of Scheier and Carver's Life Orientation Test in a samples of West Australian prisoners. *Psychological Reports*, 103, 305-322.

Austin, B. A. (1983). Factorial structure of the UCLA Loneliness Scale. *Psychological Reports*, 53, 883-889.

Barnette, J. J. (2000). Effects of stem and likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, 60, 361-370.

Bryant, F. B., & Cvengros, J. A. (2004). Distinguishing hope and optimism: Two sides of a coin, or separate coins? *Journal of Social and Clinical Psychology*, 23, 273-302.

Burke, K. L., Joyner, A. B., Czeck, D. R., & Wilson, M. J. (2000). An investigation of concurrent validity between two optimism/pessimism questionnaires: The Life Orientation Test-Revised and the Optimism/Pessimism Scale. *Current Psychology*, 19, 129-136.

Campostrini, S., & McQueen, D. V. (1993). The wording of questions in a CATI-based lifestyle survey: Effects of reversing the polarity of AIDS-related questions in continuous data. *Quality & Quantity*, 27, 157-170.

Chang, L. (1995a). Connotatively consistent and reverse connotatively inconsistent items are not fully equivalent: Generalizability study. *Educational and Psychological Measurement*, 55, 991-997.

Chang, L. (1995b). Connotatively inconsistent test items. *Applied Measurement in Education*, 8, 199-209.

Chang, L., & McBride-Chang, C. (1996). The factor structure of the LOT. *Educational and Psychological Measurement*, 56, 325-329.

Cloud, J., & Vaughan, G. M. (1970). Using balanced scales to control acquiescence. *Sociometry*, 33, 193-202.

Cordery, J. L., Sevastos, P. P. (1993). Responses to the original and revised Job Diagnostic Survey: Is education a factor in responses to negatively worded items? *Journal of Applied Psychology*, 78, 141-143.

Creed, P. A., Patton, W., Bartrum, D. (2002). Multidimensional properties of the LOT-R: Effects of optimism and pessimism on career and well-being related variables in adolescents. *Journal of Career Assessment*, 10, 42-61.

Ford, J. K., MacCallum, R. C., & Trait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39, 291-314.

Goldsmith, R. E. (1986). Dimensionality of the Rosenberg Self-Esteem Scale. *Journal of Social Behaviour and Personality*, 1, 253-264.

Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, 35, 1241-1254.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology*, 10, 249-254.

Herzberg, P. Y., Glaesmer, H., & Hoyer, J. (2006). Separating optimism and pessimism: A robust psychometric analysis of the Revised Life Orientation Test (LOT-R). *Psychological Assessment*, 18, 433-438.

Hjelle, L., Belongia, C., & Nesser, J. (1996). Psychometric properties of the Life Orientation Test and Attributional Style Questionnaire. *Psychological Reports*, 78, 507-515.

Holden, R. R., Fekken, G. C., & Jackson, D. N. (1985). Structured personality test item characteristics and validity. *Journal of Research in Personality*, 19, 386-394.

Ibrahim, A. M. (2001). Differential responding to positive and negative items: The case of a negative item in a questionnaire for course and faculty evaluation. *Psychological Reports*, 88, 497-500.

Knowles, E. S., & Nathan, K. T. (1997). Acquiescent responding in self-reports: Cognitive style or social concern. *Journal of Research in Personality*, 31, 293-301.

Kubzansky, L. D., Kubzansky, P. E., & Maselko, J. (2004). Optimism and pessimism in the context of health: Bipolar opposites of separate constructs. *Personality and Social Psychology Bulletin*, 30, 943-956.

Lai, J. C. L. (1994). Differential predictive power of the positively versus the negatively worded items of the Life Orientation Test. *Psychological Reports*, 75, 1507-1515.

Lai, J. C. L., Cheung, H., Lee, W., & Yu, H. (1998). The utility of the revised Life Orientation Test to measure optimism among Hong Kong Chinese. *International Journal of Psychology*, 33, 45-56.

Lai, J. C. L., & Yue, X. (2000). Measuring optimism in Hong Kong and mainland Chinese with the revised Life Orientation Test. *Personality and Individual Differences*, 28, 781-796.

Magazine, S. L., Williams, L, J., & Williams, M. L. (1996). A confirmatory factor analysis examination of reverse coding effects in Meyer and Allen's Affective and Continuance Commitment Scales. *Educational and Psychological Measurement*, 56, 241-250.

Marshall, G. N., & Lang, E. L. (1990). Optimism, self-mastery, and symptoms of depression in women professionals. *Journal of Personality and Social Psychology*, 59, 132-139.

Marshall, G. N., Wortman, C. B., Kusulas, J. W., Hervig, L. K., & Vickers, R. R. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of Personality and Social Psychology*, 62, 1067-1074.

Mehrabian, A., & Ljunggren, E. (1997). Dimensionality and content of optimism-pessimism analysed in terms of the Pad Temperament Model. *Personality and Individual Differences*, 23, 729-737.

Messick, S., & Jackson, D. N. (1961). Acquiescence and the factorial interpretation of the Minnesota Multiphasic Personality Inventory. *Psychological Bulletin*, 58, 299-304.

Meyer, J. P., & Allen, N. J. (1984). Testing the "side-bet-theory" of organizational commitment: Some methodological considerations. *Journal of Applied Psychology*, 69, 372-378.

Mook, J., Kleijn, W. C., & Van Der Ploeg, H. M. (1992). Positively and negatively worded items in a self-report measure of dispositional optimism. *Psychological Reports*, 71, 275-278.

Nakano, K. (2004). Psychometric properties of the Life Orientation Test-Revised in samples of Japanese students. *Psychological Reports*, 94, 849-855.

Nunnally, J. C., & Berstein, I. H. (1994). *Psychometric Theory (3rd Ed.)*. New York: McGraw-Hill.

Owens, T. J. (1993). Accentuate the positive-and the negative: Rethinking the use of self-esteem, self-deprecation, and self-confidence. *Social Psychology Quarterly*, 56, 288-299.

Pilotte, W. J., & Gable, R. K. (1990). The impact of positive and negative item stems on the validity of a Computer Anxiety Scale. *Educational and Psychological Measurement*, 50, 603-610.

Rauch, W. A., Schweizer, K., & Moosbrugger, H. (2007). Method effects due to social desirability as a parsimonious explanataion of the deviation from unidimensionality in LOT-R scores. *Personality and Individual Differences*, 42, 1597-1607.

Reiser, M., Wallace, M., & Schuessler, K. (1986). Direction-of-wording effects in

dichotomous social life feeling items. *Sociological Methodology*, 16, 1-25.

Robinson-Whelan, S, Kim, C., MacCallum, R. C., & Kiecolt-Glaser, J. K. (1997). Distinguishing optimism from pessimism in older adults: Is it more important to be optimistic or not to be pessimistic? *Journal of Personality and Social Psychology*, 73, 1345-1353.

Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The reversal of social anxiety is not always the opposite: The reverse-scored items of the Social Interaction Anxiety Scale do not belong. *Behaviour Therapy*, 38, 192-206.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Rundquist, E. A. (1966). Item and response characteristics in a attitudinal and personality measurement: A reaction to L.G. Rorer's "The greatest response-style myth". *Psychological Bulletin*, 66, 166-177.

Russell, D. A., Peplan, L. A., & Ferguson, M. L. (1978). Developing a measure of lonliness. *Journal of Personality Assessment*, 42, 290-294.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, 4, 219-247.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem: A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, 67, 1063-1078.

Schmitt, N., & Schults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, 9, 367-373.

Schriesheim, C. A., & Eisenbach, R. J. (1995). An exploratory and confirmatory factor-analytic investigation of item wording effects on the obtained factor structure of survey questionnaire measures. *Journal of Management*, 21, 1177-1193.

Schriesheim, C. A., Eisenbach, R. J., & Hill, K. D. (1991). The effect of negation and polar opposites item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement*, 51, 67-78.

Schriesheim, C. A., & Hill, K. D. (1981). Controlling for acquiescence response bias by item reversals: The effect on questionnaire validity. *Educational and Psychological Measurement*, 41, 1101-1115.

Smith, D. H. (1967). Correcting for social desirability response sets in opinion-attitude survey research. *The Public Opinion Quarterly*, 31, 87-94.

Smith, T. W., Pope, M. K., Rhodewalt, F., & Poulton, J. L. (1989). Optimism, neuroticism, coping, and symptom reports: An alternative interpretation of the Life Orientation Test. *Journal of Personality and Social Psychology*, 56, 640-648.

Spector, P. E., Van Katwyk, P. T., Brannick, M. T., & Chen, P.Y. (1997). When two factors don't reflect two constructs: How item characteristics can produce artifactual factors. *Journal of Management*, 23, 659-678.

Steed, L. G. (2002). A psychometric comparison of four measures of hope and optimism. *Educational and Psychological Measurement*, 62, 466-482.

Vautier, S., & Raufaste, E. (2006). Configural instability of data from the Life Orientation Test-Revised. *Personality and Individual Differences*, 40, 1511-1518.

Vautier, S., Raufaste, E., & Cariou, M. (2003). Dimensionality of the Revised Life Orientation Test and the status of filler items. *International Journal of Psychology*, 38, 390-400.

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioural Assessment*, 28, 189-94.

Guidelines for Contributions by Authors

Current Psychology

A Journal for Diverse Perspectives on Diverse Psychology Issues

Executive Editor: J.A. Schaler

Aim and Scope

Current Psychology is an international forum for rapid dissemination of information at the cutting edge of psychology. It includes significant empirical contributions from the major areas of psychology, including social psychology; small groups and personality; human development; sensation; perception; and cognition; clinical and abnormal psychology; and methodology and field research.

Manuscript Style

Type double-spaced using generous margins on all sides. The entire manuscript, including quotations, references, figure-caption list, and tables, should be double-spaced. Manuscript length, except under unusual circumstances, should generally be about 4000–9000 words. Empirical articles should include standard sections, such as Introduction, Methods, Results, and Discussion. Number all pages consecutively with Arabic numerals, with the title page being page 1. In order to facilitate masked (previously termed "double-blind") review, leave all identifying information off the manuscript, including the title page and the electronic file name. Upon initial submission, the title page should include only the title of the article. An additional title page should be provided as a separate submission item and should include the title of the article, author's name, and author's affiliation. Academic affiliations of all authors should be included. The affiliation should comprise the department, institution (usually university or company), city, and state (or nation) and should be typed as a footnote to the author's name. This title page should also include the complete mailing address,

telephone number, fax number, and e-mail address of the one author designated to review proofs. An abstract is to be provided, preferably no longer than 200 words.

References

List references alphabetically at the end of the paper and refer to them in the text by name and year in parentheses. The style and punctuation of the references should conform to strict APA style. In general, the journal follows the recommendations of the 2001 Publication Manual of the American Psychological Association (5th ed.), and it is suggested that contributors refer to this publication. For articles in a periodical, references in the citation list should include (in this order): last names and initials of all authors (for up to and including six authors), year published (in parentheses), title of article (roman type), name of publication (italics), volume number (italics), and inclusive pages (roman type).

Running head: FACTOR STRUCTURE OF THE LOT

The Influence of Consistent Framing on the Factor Structure of the Life Orientation Test and Life Orientation Test- Revised

Jamie Moore

Abstract

The Life Orientation Test (LOT) and Life Orientation Test-Revised (LOT-R) were investigated to see how item framing influenced both scales factor structure. Two modified versions of both scales were created, one with consistent positive framing and the other with consistent negative framing. In both scales the original meaning (keying) of items was maintained so that each version had a balance of positively and negatively keyed items. Confirmatory factor analysis results indicated that a two-factor model was a significantly better fit to the data from the positively and negatively framed LOT and positively framed LOT-R. It was suggested that participants do not respond to item framing but instead item-keying direction when completing both scales. Furthermore participants responded differently to negatively framed items, perhaps due to the increased semantic complexity.

Author: Jamie Moore
Supervisors: Ricks Allan
and Craig Harms
Submitted: October, 2008

The Influence of Consistent Framing on the Factor Structure of the Life Orientation Test and Life Orientation Test- Revised

Introduction

Self-report verbal measures, in spite of their shortcomings, are commonly used in research and therapeutic assessments. Likert-type scales are most common among verbal self-report measures, mainly because the Likert method is conceptually simple and practically straightforward (Ahlawat, 1984). The major source of criticism of self-report data centers around the susceptibility of self-report measures to various *response sets* that pose a continuing threat to the construct validity of such measures and distort the interpretation and conclusions based on such data (Ahlawat, 1984). Response sets refer to a personal tendency to respond in a specified way within a testing or interview situation that is independent of the content of the item or question presented (Smith, 1967). The endorsement of a certain response to an item may therefore not reflect the respondent's position on the construct but instead reflects their specific response set.

One response set of particular interest was *acquiescence* or the tendency to agree (*yea-saying*) or disagree (*nay-saying*) with items on a scale, regardless of their content (Smith, 1967). An example of yea-saying would be when a respondent endorses the question "I am very happy", and later endorses its opposite "I am very sad". Importantly, if acquiescence is uncontrolled within a psychological scale responses to items may lose their meaning and the respondent's answers are uninterpretable (Knowles & Nathan, 1997).

Development of Balanced Scales

It has been realised that response sets do account for a certain portion of test score variance, which affects the construct validity of the instrument, therefore measures should be taken to free the instrument from this stylistic variance (Nunnally, 1978). One measure almost universally adopted in verbal self-report measures to minimise the influence of the response

sets, such as the tendency to agree or disagree, is to include an equal number of positively (e.g., "I am happy") and negatively (e.g., "I am sad") keyed items in the scale (Ahlawat, 1984). Using this method the trait under measurement is indicated by *yes*, *true*, or *agree* for half the items and *no*, *false*, or *disagree* for the other half (Cloud & Vaughan, 1970; Nunnally & Bernstein, 1994). In terms of the construct being measured, positively keyed items thus have a positive meaning, and negatively keyed items a negative meaning. When interpreting respondent's overall scores, negatively keyed items are reversed scored so that endorsing *strongly agree* or *yes* on a positively keyed item is equal to endorsing *strongly disagree* or *no* on a negatively keyed item.

While this technique does not eliminate acquiescence it does distribute it equally across the scale's items so that the trait scores are relatively free of its effects (Rundquist, 1966). This technique not only balances out the effect of acquiescence but also forces respondents to consider the content of each item carefully and respond accordingly, instead of just responding according to their general feeling about what they perceive is the intended construct (Barnette, 2000). Accordingly many psychological measurement scales have adopted the balanced item technique including the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1940), the RSES (Rosenberg, 1965), the UCLA Loneliness Scale (Russell, 1978), the Meyer and Allen Affective and Continuance Scale (Meyer & Allen, 1984), to name a few.

Issues Concerning Factor Structure

Despite their wide usage, balanced scales do present various problems concerning the item reliability, construct validity, and particularly the factorial validity of scales. Most notably it is not easy to determine that the meaning of an item has actually been reversed (Ahlawat, 1984). Rorer (1965) has provided many examples of reversed pairs of items that on close scrutiny do not turn out to be reversals.

In regards to the factor structure of balanced scales Schmitt and Schults (1985) suggested that wording changes in an effort to create a balanced scale may cause significant changes in the intended factor structure of a scale. This is often the case when factor analysis reports a two-dimensional scale structure, where a one-dimensional structure is hypothesised. Factor analysis of various balanced scales has supported this suggestion that a two-dimensional factor structure will emerge, even when the intended factor structure is one-dimensional. Rodebaugh, Woods, Thissen, Heimberg, Chambless, and Rapee (2004) investigated the factor structure of the Original Fear of Negative Evaluation Scale (FNE) and Brief Fear of Negative Evaluation Scale (BFNE), both of which use a Likert-type format, with half the items positively-keyed and half negatively-keyed. It was hypothesised that the reverse scored items on both scales would load on a distinct factor, not related to the construct under investigation. They used college undergraduates (915 completed the FNE and 1049 completed the BFNE) and tested both a one-factor model and two-factor model defined by item-keying-direction on the data. They found that for both scales the two-factor model was a significantly better fit to the data than a one-factor model. Furthermore the fit of the one-factor model on the BFNE was poor. They suggested that this two-factor model was supported as participants had problems responding to the reverse worded items due to the use of double negatives. They suggested there is a difference in how individuals respond to straightforwardly worded items versus reverse worded and this results in a two-factor model defined by item-keying-direction being supported.

This effect has been noted in various other scales with an intended one-dimensional factor structure such as the UCLA Loneliness Scale (Austin, 1983), the Social Interaction Anxiety Scale (Rodebaugh, Woods, & Heimberg, 2007), Meyer and Allen's Affective and Continuance Commitment Scales (Magazine, Williams, & Williams, 1996) and subscales of the MMPI (Messick & Jackson, 1961) to name a few. Also there have been several other

explanations for why negatively-keyed items often load on a separate distinct factor including careless responding (Schmitt & Schults, 1985), the tendency for respondents to endorse a negative item rather than reject a positive item (Campostrini & McQueen, 1993), and the fact that respondents may not perceive the subtle differences in the semantics of positive and negative items (Schriesheim & Hill, 1981). Regardless of the scale under investigation or the reason for it occurring, this issue surrounding the discrepancy between the intended and obtained factor structure of balanced scales is of much importance in psychometric research and scale development. Two scales that have undergone extensive factor analytic research are the Life Orientation Test (LOT) and its substitute the Life Orientation Test-Revised (LOT-R)

Life Orientation Test and Life Orientation Test-Revised

The LOT was first developed by Scheier and Carver (1985) to assess the construct of dispositional optimism, which they defined as positive outcome expectancies. The LOT is the most widely used scale for assessing optimism in research having been used in the US (e.g., Dolbier, Soderstom, & Steinhardt, 2001), United Kingdom (e.g., Lancaster & Boivin, 2005), Canada (e.g., Long & Schultz, 1995), the Netherlands (e.g., Tromp & Brouha, 2005), Switzerland (e.g., Irani, Mahler, Goetzmann, Russi, & Boehler, 2006), Japan (Sumi, 2004) and China (Hamid & Chang, 1996). The scale is designed to measure optimism along a continuum with high scores reflecting optimism and low scores reflecting pessimism. It consists of eight items, plus four filler items that were included to disguise the underlying purpose of the test (Scheier & Carver, 1985). Of these eight items four are keyed in a positive direction reflecting an optimistic outlook (e.g., "I'm always optimistic about my future") and four are keyed in a negative direction reflecting a pessimistic outlook (e.g., "I hardly ever expect things to go my way"). Respondents are asked to indicate their agreement with an item on a 5-point Likert scale with the following format: 4 = *strongly agree*, 3 = *agree*, 2 =

neutral, 1 = *disagree*, 0 = *strongly disagree*. Scores can range from a high of 32 to a low of zero, with all negatively keyed items being reverse scored.

A revised version of the LOT has been developed, the Life Orientation Test- Revised (LOT-R; Scheier, Carver, & Bridges, 1994), to improve some of the issues that arouse concerning the LOT, such as its high correlations with trait anxiety (Smith, Pope, Rhodewalt, & Poulton, 1989) and the independence of its predictive validity from measures of self-blame and neuroticism (Robbins, Spence, & Clark, 1991). Scheier et al. (1994) rectified these issues by removing items four "I always look on the bright side of things", and eleven "I'm a believer that every cloud has a silver lining", as it was considered that these items did not explicitly refer to the expectation of positive outcomes. Instead they referred to a particular way of reacting to problems and stress (Scheier, et al. 1994). By deleting these items, only two positively-keyed items remained so a negatively-keyed item was deleted "Things never work out the way I want them to", and another positively-keyed item was created "Overall I expect more good things to happen to me than bad". The LOT-R is thus a six item scale plus the four original fillers, with three items keyed positively and three keyed negatively. Scores can range from a high of 24 to a low of zero. Despite the revision of the LOT, both versions are still used in current research (e.g., Hart, Vella, & Mohr, 2008; Xanthopoulou, Bakker, Demerouti, & Schaufeli, 2007).

Though optimism has been associated with positive outcomes such as an increase in active coping (Mosher, Prelow, Chen, & Yackel, 2006), and an increase in social relationships (Sumi, 2006), issues concerning the factorial validity of both the LOT and LOT-R have emerged. The LOT and LOT-R were created to conform to the theoretical definition of optimism as a one-dimensional construct aligned along a continuum with pessimism in order to maintain factorial validity. However, many authors (Creed, Patton, & Bortrum, 2002; Herzberg, Glaesmer, & Hoyer, 2006; Mehrabian & Ljunggren, 1997; Vautier & Raufaste,

2006), including the authors of the LOT (Scheier & Carver, 1985), and LOT-R (Scheier, Carver, & Bridges, 1994), have found that factor analysis of both scales produces a two-factor solution. This is not in line with Hoyle's (2005) definition of factorial validity, as the factor structure of both scales does not conform to the theoretical definition of optimism. This issue has centred on whether the obtained factor structure is driven by measurement error as a result of item-keying-direction or whether it reflects differences in the content of the items. I will first discuss the factor structure of the LOT, and then reflect on the factor structure of the LOT-R.

Factor Structure of the Life Orientation Test

The original authors, Scheier and Carver (1985), created the scale and administered it to 624 undergraduates to assess its psychometric properties. Using principal-axis factor analysis they extracted two factors, the first factor defined by items keyed negatively and the second factor by items keyed positively. They subsequently tested the data using confirmatory factor analysis (CFA) to compare a one-factor model with a two-factor model defined by item-keying direction. While both the one-factor and two-factor models yielded an acceptable fit to the data, it was found that the two-factor model was significantly better. While they did suggest there was justification for examining the two halves of the scale separately, they still suggested it should be treated as one-dimensional for most purposes as the two factors extracted had a high positive correlation (r = .64) and the factors reflected item-keying direction not differences in meaningful item content. Several subsequent studies also supported the two-factor model of the LOT over the one-factor model in a range of samples (Bryant & Cvengros, 2004; Lai, Cheung, Lee, & Yu, 1998; Lai & Yue, 2000; Marshall, Wortman, Kusulas, Hervig, & Vickers, 1992; Robinson-Whelan, Kim, MacCallum, & Kiecolt-Glaser, 1997; Steed, 2002).

Factor Structure of the Life Orientation Test - Revised

Similar results have also been reported for the LOT-R, a revised version of the LOT that contains 6-items, half positively keyed and half negatively keyed, as well as four fillers. The original authors Scheier, Carver, and Bridges (1994) extracted two factors using exploratory factor analysis, with a two-factor model also being favoured over a one-factor model with subsequent CFA. Both analyses supported the idea that the two-factor model comprised of negatively and positively keyed items respectively and the factors were defined by item-keying. They did suggest that the scale be treated as one-dimensional however, as the one factor model was favoured when correlated errors among positive items were allowed. However, Vautier, Raufaste, and Cariou (2003) argued that if correlated errors are allowed in CFA these errors tell us that one-dimensionality has been lost. Other studies that suggest a one-dimensional LOT-R factor structure are plagued by issues regarding small sample sizes, inappropriate correlational techniques and reliance on unconservative goodness-of-fit indexes (Mehrabian & Ljunggren, 1997; Rauch, Schweizer, & Moosbrugger, 2007). Similar to the LOT, studies with acceptable sample sizes found data from the LOT-R yielded a poor to moderate one-factor model fit and an acceptable to high two-factor model fit (Creed, Patton, & Bortrum, 2002; Herzberg, Glaesmer, & Hoyer, 2006; Mehrabian & Ljunggren, 1997; Vautier & Raufaste, 2006). Therefore, while the original authors still support one-dimensionality, they and others suggest that the responses to the LOT-R can be scored along two subscales reflecting trait optimism and trait pessimism (Scheier et al.1994; Burke, Joyner, Czeck, & Wilson, 2000). Once again some argue though that these two factors emerge due to measurement error as a result of item-keying-direction and not meaningful differences in item content (Nakano, 2004).

Competing Explanations for Two-Factor Structure

While Scheier and Carver (1985) found the one-factor model yielded an acceptable fit to their data, other investigations suggest the one-factor model fit was moderate to poor (Bryant &

Cvengros, 2004; Marshall & Lang, 1990; Marshall, Wortman, Kusulas, Hervig, & Vickers, 1992; Robinson-Whelan, Kim, MacCallum, & Kiecolt-Glaser, 1997). Therefore, while it has been consistently found that a two-factor model provided the best fit for the LOT and LOT-R data, research is divided on what causes the items to divide into two-factors, one defined by optimism and the other by pessimism. Some have argued that the two-factor solution emerges due to measurement error as a result of items being keyed in opposite directions, resulting in positively keyed, optimism items loading on one factor and negatively keyed, pessimism items loading on another factor (Chang & McBride-Chang, 1996; Scheier & Carver, 1985; Scheier, Carver, & Bridges, 1994). Others, however, have argued that the two-factor solution emerges not as a result of measurement error, but due to differences in the content of positive and negative items, resulting in positively keyed, optimism items measuring a different aspect of the construct than negatively keyed, pessimism items (Lai, 1994; Mook, Kleijn, & Van Der Ploeg, 1992). Both of these arguments are discussed below, with a suggestion for future research to further determine which argument is most supported.

Item-Keying Direction

The argument that the two factors emerge due to measurement error and not differences in item content comes from studies that investigated the LOT based on the idea of connotatively consistent (CC) and connotatively inconsistent (CI) items (Chang, 1995a). Chang viewed CI items as those that did not have the same keying direction as the majority of other items on the scale. Chang (1995b) argued that in relation to the LOT a rating of "0" for a negatively keyed item, would not truly be equal to a "5" if the item was reworded to be positively keyed. He suggested that the two-factor model of the LOT may be caused by the use of CI items that do not measure the same aspects of the construct as CC items due to their inconsistent connotation, therefore endorsing an item reflecting optimism would not be equivalent to disagreeing with an item reflecting pessimism. Chang and McBride-Chang (1996) continued

this line of research by investigating the factor structure of the LOT in its original form and when the items were reworded to reflect consistent pessimism or consistent optimism. The consistent optimism version was created by rewriting the four pessimism items to reflect optimism. For instance, the item "Things never work out the way I want them to" was changed into "Things always work out the way I want them to". The consistent pessimism version was created by rewriting the four optimism items to reflect pessimism. For example, "In uncertain times, I usually expect the best" became "In uncertain times, I usually expect the worst". The two rewritten versions therefore contained CC items, whereas the original contained CI items. Chang and McBride-Chang had undergraduates complete one version of consistent pessimism (n = 149), or consistent optimism (n = 129) on a 4-point scale ranging from 1 = *strongly disagree* to 4 = *strongly agree*. Another sample completed the original LOT (n = 108) using both a 4-point and 6-point scale. Confirmatory factor analysis indicated an overwhelming superiority of the two-factor model on the original; however the two CC versions supported the one-factor model. They suggested that the two-factor structure of the LOT was therefore a result of response bias due to the use of oppositely-keyed items, rather than a reflection of substantial differences in item content. It could not be determined whether this factor structure was driven by meaning or framing though as both were changed in the rewritten versions. Therefore, as meaning and framing were not split it is unknown whether respondents responded to the content of the item, the framing, or both. Also it is unknown why a 4-point and 6-point scale were used instead of the original 5-point scale and why the response options were reversed, to indicate strongly disagree first. It is possible these subtle changes of the LOT may have confounded the obtained results.

Allan and Giles (2008) addressed the issue further by investigating the factor structure of the LOT in a sample of West Australian prisoners to determine whether it was affected by measurement error. They read items of the LOT to (n = 453) participants, which the

participants responded to on a 5-point Likert scale. The results of CFA indicated a two-factor model corresponding to item-keying-direction was a better fit to the data than a one-factor model, once again demonstrating that the structure of the LOT is not in line with the one-dimensional construct it was intended to measure. When they removed participants who demonstrated a tendency to consistently agree or disagree however, the fit of the one-factor model improved while the two-factor model fit decreased. They suggested this indicated that the two factor structure of the LOT is a result of measurement error due to item-keying-direction rather than reflecting substantial differences in the content of positively and negatively keyed items.

Differences in Item Content

Furthermore it has been suggested that the two-factor model emerges due to differences in the content of positive and negative items, rather than reflecting measurement error due to their opposite keying directions (Lai, 1994). Mook, Kleijn, and Van Der Ploeg (1992) suggested that rather than measuring optimism, negatively keyed items measured a lack of pessimism. They argued that agreeing that one rarely feels pessimistic is not equivalent to saying one feels optimistic nearly all the time. However, their sample size was much smaller than in the original assessment of the LOT and they did not conduct the more stringent CFA on their data, therefore their results may be misleading.

Kubzansky, Kubzansky, and Maselko (2004) followed on from Chang and McBride-Chang (1996) by distinguishing between framing and meaning in their study. Kubzansky et al. (2004) aimed to investigate whether the two-factor structure of the LOT was determined by response bias due to item-keying direction or differences in meaningful content. They stated that a positively framed item uses words with positive connotations (e.g., always or right), whereas a negative framed item uses words with negative connotations (e.g., never or wrong). They suggested that an item could be negatively framed while still maintaining its

positive meaning and vice versa. For example the item "I'm never pessimistic about my future" is negatively framed but has a positive meaning. In their study they teased apart the meaning and framing of items, by administering three versions of the LOT to undergraduates. The original version was administered (n = 146), as well as the two reworded versions. The first (n = 141) was derived by changing the framing of half of the items on each subscale but preserving the meaning. This resulted in each subscale having two negatively-framed and two-positively-framed items while still maintaining consistent optimistic or pessimistic meaning. The other (n = 142) was derived by changing the framing of all items but preserving the meaning so that framing and meaning were inconsistent. They tested three models with CFA; the bipolar model, reflecting a one-dimensional view of the LOT, a method artifact model, that suggested items cluster together because they are similarly framed and a bivariate model that suggests items cluster together because they have similar meaning. The bipolar model had an unacceptable fit for all three versions. However for the two derived versions the bivariate model was a better fit than the method artifact model. Items with optimistic content loaded on one factor and items with pessimistic content loaded on another factor regardless of which direction they were framed. They suggested this strongly implies the factor structure of the LOT is driven by item meaning and not measurement error due to item-keying direction and that optimism as measured by the LOT does not exist along a single continuum with pessimism. One problem with this study is that all versions administered were CI, not CC, as has been recommended by Chang (1995a) and Chang (1995b). It is assumed that the LOT-R would perform similarly under these conditions, due to similar results found for the LOT and LOT-R in other studies that assessed factor structure.

Current Research

There is research that supports both positions in the debate over whether the two-factor structure of the LOT and LOT-R emerges due to measurement error or substantial differences in item content. The present study was designed to examine how the factor structure is affected by keeping framing constant (all items either positively or negatively framed) while preserving the CI content of the items (retaining the meaning of positively and negatively keyed items). It was anticipated that the two-factor model would provide the best fit to the data as this is consistent with other research where the content of the LOT items was CI (Chang, 1995; Chang, & McBride-Chang, 1996). While no study has looked at the effect of consistently framed items, the items still contained mixed content, which usually produces a two-factor structure. If the results do not support a one-factor structure even when all items are consistently framed, then framing could be considered to have little effect on the factor structure of the LOT. If the results do support a one-factor structure it will suggest that framing does affect the factor structure of the LOT and it should be kept consistent if the test is still to be considered as measuring optimism-pessimism as a one-dimensional construct.

Method

Research Design

The research involved a quantitative investigation of the psychometric properties of the LOT and LOT-R. The study looked specifically at the factor structure of both tests, which enabled its factorial validity to be investigated.

Participants

Participants were first, second and third year undergraduate psychology students enrolled at Edith Cowan University in Western Australia. A convenience sample was drawn from this specific population rather than a random community sample because responses to the LOT and LOT-R can vary depending on the personal situation of an individual (Schulz & Tompkins, 1988). First year students (n = 100) completed the negatively framed version of

the LOT/LOT-R and second and third year students (n = 100) completed the positively framed version. The final sample consisted of 152 females (76.0%) and 48 males (24.0%). Participants were predominantly Caucasian (Caucasian = 81.5%, Asian = 12%, Other = 6%), primarily born in Australia (Born in Aus = 65%, Born overseas = 35%) and the majority were aged 18-24 years.

Materials

Two modified versions of the LOT and LOT-R (Scheier & Carver, 1985; Scheier, Carver, & Bridges, 1994) were created and administered as a single 13-item scale. In the original LOT and LOT-R the framing of items matched their content, where as items in the current study were consistently framed negatively or positively irrespective of their content. This was done by taking the original LOT items and combining them with the reversed frame items from their opposite subscale created by Kubzansky, Kubzansky, and Maselko (2004; Appendix A). This procedure resulted in each modified version still containing four optimism and four pessimism items, but all of the items were either positively framed or negatively framed (Appendix C; Appendix D). To enable the LOT-R to be tested, the additional positively keyed item was added at the end of each modified scale in either a positively framed or negatively framed format. For the fourth optimism item "I'm a believer in the idea that every cloud has a silver lining" no reverse frame item could be created without affecting the item's meaning, therefore it was left in its original form (Kubzansky, Kubzansky, & Maselko, 2004). Extra demographic questions such as age, gender, country of birth and ethnicity were also included in the questionnaire.

Procedure

Participants were obtained by seeking permission from lecturers to present the research during a lecture and ask if students wished to participate (Appendix B). Participants from first year classes were approached with the positively framed version of the scale and second and

third year classes with the negatively framed version. This was done to ensure that participants did not complete both versions of the scale. Participants were also provided with an information letter describing the scale and whom they can contact in regards to the research. The information letter advised students that by completing the scale they were giving their informed consent to participate in the research, which they were told was investigating the levels of optimism and pessimism in first, second and third year students. The study received ethical approval from both the Edith Cowan and School of Psychology Ethics Board.

Analysis

Data from the two modified versions of the LOT was processed using PRELIS 2.72 (Joreskog & Sorbom, 2005), to generate polychoric correlation matrices for the test items. As the data is ordinal, polychoric correlations and diagonally weighted least squares estimations were utilised as suggested by Wang and Cunningham (2005). Next, confirmatory factor analysis (CFA) was conducted on both sets of data using LISREL 8.80 (Joreskog & Sorbom, 2007). In a CFA, a priori structure is posited and how well the data fits the structure is tested. The purpose was to evaluate the two competing interpretations reported for the factor structure of dispositional optimism (as measured by the modified versions of the LOT and LOT-R), that is Model 1, that the LOT is a one-dimensional measure (Scheier & Carver, 1985), and Model 2, that the LOT has a two-factor structure. For Model 1, the eight LOT items and six LOT-R items were allowed to load freely on a single latent factor representing Optimism. For Model 2 the four optimism items of the LOT were allowed to load freely on a latent factor representing Optimism, and the four pessimism items were allowed to load freely on a latent factor representing Pessimism. The same procedure was used for the three optimism and three pessimism items of the LOT-R. The quality of fit for each model was assessed using the following goodness-of-fit measures; the root mean square error of

approximation (RMSEA); the goodness-of-fit index (GFI); the adjusted goodness-of-fit index (AGFI); the non-normed fit index (NNFI); and the chi-square result.

The interpretation guidelines suggested by Marsh, Balla, and McDonald (1988) were used when evaluating the GFI, AFI and NNFI, all of which must be greater than .90 to provide an acceptable fit, and as suggested by Wegener and Fabrigar (2000) an NNFI above .95 to provide a close fit. RMSEA values less than .05 were interpreted as an appropriate fit, values between .05 and .08 were interpreted as a reasonable fit, while values over .1 were considered a poor fit to the data (Browne & Cudeck, 1993). Bollen and Lang (1993) suggested that a chi-square greater than .05 is indicative of good fit, whereas a chi-square less than .05 suggests a poor fit. All these measures were considered for both models, as CFA is best used when testing rival models (e.g., one-factor vs. two-factor) and results are strengthened when various statistical-fit-indices are acceptable (Thompson, 2004). Finally to test whether Model 1 or Model 2 provided a significantly better fit, a nested model test (Bentler & Bonett, 1980) comparing the fit of each model to the other was conducted for each modified version.

Results

The test scores on the positively framed version ranged from 6 – 31 on the LOT ($M = 20.96$, $S.D. = 4.94$), with a Cronbach alpha of .80 and from 5 – 23 on the LOT-R ($M = 15.91$, $S.D. = 3.94$), with a Cronbach alpha of .60. Test scores on the negatively framed version ranged from 10 - 29 on the LOT ($M = 20.09$, $S.D. = 4.10$), with a Cronbach alpha of .52 and from 6 - 24 on the LOT-R ($M = 15.54$, $S.D. = 3.51$), with a Cronbach alpha of .50. These were comparable to the normative scores on the LOT ($M = 21.03$, $S.D. = 4.57$), and the LOT-R ($M = 14.33$, $S.D. = 4.28$) (Scheier & Carver, 1985; Scheier, Carver, & Bridges, 1994).

Confirmatory factor analyses (CFA) were performed separately on the positively framed and negatively framed versions of the LOT and LOT-R data, to compare the fit of

one- and two-factor models. Figure 1 is an example of a one-factor model and Figure 2 is an example of a two-factor model.

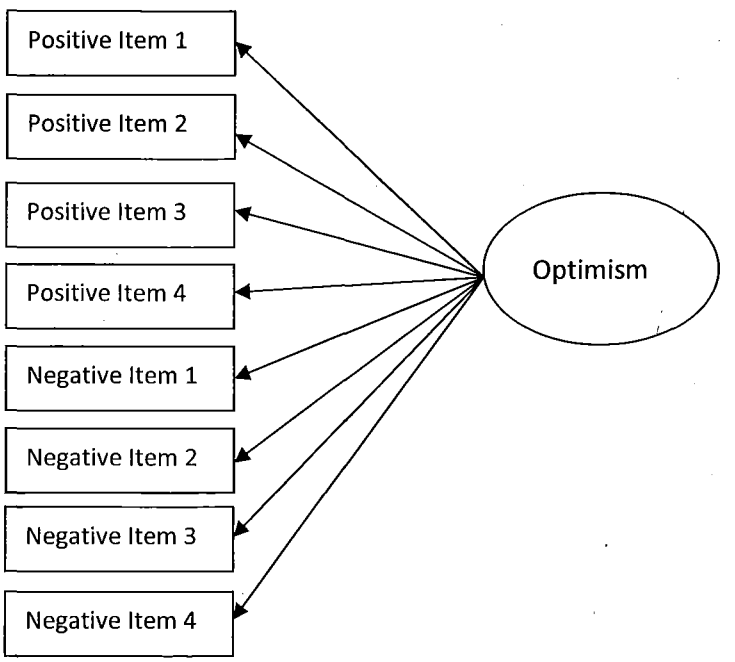

Figure 1. One-factor LOT Model.

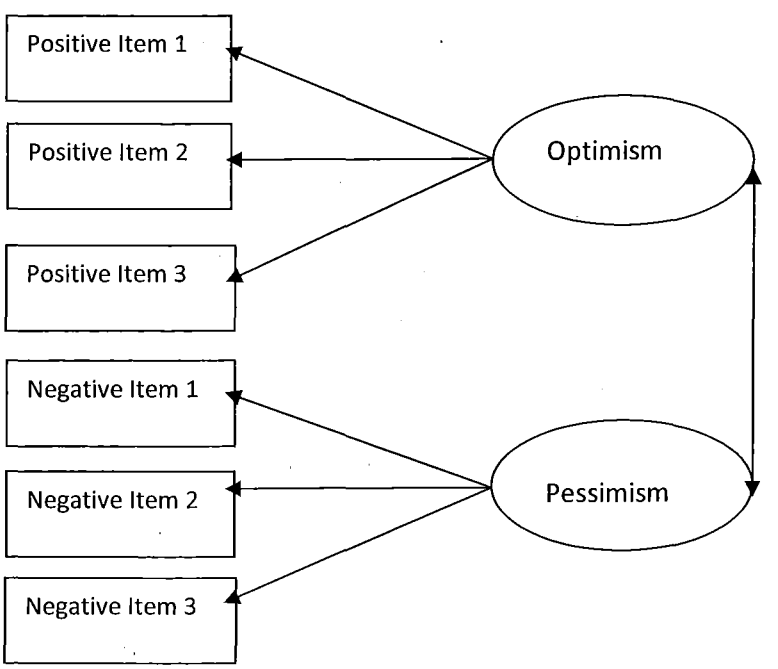

Figure 2. Two-factor LOT-R Model.

The CFA results are provided in *Table 1* along with the degrees of freedom (df), internal reliability (ρ), internal consistency (Cronbach alpha; α), and latent factor correlation (*r*) of each model. The construct reliability of each model was analysed using the guidelines proposed by Fornell and Larcker (1981) who suggest that as a general rule, construct reliability should exceed .50 if researchers want to estimate how reliably the model *indicates* the latent construct. Results indicated all models had moderate to acceptable construct reliability. An inspection of Table 1 reveals that across both versions and models the AGFI, GFI and NNFI indicated a good fit. Therefore the RMSEA and chi-square are discussed as they give the most information concerning the differences in the models.

Table 1.

Fit indices for Model 1 and Model 2 across the LOT and LOT-R

Model	χ^2	df	p	RMSEA	GFI	AGFI	NNFI	*r*	α	ρ
					Positive Version (n = 100)					
1-F LOT	36.90*	20	.01	.09	.99	.98	.96		.80	.88
2-F LOT	17.81	19	.53	.00	1.00	.99	1.00	.68	.80	.72/.91^
1-F LOT-R	21.61*	9	.01	.12*	.98	.96	.93		.60	.85
2-F LOT-R	10.70	8	.22	.06	.99	.97	.98	.74	.60	.68/.87^
					Negative Version (n = 100)					
1-F LOT	26.49	20	.15	.06	.97	.95	.97		.52	.80
2-F LOT	22.59	19	.26	.04	.98	.95	.99	.69	.52	.43/.87^
1-F LOT-R	14.88	9	.09	.08	.98	.96	.96		.50	.79
2-F LOT-R	11.06	8	.20	.06	.99	.97	.98	.78	.50	.50/.85^

*Poor fit

^ For two factors models; optimism item reliabilities are listed first and pessimism item reliabilities are listed second.

Positively Framed Version

On the positively framed version the two-factor model provided an appropriate fit to the data χ^2 (19, N = 100) = 17.80, p > .05 from the LOT, whereas the chi-square χ^2 (20, N = 100) = 36.90, p < .05 and RMSEA indicated a poor fit of the one-factor model. The chi-square difference test yielded a significant result χ^2 (1, N = 200) = 19.09, p <.05, indicating that the models are different and the two-factor model serves to better explain the positively framed LOT data.

On the LOT-R data from the positively framed version all fit indices from the two-factor model indicated an appropriate fit, except the RMSEA (.06) which indicate a reasonable fit, where as the chi-square χ^2 (9, N = 100) = 21.61, p < .05 and RMSEA (.12) indicated a poor fit of the one-factor model. Once again the chi-square difference test yielded a significant result χ^2 (1, N = 200) = 10.91, p < .05, indicating that the models are different and the two-factor model serves to better explain the positively framed LOT-R data.

The two-factor models had moderate latent factor correlations between optimism and pessimism on both the LOT (r = .68) and LOT-R (r = .74). Comparison of the LOT and LOT-R two-factor models via a chi-square difference test did not yield a significant result, indicating there was no difference in the fit of the two-factor models.

Negatively Framed Version

On the negatively framed version the two-factor model and one-factor model provided an appropriate fit to the data from the LOT, although the one-factor model RMSEA (.06) indicated a reasonable fit. Results of the chi-square difference test indicated a significant result χ^2 (1, N = 100) = 3.9, p < .05, indicating that the models are different and the two-factor model serves to better explain the negatively framed LOT data.

On the LOT-R data from the negatively framed version the one-factor fit was reasonable with the chi-square χ^2 (9, N = 100) = 14.88, p = .094 and RMSEA (.08) indicating a reasonable fit to the data. The two-factor LOT-R was considered an appropriate fit, with all fit indices indicating an appropriate fit except the RMSEA (.06), which was reasonable. However for data on the negatively framed LOT-R the chi-square difference test χ^2 (1, N = 200) = 3.82, p > .05 did not indicate a significant difference between the fit of the one-factor and two-factor models.

Once again both of the two-factor models had high latent variable correlations on the LOT (r = .69) and LOT-R (r = .78). Comparison of the LOT and LOT-R two-factors models via a chi-square difference test did not yield a significant result, indicating there was no difference in the fit of the two-factor models.

Comparison of the Positive and Negative Versions

Comparisons of the one-factor models and two-factor models across both versions of the LOT and LOT-R were conducted using a chi-square difference test. For the LOT the one-factor model was a significantly better fit to the positive version compared to the negative version χ^2 (1, N = 200) = 10.41; p < .01, and the two-factor model was a significantly better fit to the negative version compared to the positive version χ^2 (1, N = 200) = 4.78, p < .05.

For the LOT-R the one-factor model was a significantly better fit to the positive version compared to the negative version χ^2 (1, N = 200) = 6.73, p < .01, but for the two-factor model the difference in fit between each version was not significant χ^2 (1, N = 200) = .36, p > .05.

Discriminant Validity

As all the two-factor models had high latent correlations between the latent factors Optimism and Pessimism, an analysis of discriminant validity was completed for each of the two-factor

models. According to Fornell and Larcker (1981) if the average variance extracted from the two latent factors is higher than the latent correlation squared, then the model has acceptable discriminant validity. Discriminant validity results are presented in Table 2. An examination of Table 2 indicated that the two-factor LOT and two-factor LOT-R had acceptable discriminant validity on the positively framed version, but unacceptable discriminant validity on the negatively framed version.

Table 2.

Discriminant validity analysis.

Model	Average variance extracted		Correlation squared
		Positive Version (N = 100)	
2-Factor LOT	.39	<	.46
2-Factor LOT-R	.43	<	.55
		Negative Version (N = 100)	
2-Factor LOT	.43	<	.48
2-Factor LOT-R	.35	<	.61

Discussion

The mean LOT and LOT-R scores obtained from both modified versions fell in the range obtained in the normative sample (Scheier & Carver, 1985; Scheier, Carver, & Bridges, 1994). While the LOT-R scores were slightly higher in the student population, they were still similar to those of other populations (Creed, Patton, & Bortrum, 2002; Herzberg, Glaesmer, & Hoyer, 2006). All scales had acceptable internal reliability (see Table 1) and were similar to those of other studies (Dolbier, Soderstrom, & Steinhardt, 2001; Lancastle & Boivin, 2005).

Reasons for Two-Factor Structure

The aim of the current study was to see how manipulating the framing of items on the LOT and LOT-R would affect the scale's factor structure. The modified versions contained either

all positively or all negatively framed items, while still maintaining their optimistic or pessimistic content. As hypothesised the two-factor model defined by item content was a more appropriate fit to the data compared to a one-factor model. Across both modified versions the two-factor model was an appropriate fit on both the LOT and LOT-R and on the positively framed version the one-factor model was a poor fit. Furthermore on all versions except the negatively framed LOT-R the two-factor model fit was significantly better than the one-factor model. These findings indicated that despite all items being consistently framed, items still loaded on two separate factors according to their optimistic or pessimistic content.

Differences in Item Content

These findings support research by Kubzansky, Kubzansky, and Maselko (2004) who suggested that participants respond to the meaning of items and that the two factors are in line with the different content of the items that reflect an optimistic or pessimistic outlook respectively. They created a version of the LOT where each subscale had two positively framed and two negatively framed items but the content of all items was retained. They found that items with positive content loaded on one factor and items with negative content loaded on another factor, regardless of framing. This is similar to the current study in that despite all items being consistently framed they still loaded on separate factors according to their optimistic (positive) or pessimistic (negative) content. This suggests that framing does not contribute to the factor structure of the LOT and LOT-R and instead participants respond to the optimistic or pessimistic meaning of the items. Those who support this idea suggest that instead of measuring optimism along a bipolar continuum, the positively keyed items independently measure optimism and the negatively keyed items independently measure pessimism, as the content of the items is different and each set of items loads on a separate factor (Kubzansky, et al. 2004; Vautier & Raufaste, 2006)

Measurement Error

Even though the two-factor structure defined by item content is supported, this does not mean that optimism as measured by the LOT and LOT-R is necessarily a two-dimensional construct. It is argued the two-factor structure may be a result of measurement error due to the use of CI items (Chang, 1995a). Chang (1995a) suggested that agreeing with a positive item is not the same as disagreeing with a negative item and that the two sets of items would load on separate factors due to their inconsistent connotation.

The modified versions in the current study had consistent positive or negative framing but the items were still CI due to their opposite-keying direction. Chang and McBride-Chang (1996) found that when items on the LOT were reworded to reflect consistent optimism (CC) or consistent pessimism (CC) a one-factor model was a better fit to the data than a two-factor model. However when the scale contained both optimism and pessimism items (CI) a two-factor model was an overwhelming better fit to the data. They suggested that the two-factor structure of the LOT was therefore a result of response bias due to the oppositely-keyed items, rather than a reflection of differences in item content. It is possible that the two-factor structure found for the LOT and LOT-R in the following study is therefore a result of measurement error rather than substantial differences in item content as both modified versions contained CI items.

Also as items were oppositely keyed the two-factor structure could have been driven by measurement error due to the tendency to agree or disagree. Allan and Giles (2008) found that by removing respondents who had a tendency to agree or disagree, the one-factor LOT structure was supported. It is possible that this response set drove the two-factor structure in the current study as items were oppositely keyed and agreeing or disagreeing respondents were not removed. Furthermore it is suggested by Campostrini and McQueen (1993) that respondents are more likely to endorse a negatively keyed item rather than reject a positively keyed item due to the influence of social desirability. As modified versions in the current

study contained both positively and negatively keyed items this response set could have caused the two-factor structure, as a result of measurement error.

Although it is unclear whether differences in item content or item-keying caused the two-factor structure, it is clear that participants did not respond to the framing of items on the LOT and LOT-R. Instead they responded to either the positive or negative content of the items, referring to optimism and pessimism respectively.

Latent Factor Correlation

Further support for a two-factor structure is indicated by the investigation of the latent factor correlation between Optimism and Pessimism in the two-factor models. Scheier and Carver (1985) in their original investigation of the LOT reported that despite a two-factor model being a better fit to the data than a one-factor model, the one-factor model should be considered favourable due to the high latent factor correlation ($r = .64$) between Optimism and Pessimism. Scheier, Carver, and Bridges (1994) also suggested that the LOT-R should be considered one-dimensional despite a two-factor model indicating a significantly better fit to the data than a one-factor model. They also reported a high latent factor correlation but that the one-factor model was a better fit to the data when correlated errors between the two latent factors were allowed.

The reasoning of the original LOT and LOT-R authors needs to be criticised as a one-factor model cannot be justified when a two-factor model is better supported by the data, simply because of a high latent factor correlation (Fornell & Larcker, 1981). Fornell and Larcker (1981) suggest that latent factor correlations in the vicinity of .80 and .90 can be indicative of a one-factor structure, however the correlations reported in the normative samples of the LOT and LOT-R are only moderate compared to this. Also allowing correlated errors in CFA tells us that one-dimensionality has been lost (Vautier, Raufaste, & Cariou, 2003).

In the current sample the latent factor correlations for the positive version and negative version were slightly higher than reported by the original authors on both the LOT and LOT-R. To provide evidence that this correlation between the latent factors Optimism and Pessimism was not indicative of a one-dimensional structure the discriminant validity of the two-factor models was investigated. Whiteley (2002) stated that if a model has acceptable discriminant validity it suggests that the latent factors are measuring separate constructs. On the positive version the two-factor LOT and LOT-R models had acceptable discriminant validity providing further support for the two-factor model as the latent factors discriminated between optimism and pessimism indicating they are different. On the negative version the two-factor LOT and LOT-R models indicated a lack of discriminant validity. It is suggested this does not indicate that the latent factors are measuring the same construct but instead that participants responded differently to negatively framed items. As the one-factor model fit on the negatively framed LOT and LOT-R was poor to moderate, it cannot be argued that the scales are in fact one-dimensional due only to a moderate latent factor correlation and lack of discriminant validity (Fornell & Larcker, 1981).

Though the original authors did not test the discriminant validity of their two-factor models the current study suggests, at least for the positively framed versions, that the two latent factors have enough discriminant validity to indicate that they measure optimism and pessimism separately. Further, the endorsement of the one-factor model in the original LOT and LOT-R samples is criticised as a one-factor model cannot be supported when it is a poor fit to the data just because a two-factor model has a moderate latent factor correlation.

Problems with Negatively Framed Items

Results of the current study indicated that when items on the LOT and LOT-R were CI, negative framing was related to further deviation away from a one-factor structure. As mentioned earlier both negatively framed versions lacked discriminant validity and internal

consistency, suggesting there may be problems associated with negatively framed items. Comparison of positively framed and negatively framed model fits indicated that on the LOT and LOT-R the fit of a one-factor model was better for the positively framed version but the fit of a two-factor model was better for the negatively framed versions. This demonstrates that on both scales a combination of negative framing and CI items resulted in a decreased support of a one-factor structure, while a combination of positive framing and CI items caused an increased support of a one-factor structure.

This is similar to findings by Chang and McBride-Chang (1996) who found that when all items on the LOT were rewritten to be positively or negatively CC a one-factor model was a better fit to the positive version than the negative version. Rodebaugh (2004) suggested problems with negatively framed items may occur because participants find it harder to respond to the double negatives they employ compared to straightforwardly worded positively framed items. Scheier, et al. (1994) also reported concern regarding negatively framed items in their study of the LOT-R. They found a higher degree of shared disturbance between positively framed items than negatively framed items. They proposed that participants may have issues responding to negatively framed items due to their increased semantic complexity and that using negatively framed items caused more measurement error or made the impact of it on the scale's factor structure greater.

These findings are further supported in the current study as positively keyed items on the negative versions had the lowest internal reliability, perhaps because they were the most semantically complex items, (e.g., "Even in uncertain times, I don't expect the worst"). It is apparent that even though participants seemed to respond to the meaning of items on the LOT and LOT-R irrespective of consistent framing, the use of negative framing further impacts on the factor structure of the scales due to the influence of measurement error.

Limitations of the Current Research

Other investigations of the LOT and LOT-R have used sample sizes as large as 450 (Allan & Giles, 2008), to over 2000 (Scheier, Carver, & Bridges, 1994). While the sample size in the current study was not this large it was comparable in size to that reported by Chang (1995b), an average of 150 for each version, and within the acceptable range of ten participants per item for CFA (Thompson, 2004). The use of a strictly student population may also have been problematic but this is also a criticism of other studies that have investigated the LOT and LOT-R (Chang, 1995b; Scheier, Carver, & Bridges, 1994). The two-factor structure has been supported in populations of prisoners (Allan & Giles, 2008), and older women (Sharpe, Hickey, & Wolf, 1994), therefore the student sample is perhaps not so problematic.

Areas for Future Research

It seems from the current study that the factor structure of the LOT and LOT-R may be influenced by measurement error in spite of consistent framing. Barnette (2000) has suggested that a possible way to eliminate/reduce measurement error that results from the use of balanced scales is to vary the direction of the response options. Accordingly half of the items would have a response scale, *strongly agree* to *strongly disagree*, and the other half, *strongly disagree* to *strongly agree*. Future research could use positively framed and positively keyed items (CC) as suggested by Chang (1995a), and vary the response options to see whether a one-factor or two-factor model is a better fit to the LOT and LOT-R.

Conclusions of the Current Research

In conclusion, the current study contributes information about the influence of item framing on the factor structure and subsequent factorial validity of the LOT and LOT-R. It appears that participants respond to the meaning of items rather than the framing and this meaning drives the two-factor structure. It is still unclear whether the two-factor structure emerges due to substantial differences in item content or due to measurement error as a result of oppositely-keyed items, but it is clear that, as they stand, the LOT and LOT-R fail to measure

optimism along a bipolar continuum. Also the current study has provided more information regarding the factor structure of the LOT-R as not as much research has been performed on it compared to the LOT. It appears that results of factor analysis on the LOT are applicable to the LOT-R as results of the current study were similar across both scales and have been similar in other studies (Scheier & Carver, 1985; Scheier, Carver, & Bridges, 1994).

References

Ahlawat, K. S. (1984). On the negative valence items in self-report measures. *The Journal of General Psychiatry*, 112, 89-99.

Allan, M. M., & Giles, M. (2008). Psychometric properties of Scheier and Carver's Life Orientation Test in a samples of West Australian prisoners. *Psychological Reports*, 103, 305-322.

Austin, B. A. (1983). Factorial structure of the UCLA Loneliness Scale. *Psychological Reports*, 53, 883-889.

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement*, 60, 361-370.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.

Bollen, K., & Lang, J. (1993). *Testing structural equation models*. Newbury Park, CA, Sage.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In A. Kenneth, J. Bollen, & S. Long. (Eds.), *Testing structural equations models* (pp. 136-160). Thousand Oaks, CA: SAGE.

Bryant, F. B., & Cvengros, J. A. (2004). Distinguishing hope and optimism: Two sides of a coin, or separate coins? *Journal of Social and Clinical Psychology*, 23, 273-302.

Campostrini, S., & McQueen, D. V. (1993). The wording of questions in a CATI-based lifestyle survey: Effects of reversing the polarity of AIDS-related questions in continuous data. *Quality & Quantity*, 27, 157-170.

Chang, L. (1995a). Connotatively consistent and reverse connotatively inconsistent items are

not fully equivalent: Generalizability study. *Educational and Psychological Measurement*, 55, 991-997.

Chang, L. (1995b). Connotatively inconsistent test items. *Applied Measurement in Education*, 8, 199-209.

Chang, L., & McBride-Chang, C. (1996). The factor structure of the LOT. *Educational and Psychological Measurement*, 56, 325-329.

Cloud, J., & Vaughan, G. M. (1970). Using balanced scales to control acquiescence. *Sociometry*, 33, 193-202.

Creed, P. A., Patton, W., Bartrum, D. (2002). Multidimensional properties of the LOT-R: Effects of optimism and pessimism on career and well-being related variables in adolescents. *Journal of Career Assessment*, 10, 42-61.

Dolbier, C. L., Soderstrom, M., & Steinhardt, M. A. (2001). The relationship between self-leadership and enhanced psychological, health and work outcomes. *The Journal of Psychology*, 135, 469-485.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-49.

Hamid, P. N., & Cheng, S. (1996). The development and validation of an index of emotional disposition and mood state: The Chinese affect scale. *Educational and Psychological Measurement*, 56, 998-1014.

Hart, S. L., Vella, L., & Mohr, D. C., (2008). Relationships among depressive symptoms, benefit-finding, optimism, and positive affect in multiple sclerosis patients after psychotherapy for depression. *Health Psychology*, 27, 230-238.

Hathaway, S. R., & McKinley, J. C. (1940). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology*, 10, 249-254.

Herzberg, P. Y., Glaesmer, H., & Hoyer, J. (2006). Separating optimism and pessimism: A robust psychometric analysis of the Revised Life Orientation Test (LOT-R). *Psychological Assessment*, 18, 433-438.

Hjelle, L., Belongia, C., & Nesser, J. (1996). Psychometric properties of the Life Orientation Test and Attributional Style Questionnaire. *Psychological Reports*, 78, 507-515.

Hoyle, R. H. (2005). Applications of structural equation modelling in clinical and health psychology research. In J. Miles & P. Gilbert (Eds.), *Handbook of research methods for clinical and health psychology* (pp. 265-275). Oxford: Oxford University Press.

Irani, S., Mahler, C., Goetzmann, L., Russi, E. W., & Boehler, A. (2006). Lung transplant recipients holding companion animals: Impact on physical health and quality of life. *American Journal of Transplantation*, 6, 404-411.

Joreskog, K., & Sorbom D. (2005). *PRELIS (Version 2.72)*. Lincolnwood, IL: Scientific Software International, Inc.

Joreskog, K., & Sorbom, D. (2007). *LISREL Student Edition (Version 8.80)*. Lincolnwood, IL: Scientific Software International, Inc.

Kubzansky, L. D., Kubzansky, P. E., & Maselko, J. (2004). Optimism and pessimism in the context of health: Bipolar opposites of separate constructs. *Personality and Social Psychology Bulletin*, 30, 943-956.

Lai, J. C. L. (1994). Differential predictive power of the positively versus the negatively worded items of the Life Orientation Test. *Psychological Reports*, 75, 1507-1515.

Lai, J. C. L., Cheung, H., Lee, W., & Yu, H. (1998). The utility of the revised Life Orientation Test to measure optimism among Hong Kong Chinese. *International Journal of Psychology*, 33, 45-56.

Lai, J. C. L., & Yue, X. (2000). Measuring optimism in Hong Kong and mainland Chinese

with the revised Life Orientation Test. *Personality and Individual Differences*, 28, 781-796.

Lancaster, D., & Boivin, J. (2005). Dispositional optimism, trait anxiety, and coping: Unique or shared effects on response to fertility treatment. *Health Psychology*, 24, 171-178.

Long, B. C., & Schutz, R. W. (1995). Temporal stability and replicability of a workplace stress and coping model for managerial women: A multiwave panel study. *Journal of Counseling Psychology*, 42, 266-278.

Magazine, S. L., Williams, L, J., & Williams, M. L. (1996). A confirmatory factor analysis examination of reverse coding effects in Meyer and Allen's Affective and Continuance Commitment Scales. *Educational and Psychological Measurement*, 56, 241-250.

Marsh, H., Balla, J., & McDonald, R. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391-410.

Marshall, G. N., & Lang, E. L. (1990). Optimism, self-mastery, and symptoms of depression in women professionals. *Journal of Personality and Social Psychology*, 59, 132-139.

Marshall, G. N., Wortman, C. B., Kusulas, J. W., Hervig, L. K., & Vickers, R. R. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of Personality and Social Psychology*, 62, 1067-1074.

Mehrabian, A., & Ljunggren, E. (1997). Dimensionality and content of optimism-pessimism analysed in terms of the Pad Temperament Model. *Personality and Individual Differences*, 23, 729-737.

Mels, G. (2006). *LISREL for Windows: Getting Started Guide*. Lincolnwood, IL: Scientific Software International, Inc.

Messick, S., & Jackson, D. N. (1961). Acquiescence and the factorial interpretation of the Minnesota Multiphasic Personality Inventory. *Psychological Bulletin*, 58, 299-304.

Meyer, J. P., & Allen, N. J. (1984). Testing the "side-bet-theory" of organizational commitment: Some methodological considerations. *Journal of Applied Psychology*, 69, 372-378.

Mook, J., Kleijn, W. C., & Van Der Ploeg, H. M. (1992). Positively and negatively worded items in a self-report measure of dispositional optimism. *Psychological Reports*, 71, 275-278.

Mosher, C. E., Prelow, H. M., Chen, W. W., & Yackel, M. E. (2006). Coping and social support as mediators of the relation of optimism to depressive symptoms among black college students. *Journal of Black Psychology*, 32, 72-86.

Sociometry, 33, 193-202.

Nunnally, J. C. (1978). *Psychometric Theory*. New York: McGraw-Hill.

Nunnally, J. C., & Berstein, I. H. (1994). *Psychometric Theory (3rd Ed.)*. New York: McGraw-Hill.

Robbins, A. S., Spence, J. T., & Clark, H. (1991). Psychological determinants of health and performance: The tangle web of desirable and undesirable characteristics. *Journal of Personality and Social Psychology*, 61, 755-765.

Robinson-Whelan, S, Kim, C., MacCallum, R. C., & Kiecolt-Glaser, J. K. (1997). Distinguishing optimism from pessimism in older adults: Is it more important to be optimistic or not to be pessimistic? *Journal of Personality and Social Psychology*, 73, 1345-1353.

Rodebaugh, T. L., Woods, C. M., & Heimberg, R. G. (2007). The reversal of social anxiety is not always the opposite: The reverse-scored items of the Social Interaction Anxiety Scale do not belong. *Behaviour Therapy*, 38, 192-206.

Rodebaugh, T. L., Woods, C. M., Thissen, D. M., Heimberg, R. G., Chambless, D. L., & Rapee, R. M. (2004). More information from fewer questions: The factor structure and item properties of the Original and Brief Fear of Negative Evaluation scale. *Psychological Assessment*, 16, 169-181.

Rorer, L. G. (1965). The great response style myth. *Psychological Bulletin*, 63, 129-156.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Rundquist, E. A. (1966). Item and response characteristics in a attitudinal and personality measurement: A reaction to L.G. Rorer's "The greatest response-style myth". *Psychological Bulletin*, 66, 166-177.

Russell, D. A., Peplan, L. A., & Ferguson, M. L. (1978). Developing a measure of lonliness. *Journal of Personality Assessment*, 42, 290-294.

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, 4, 219-247.

Scheier, M. F., & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy and Research*, 16, 210-228.

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem: A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, 67, 1063-1078.

Schmitt, N., & Schults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, 9, 367-373.

Schriesheim, C. A., & Eisenbach, R. J. (1995). An exploratory and confirmatory factor-analytic investigation of item wording effects on the obtained factor structure of survey questionnaire measures. *Journal of Management*, 21, 1177-1193.

Sharpe, P. A., Hickey, T., & Wolf, F. M. (1994). Adaptation of a general optimism scale for use with older women. *Psychological Reports*, 74, 931-937.

Smith, D. H. (1967). Correcting for social desirability response sets in opinion-attitude survey research. *The Public Opinion Quarterly*, 31, 87-94.

Smith, T. W., Pope, M. K., Rhodewalt, F., & Poulton, J. L. (1989). Optimism, neuroticism, coping, and symptom reports: An alternative interpretation of the Life Orientation Test. *Journal of Personality and Social Psychology*, 56, 640-648.

Steed, L. G. (2002). A psychometric comparison of four measures of hope and optimism. *Educational and Psychological Measurement*, 62, 466-482.

Sumi, K. (2004). The Japanese version of the revised Life Orientation Test: Reliability and construct validity. *Psychological Reports*, 95, 86-88.

Sumi, K. (2006). Correlations between optimism and social relationships. *Psychological Reports*, 99, 938-940.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington DC: American Psychological Association.

Tromp, D. B., & Brouha, X. D. R. (2005). Patient factors associated with delay in primary care among patients with head and neck carcinoma: A case-series analysis. *Family Practice*, 22, 554-559.

Vautier, S., & Raufaste, E. (2006). Configural instability of data from the Life Orientation Test-Revised. *Personality and Individual Differences*, 40, 1511-1518.

Wang, W. C., & Cunningham, E. G. (2005). Comparison of alternative estimation methods in confirmatory factor analyses of the general health questionnaire. *Psychological Reports*, *97*, 3-10.

Wegner, D. T., & Fabrigar, L. R. (2000). Analysis and design for non-experimental data: Addressing casual and non-casual hypotheses. In H. T. Reis & C.M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 412-450). New York: Cambridge University Press.

Whiteley, B. E. (2002). *Principles of research in behavioural science*. (7th ed.). New York, NY: McGraw Hill.

Xanthopoulou, D., Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2007). The role of personel resources in the job demands-resources model. *International Journal of Stress Management*, 14, 121-141.

Appendix A

Framing and Content of Original and Revised Versions of the LOT and LOT-R

Original LOT Framing/Content	Positively Framed Version Framing/Content	Negatively Framed Version Framing/Content
In uncertain times, I usually expect the best. +/+	In uncertain times, I usually expect the best. +/+	Even in uncertain times, I don't expect the worst. -/+
I always look on the bright side of things. +/+	I always look on the bright side of things. +/+	I never look on the dark side of things. -/+
I'm always optimistic about my future. +/+	I'm always optimistic about my future. +/+	I'm never pessimistic about my future. -/+
I'm a believer in the idea that "every cloud has a silver lining". +/+	I'm a believer in the idea that "every cloud has a silver lining". +/+	Used original +/+
If something can go wrong for me, it will. -/-	It somehow seems that if something can go right for me, it won't. +/-	If something can go wrong for me, it will. -/-
I hardly ever expect things to go my way. -/-	I almost always expect that things won't go my way. +/-	I hardly ever expect things to go my way. -/-
Things never work out the way I want them to. -/-	Things always work out the way I don't want them to. +/-	Things never work out the way I want them to. -/-
I rarely count on good things happening to me. -/-	I often count on bad things happening to me. +/-	I rarely count on good things happening to me. -/-
Overall, I expect more good things to happen to me than bad. +/+	Overall, I expect more good things to happen to me than bad. +/+	Overall, I don't expect more bad things to happen to me than good. -/+

Note: The -/+ symbols refer to the items framing and keying direction.

Appendix B

Dear Participant,

My name is Jamie Moore and I am currently undertaking a research project as part of the requirements of completing Honours in Psychology at Edith Cowan University. My research involves analysing certain properties of the Life Orientation Test, a brief self-report measure of how people perceive their environment and their experiences.

If you choose to participate in this study, you will be required to answer a few demographic questions and to complete a 13-item questionnaire. This should take less than 10 minutes of your time.

Your participation is entirely voluntary and completely anonymous. As you will not be required to provide any identifying information, completion of the questionnaire will indicate your consent. Participation or refusal to participate will not have any bearing on your current or future academic outcomes or receipt of university services. You can withdraw at any time without consequences by submitting an uncompleted or partially completed questionnaire.

The Ethics Committee of the ECU Faculty of Computing, Health and Science has approved this project, and there are no known risks associated with this project. If you would like to speak to someone independent of the research please contact Dr Justine Dandy of the School of Psychology and Social Science on 6304 5105 or via email at j.dandy@ecu.edu.au.

Your assistance is greatly appreciated.

Thank you

Jamie Moore
Principal Researcher:
Edith Cowan University
Email: jmoore4@student.ecu.edu.au

Supervisor: **Dr Ricks Allan**
School of Psychology, ECU
Ph; 6304 5048
Email: m.allan@ecu.edu.au

Appendix C

Negatively Framed Version

Please provide the following demographic information

1. Age in years

   _____

2. Gender (please circle)

   Male Female

3. Born in Australia

   Yes No

4. Cultural affiliation (please circle)

   Caucasian Indigenous Asian Other

For each of the following statements, state your feelings 'strongly agree', 'agree', 'neutral, 'disagree', or 'strongly disagree', by circling the corresponding number. There are no right or wrong answers, but try to be as accurate and honest as possible, without letting your answer to one-question influence answers to the others.

Even in uncertain times, I don't expect the worst

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

It's easy for me to relax

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

If something can go wrong for me it will.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I never look on the dark side of things.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I'm never pessimistic about my future.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I enjoy my friends a lot.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

It's important for me to keep busy.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I hardly ever expect things to go my way.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

Things never work out the way I want them to.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I don't get upset too easily.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I'm a believer in the idea that every cloud has a silver lining.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I rarely count on good things happening to me.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

Overall, I don't expect more bad things to happen to me than good.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

(Adapted from Scheier & Carver, 1985; Scheier, Carver, & Bridges, 1994, and Kubzansky, et al, 2004).

Appendix D

Positively Framed Version

Please provide the following demographic information

1. Age in years

   _____

2. Gender (please circle)

   Male Female

3. Born in Australia

   Yes No

4. Cultural affiliation (please circle)

   Caucasian Indigenous Asian Other

For each of the following statements, state your feelings 'strongly agree', 'agree', 'neutral, 'disagree', or 'strongly disagree', by circling the corresponding number. There are no right or wrong answers, but try to be as accurate and honest as possible, without letting your answer to one-question influence answers to the others.

In uncertain times, I usually expect the best.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

It's easy for me to relax

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

It seems that if something can go right for me it won't.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I always look on the bright side of things.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I am always optimistic about my future.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I enjoy friends a lot.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

It's important for me to keep busy.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I almost always expect that things won't go my way.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

Things always work out the way I don't want them to.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I don't get upset too easily.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I'm a believer in the idea that every cloud has a silver lining.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

I often count on bad things happening to me.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

Overall, I expect more good things to happen to me than bad.

Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
4	3	2	1	0

(Adapted from Scheier & Carver, 1985; Scheier, Carver, & Bridges, 1994, and Kubzansky, et al, 2004).

Guidelines for Contributions by Authors

Educational and Psychological Measurement

Executive Editor: X. Fan

Aim and Scope

Educational and Psychological Measurement discusses problems in the measurement of individual differences (including SEM, IRT, and "reliability generalization" studies), research on the development and use of tests and measurements (validity studies), testing programs (computer studies) being used for a variety of programs, and new and improved methods or items for treating test data. The journal also publishes statistics articles dealing with issues relevant to construct validity, broadly conceived.

Some of the significant topics covered include:

- Ways of thinking about describing score reliability
- Appropriate use of statistical significance scores and effect size measures
- Refusal to use stepwise methods either to select variables or to infer order of variable importance
- Suggested practices in conducting and reporting exploratory and confirmatory factor analyses

Manuscript Style

Type double-spaced using generous margins on all sides. The entire manuscript, including quotations, references, figure-caption list, and tables, should be double-spaced. Manuscript length, except under unusual circumstances, should generally be about 4000–9000 words. Empirical articles should include standard sections, such as Introduction, Methods, Results, and Discussion. Number all pages consecutively with Arabic numerals, with the title page being page 1. In order to facilitate masked (previously termed "double-blind") review, leave all identifying information off the manuscript, including the title page and the electronic file

name. Upon initial submission, the title page should include only the title of the article. An additional title page should be provided as a separate submission item and should include the title of the article, author's name, and author's affiliation. Academic affiliations of all authors should be included. The affiliation should comprise the department, institution (usually university or company), city, and state (or nation) and should be typed as a footnote to the author's name. This title page should also include the complete mailing address, telephone number, fax number, and e-mail address of the one author designated to review proofs. An abstract is to be provided, preferably no longer than 200 words.

References

List references alphabetically at the end of the paper and refer to them in the text by name and year in parentheses. The style and punctuation of the references should conform to strict APA style. In general, the journal follows the recommendations of the 2001 Publication Manual of the American Psychological Association (5th ed.), and it is suggested that contributors refer to this publication. For articles in a periodical, references in the citation list should include (in this order): last names and initials of all authors (for up to and including six authors), year published (in parentheses), title of article (roman type), name of publication (italics), volume number (italics), and inclusive pages (roman type).