Emotional and Psychological Wellbeing in Children:

The Standardisation of the Stirling Children's Wellbeing Scale

Dr Ian Liddle and Greg Carter   *Stirling Council Educational Psychology Service*

Acknowledgements

Summary

Introduction

The Stirling Children's Wellbeing Scale (SCWBS) was a project initiated by the Educational Psychology Service with the objective of creating a holistic, positively worded scale measuring emotional and psychological wellbeing in children aged of 8 to 15 years. The scale was developed and standardised by administering the questionnaire across 18 schools in the Stirling Council area with a population sample of 1 800 children.

Background

The Scottish Government's Curriculum for Excellence (2004) is designed to "enable young people to become…confident individuals…with a sense of physical, mental and emotional well-being... with resilience and self-reliance."

The need for a holistic positively worded scale measuring emotional and psychological wellbeing has risen directly out of the need to examine outcomes and assess progress in the health and wellbeing area of the Curriculum for Excellence. In addition, through Stirling Educational Psychology Service's work with *Friends for Life* project, a school based cognitive behavioural therapy programme, it became apparent that there was a need for a measure of wellbeing that encompassed both emotional aspects and psychological aspects in the school-age population.

In this context it was felt to be important to take a Positive Psychology perspective, i.e. to focus on positive aspects of wellbeing rather than a deficit based approach on mental health. Mental health, often interpreted as mental illness such as depression or anxiety, has been shown to be on a separate rather than continuous spectrum to that of a positive wellbeing. Mental health scales additionally are prone to ceiling effects and are not effective measures in showing changes or improvements in wellbeing.

The Educational Psychology Service, with agreement from Stirling Council and partly supported by a grant from the Scottish Government, has developed and standardised a new scale measuring both subjective and psychological aspects of wellbeing. This would enable education professionals to examine the effects that interventions or projects were having on children's wellbeing. The scale was developed to be a short and robust measure with positively worded items and with a range that allows progress to be measured.

Findings

The resulting 12-item scale met the statistical requirements for standardisation, proving to be a robust, reliable and a valid measurement for emotional and psychological wellbeing. Each item was rated on a Likert scale of 5 levels with the minimum score being 12 and the maximum score being 60. The scale covered areas of wellbeing including: optimism, cheerfulness and relaxation; satisfying Interpersonal relationships; clear thinking and competence. Overall this formed a single dimensioned scale with two sub-components described as *Positive Outlook* and *Positive Emotional State*. The mean score was 43.95 with a standard deviation of 7.29 and an inter-quartile range of 9. The scale should provide a useful tool for education professionals to assess any changes in wellbeing from a positive psychological perspective.

Further Research
The scale produced excellent standards of reliability and validity; however further research is required to determine the scale's sensitivity to change and discriminant validity. With ongoing research and the coalition of studies, this can be established and benchmark levels set. The scale is not intended to be used for diagnostic purposes but with ongoing research it may become useful as an initial indicator of poor mental health, which could then be assessed using established diagnostic measures.
 Further research is needed also into the components of wellbeing and their evolution as children become adolescents and young adults.

The SCWBS: Administering and Scoring
The SCWBS consists of 12 items measuring emotional and psychological wellbeing. The scale can be found in Appendix B and the key in Appendix C. The scale consists of two sub-components consisting of 6 items each relating to emotional and psychological wellbeing, namely Positive Emotional State and Positive Outlook. The scale additionally includes a social desirability sub-scale in order to determine whether any participants' scores have a response set – that is, there is a bias in the way they answer – or a predominance of socially desirable answers. The social desirability sub-scale consists of three items. All items on the scale are rated on a 5-point Likert-based scale.

The SCWBS was developed using both an electronic voting system, *PowerVote*, and the traditional paper and pencil method. The scale can be administered electronically or by paper and pencil methods. The scale performed optimally when using an electronic voting system, due to the increased accessibility to younger children and children with reading difficulties.
It is the authors' intention to make the scale available online.

The items on the scale are designed to measure any effects which projects and interventions are having on wellbeing, and as such relate to how the participants have been feeling and acting 'over the last couple of weeks'. The items relate to life both in school and outwith it. The minimum score for the SCWBS scale is 12 and the maximum 60 with a mean score of 44 having been recorded in this study. 50% of all scores fell between 39 and 48. For children scoring low on the scale, the option of further mental health assessment should be considered.   Any participants measuring 3 or 14-15 on the social desirability sub-scale are likely to be answering the questions with a response set or giving socially desirable (or undesirable) answers and their scores on the wellbeing scale should therefore be treated with caution. Any measure of effectiveness in projects and interventions can be analysed by using an appropriate statistical procedure such as a student's t-test (available in Excel) between the pre-intervention test scores and the post-intervention test scores.

Conclusion
The scale proved to be both reliable and valid, meeting the benchmark criteria set for standardising measures. The SCWBS consists of 12 items measuring emotional and psychological wellbeing and 3 items forming a social desirability sub-scale overall, providing a short and easy to administer scale. The scale is suitable for educational professionals looking to measure the effectiveness of projects and interventions with children aged from 8 to 15 years. The scale is founded on a positive psychological perspective measuring positive aspects of wellbeing as opposed to a deficit-based mental health model. With further research the scale's sensitivity and discriminant validity can be established; some diagnostic features may emerge. The scale is suitable for paper and pencil testing, although electronic administration provided increased accessibility for young children and children with reading difficulties.

Contents

Introduction

The Stirling Children's Wellbeing Scale (SCWBS) was initiated by the Stirling Educational Psychology Service with the objective of creating a holistic, positively worded scale measuring emotional and psychological wellbeing in children aged 8 to 15 years. The majority of current scales fail to do this due to a focus on mental illness rather than on wellbeing (van Dierendonck, 2003; Keyes, 2002; McDowell, 2009; Ryan & Deci, 2001; Springer & Hauser, 2006; Tennant, Hiller, Fishwick, Platt, Joseph, Weich, Parkinson, Secker & Stewart-Brown, 2007). Drawing on currently available theory to develop the scale, the resulting Stirling Children's Wellbeing Scale (SCWBS) should provide health and education professionals with a concise, robust measure of wellbeing.

Background

Historically the understanding of what constitutes wellbeing has been the subject of a very long and lively debate. The debate has essentially focussed on two predominant views, that of the hedonic and the eudaimonic perspectives (van Dierendonck, 2003; Keyes, 2002; McDowell, 2009; Ryan & Deci, 2001; Springer & Hauser, 2006; Tennant et al., 2007). Ryan & Deci (2001) described the two approaches as *hedonic* wellbeing, which is primarily concerned with the immediate states of pleasure and happiness, and *eudaimonic,* which is concerned with the actualization of human potentials. The modern encapsulation of hedonic wellbeing can be seen in what is often termed "subjective wellbeing" (SWB) (van Dierendonck, 2003; Keyes, 2002; McDowell, 2009; Ryan & Deci, 2001; Springer & Hauser, 2006; Tennant et al., 2007). SWB is seen to comprise life satisfaction, the presence of positive mood, and the absence of negative mood (Ryan & Deci, 2001). Psychological wellbeing (PWB) on the other hand, is based on the eudaimonic perspective and is seen as having the components of autonomy, personal growth, self-acceptance, life purpose, mastery, and positive relatedness (Ryff & Keyes, 1995). Although historically these have been opposing perspectives, there is a growing understanding that it is the combination of these two that constitutes a more complete understanding of a general psychological wellbeing (van Dierendonck, 2003; Keyes, 2002; McDowell, 2009; Ryan & Deci, 2001; Springer & Hauser, 2006; Tennant et al., 2007). This perspective is taken in the World Health Organisation's (WHO) definition of positive mental health, which is "a positive state of wellbeing, one which allows individuals to fully engage with others, cope with the stresses of life and realise their abilities" (2001). Positive mental health is often interchangeably used with psychological wellbeing to connote overall wellbeing. For the sake of clarity, in what follows the holistic view of wellbeing incorporating both SWB and PWB will simply be described as psychological wellbeing (PWB).

It is worth exploring the WHO's use of the phrase *positive* mental health further as it highlights a key factor in today's psychology. Exponents of Positive Psychology have argued that modern psychology is largely deficit-based (Seligman & Csikszentmihalyi, 2000) and that the majority of modern psychology is focussed on research into, and the prevention or curing of mental illnesses. They have argued that there is a need to understand and promote positive emotion, positive character and positive institutions (Seligman & Csikszentmihalyi, 2000). The confusion of mental health with mental illness has led to the need to differentiate it from that of a deficit-based understanding (van Dierendonck, 2003; Keyes, 2002; McDowell, 2009; Ryan & Deci, 2001; Springer & Hauser, 2006; Tennant et al., 2007) and to focus on wellbeing as a positive measure of healthy functioning (van Dierendonck, 2003; Keyes, 2002; McDowell, 2009; Ryan & Deci, 2001; Springer & Hauser, 2006; Tennant et al., 2007). With continuing research into wellbeing there has been some

evidence to support the notion that PWB is located on a separate dimension to mental illness rather than at the opposite end of a continuum (McDowell, 2009; Keyes, 2002). If mental illness and mental health are separate psychological phenomena then there is even a greater need to break out of a purely deficit-based model and to embrace an inclusion of positive psychology. This has been a major consideration in the development of the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) (Tennant, Hiller, Fishwick, Platt, Joseph, Weich, Parkinson, Secker, and Stewart-Brown (2007)) and the WHO (Five) Wellbeing Index (Bech, 2004), and the importance they placed on the need for positively worded questions in order to measure positive aspects of PWB. This needs also to be a key element in the development of a scale of PWB for children. It is a consideration that does however need to be weighed against the susceptibility of children in providing answers that are socially desirable or that fall victim to response set (Long, 1972; Merrell, 2007). The structure of the scale and the inclusion or exclusion of negatively worded items will be explored further in the design section of the report. This study does however aim to construct a scale that measures positive changes in children's wellbeing, and that is based on Positive Psychology principles.

Design of the SCWBS

In the design and standardisation of a scale there are a number of key requirements as laid out in Table 1 (Coolican, 2005). The majority of the requirements are statistical in nature and will be discussed in the method section of the report. The first requirement however is that of face validity and is relevant in the design of the scale. Face validity deals with whether the scale looks as though it is measuring what it is intending to measure. In order to achieve face validity there needs to be a clearly defined construct of what is being measured, and the scale itself must be recognisable and understandable to the participants as measuring that construct. This normally entails drawing upon current research and undertaking qualitative research in order to determine what that construct is and people's perception of that construct. In developing the SCWBS there was very little background research into what constituted PWB for children and how this may differ through the stages of development towards adulthood from a positive psychology perspective. The majority of what research was available focussed on adult models of wellbeing. This necessitated a reliance on literature on the wellbeing of adults and the adaptation of this into a scale suitable for children aged between 8 and 15 years. Although this was not ideal it provided a good starting point in developing a measure of wellbeing in children. In order to ensure that the items were suitable and relevant to children's wellbeing the items were piloted with individual children and assessed in light of their feedback. Items that children found difficult or had meanings that were open to interpretation were altered or omitted.   This gave the scale a good theoretical grounding, and ensured that it was understood by children and was perceived by children to be measuring wellbeing.

In looking at the structure for the key components of wellbeing in children the components of SWB and PWB were considered. Researchers developing the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) based the development of their scale largely on a meta-analysis carried out in 1983 by Kammann & Flett, where through regression analysis ten aspects of wellbeing were identified (Tennant et al. 2007). Tennant et al. have termed SWB and PWB as Mental Well-Being (MWB). The aspects identified were: confidence, usefulness, interest in life, problem solving, autonomy, positive relationships, thinking clearly and creatively, energy, happiness and optimism. Through their own analysis Tennant et al. (2007) derived the following structure: positive affect comprising of feelings of optimism, cheerfulness, and relaxation; satisfying interpersonal relationships; and positive functioning comprising energy, clear thinking, self acceptance, personal development, competence, and

autonomy. This presented itself as a neat breakdown of the components of a positive and holistic approach to MWB. There is an increasing debate however as to the dimensionality and factoring of components in wellbeing (Keyes, 2002; Lucas, Diener, & Suh, 1996; Ryan & Deci, 2001; Springer & Hauser, 2006; Van Dierendonck, 2003). Currently there are empirically supported arguments for a single dimensional construct; a 3 dimensional construct with sub-dimensions; and a construct comprising of 6 separate dimensions. The uncertainty over the construct of wellbeing can perhaps be put down to the interpretative nature of factor analysis; it may also be explained by the generality of what wellbeing is. This is not specific to wellbeing; a look at the history of personality constructs gives an indication of the problematic nature of defining components of a complex construct. Personality constructs have developed from Allport's 171 trait names to Eysenck's three personality supertraits (Maltby, Day, & Macaskill, 2007). A further complication in identifying the components of a construct is the definition of the construct itself. A loosely defined construct results in loosely structured components. Considering the lack of available theory on children's positive wellbeing, and the limitations of the study, the study relied on the structure as laid out in the development of the WEMWBS. The structure is based on sound empirical research and should be a reasonable basis for the development of a children's wellbeing scale, which can then be investigated through Principal Components Analysis for dimensionality and component structure.

Table 1
Measures for Assessing the Requirements for the Standardisation of the SCWBS

Criterion	Purpose of Criterion	Method for Meeting the Criterion
Face Validity	Ensure the scale is targeting the construct it is intending to	Theoretical foundation and interpretation of measure.
Construct Validity	Compare the scale to existing scale	Run a Pearson's Correlation between the scale and existing scales.
Internal Reliability	Assesses whether the scale consistently measures the same construct	Scale reduction through Item Analysis
		Assess the overall consistency using a Cronbach's Alpha
		Assess the overall consistency using a Pearson's Correlation between split halves of the scale
		Assess the dimensionality through Principal Components Analysis
External Reliability	Assess whether the scale is stable over time	Test-Retest using Pearson's correlation
Sensitivity to Change	Assess whether the scale is sensitive enough to measure changes.	Measure variance between known groups and effects of interventions using an analysis of variance.

(Coolican, 2005)

Overall 24 statements were constructed for the initial scale. The scale items themselves were based on each of the components and sub-components. The scale items consisted of a statement written in the first person that was designed to encapsulate that component in a language that is easily understood by a child with an average reading age of 8 years (Davis-Kean & Sandler, 2001). The statements were rated on a Likert scale with the first response being "at no time" and the last

response being "all of the time" (see Appendix A for Likert Scale ratings). A number of statements were created for each of the components to allow for the strongest items to be selected for the scale.

In the construction of the statements the question of whether to include negative items arose. As mentioned above, negative items have a tendency to measure aspects of mental illness as opposed to wellbeing. There is however an argument for including negative items in order to minimise socially desirable answers or response set. Social Desirability is observed when participants give answers that they think the researchers want to hear or that give them a desirable result. Response set occurs where participants simply rate all items with the same response. This can result in ratings that are not a true representation of the subjective state of the individual. Response set is more prevalent in younger children although studies have shown that it is only becomes a major detrimental factor for children aged below 8yrs (Chambers & Johnston, 2002). Negatively worded statements break this cycle and actively encourage the participant to engage in the measure. This posed a problem in the development of the scale. In order to ensure that the participants were actively engaging in the measure negatively worded items would be an asset, however negatively worded items may well be measuring a different construct. A potential way in which to have a negatively worded item that does not measure negative aspects is through the use of double negative statements. However, this is problematic because double negatives are especially difficult for children to process and may well not be suitable for children aged 8 years (Benson & Hocevar, 1985; Marsh, 1986). A possible solution to this is to create statements that measure negative aspects that are not potential measures of mental difficulties, that is, to target the lower end of the wellbeing continuum. Keyes' (2002), notion of "languishing", may have identified the nature of these aspects. Languishing is a state that lacks positive functioning and has an emphasis on the individual merely existing from day to day. An example of a negative statement fitting this context would be "I'm bored with the people around me."

An alternative method for reducing response set might be to include a social desirability sub-scale within the main scale. This would be a set of items that most people would not normally always agree to such as "I always tell the truth." This encourages a break in the response pattern and an active engagement with each item. Additionally it serves as a measure of socially desirable and response set responses. Participants who score high on the social desirability sub-scale can then have their overall responses assessed and either discarded or treated with some caution. In the construction of the scale it was decided to trial the inclusion of negative items, such as languishing, in order to fully engage the participants. Should the items prove to be unreliable or problematic, the alternate option of using a social desirability sub-scale would be considered.

Validity and reliability were tested through two separate studies; the first was an initial pilot study, which aimed to produce the initial scale and to assess the internal reliability, construct validity and dimensionality of the scale. The second study aimed to further assess the internal reliability of the scale while testing the external reliability and construct validity of the scale. The sensitivity of the scale could not be tested in this study as this involves running an intervention or discriminating between two known groups. The sensitivity of the scale will be assessed at a later point by undertaking research into the effectiveness of various interventions and programmes such as the Friends for Life programme, and using the SCWBS alongside other measures of wellbeing. This can be compared and collated with research undertaken by other parties to form an overall picture of the scale's sensitivity.

Initial Study

The aim of the initial study was to assess the construct of the scale and to identify which items were to be included in order to produce a short, robust and valid scale. The items were assessed for internal consistency through Item Analysis. Items that lowered the overall reliability of the scale would be discarded. The internal reliability of the scale was determined by using a factor analysis with a benchmark Cronbach's Alpha value set at greater than 0.8. The dimensionality and component structure was assessed using Principal Components Analysis with an Eigen value set to 1. Items that loaded onto components with a correlation of less than 0.4 were then discarded. An initial test for the construct validity of the scale was run using a Pearson's Correlation between the WEMWBS, the DuBois Self-Esteem Scale and the revised SCWBS with a benchmark greater than r = 0.7 being set. It was predicted that the correlation to the WEMBWS would be stronger than to the DuBois Self-Esteem scale. This is because the DuBois scale measures self-esteem as opposed to general wellbeing. The construct validity was only assessed for the Secondary School participants due to the WEMWBS only being suitable for children of 13 years and over. The construct validity was tested for all children in the second phase of the study. All analyses were carried out using SPSS.

Participants
For the pilot study the initial scale was administered to Primary 4 to Secondary 4 children across 12 schools in the Stirling Council area. Letters were written to the parents and pupils in the schools detailing the study and providing them with an opportunity to opt out of the study. The only personal details taken from the participants were their gender, age in years and keypad number to ensure anonymity. All the children were administered the scale unless they or their parents opted out. In the secondary school the participants were chosen by convenience of their respective class timetables. The schools ranged from rural to urban and from affluent to deprived catchment areas, in order to reflect a representative cross-section of the Stirling schools population (see Appendix A for the list of schools and sample sizes).

Procedure and Materials
The scale was administered to the participants using PowerVote. PowerVote is an electronic voting system where participants are able to cast their votes by using a remote keypad. For the study a PowerPoint presentation was written that incorporated the initial SCWBS. The presentation gave an overview of the study, instructions on using the keypads and two practice questions in order for the participants to familiarise themselves with the keypads and the Likert ratings. The practice questions were of lifestyle activities that related to the last two weeks but did not involve aspects of wellbeing. The practice statements comprised of a positively worded and a negatively worded statement. For the main part of the study each statement of the SCWBS was then projected onto a white screen and read out to the participants. The participants were again clearly told that the statements related to the past two weeks both in and outwith the school setting. The participants then rated their response to the statement using their keypads. Participants from the secondary school were additionally administered the DuBois Self-Esteem Scale using PowerVote and the WEMWBS in a paper and pencil format. The data was then entered into the SPSS database where the analysis was carried out.

Results

The scale was administered to 1162 participants, of whom 11 participants were shown to have response set/socially desirable answers by answering 5, 3, or 1 for all of the items regardless of whether they were negative or positively worded statements. The resulting sample size was therefore 1151. Of the 1151 participants 48% were female and 52% male. The breakdown of participants by age was 7 years = 7, 8 years = 179, 9 years = 238, 10 years = 243, 11 years = 264, 12 years = 45, 13 years = 58, 14 years = 45, 15 years = 69, and 16 years = 3. For a breakdown of ages by school see Appendix A. An initial inspection of the dispersion of scores for each item gave indications of ceiling effects in some statements. For example Statement 15 "I am happy to be me" showed a strong grouping for the answer of 5 "all of the time" (see figure 1). Where there were indications of ceiling effects the decision was taken to omit these statements from the final scale in order to maximise the sensitivity to positive changes and to ensure a more normal distribution. An inspection of the items on the scale showed that when broken down by age, there was a higher tendency for ceiling effects in responses from younger children. An ad-hoc Pearson correlation between age and overall score on the initial SCWBS showed a weak but significant negative correlation (r = - 0.060, N = 1151, p = 0.02) indicating that with the increase in age there was a decrease in the overall score. Where the items had a continuous ceiling effect across all ages they were omitted from the final scale and in total 7 items were omitted. Overall the distribution of scores showed a fairly normal distribution with negative skew (see Figure 2).

Figure 1.
Distribution of Scores for Question 15

Distribution of Scores for Question 15

Principal Components Analysis

A Principal Components analysis was run for the remaining 17 items. The scale showed acceptable goodness of fit with the Bartlett's test of sphericity at p < 0.001. An initial analysis of the Scree Plot indicated a uni-dimensional scale. A Varimax rotated component analysis was run indicating 3 sub-components (see Figure 3).

Inspection of the components indicated that the first component could be described as Positive Outlook with the second component being described as Positive Emotional State. The third component was a cluster of negatively worded items and upon closer inspection was most likely indicators of depression or anxiety. This indicated potential difficulties with the inclusion of the negative items and they were subsequently omitted from the scale. This resulted in a 12-item scale. Principal Components Analysis was run again, producing a single dimensional construct comprising two sub-components of Positive Emotional State and Positive Outlook. The sub-components each consisted of 6 items. All the items loaded onto the two components with values greater than 0.4 indicating good internal consistency.

Figure 2.
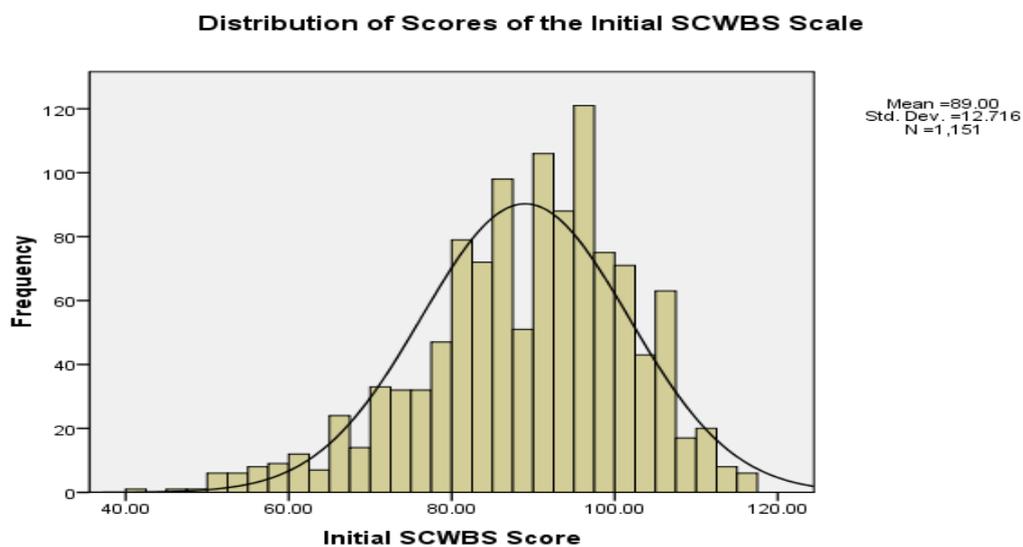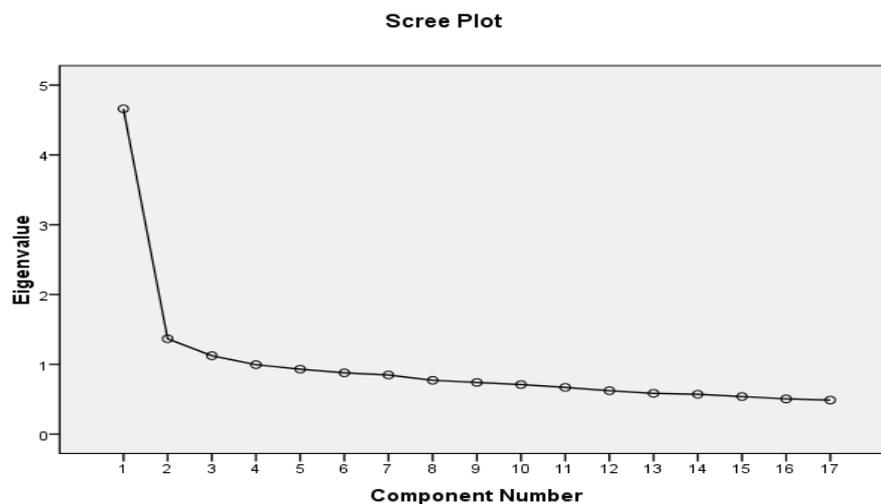*Distribution of Scores for the Initial SCWBS Scale*

Figure 3
Scree Plot of SCWBS Items

Item Analysis using Factor Analysis

Factor analysis was conducted to test the scale's internal reliability and to assess whether any items could be omitted to improve the overall reliability of the scale. The factor analysis showed that the scale had good internal reliability with a Cronbach's Alpha of 0.825, which was above the benchmark of 0.8. Item analysis indicated that removing any further items would weaken the internal reliability of the scale. Inter-item correlations did not give any indication of item redundancy. Overall the scale showed very reasonable internal reliability.

Construct Validity

Construct validity was tested by correlating the revised SCWBS with the DuBois Self-Esteem Scale and the WEMWBS using a Pearson's Bivariate correlation. The total number of participants for the SCWBS was N = 204, DuBois Self-Esteem Scale N = 176 and the WEMWBS N = 200. The discrepancy in the participant numbers was a result of children being withdrawn from class to attend other activities, keypads malfunctioning, and one participant deciding to opt out of the study. There was a strong significant positive correlation between the SCWBS and the WEMWBS ($r = 0.750$, N = 200, $p < 0.01$); a strong significant positive correlation between the SCWBS and the DuBois Self-Esteem Scale ($r = 0.647$, N = 176, $p < 0.01$); and a strong significant positive correlation between the WEMWBS and the DuBois Self-Esteem Scale ($r = 0.646$, N = 176, $p < 0.01$). This suggested good construct validity with a correlation above the $r = 0.7$ benchmark between the WEMWBS and the SCWBS and as predicted, a strong but weaker correlation between the SCWBS and the DuBois Self-Esteem Scale.

Discussion

Overall the resulting 12-item scale showed good internal reliability and good construct validity and was seen to be uni-dimensional with 2 sub-components. The 12-item scale was reduced down from the initial 24-item scale. The initial 24-item scale showed the desired bell shape distribution of scores but was positively weighted with a negative skew. This positive weighting falls in line with other adult scales measuring PWB such as the WEMWBS and the WHO (five). With the removal of items showing ceiling effects there was a reduction in the positive weighting of the scale. The mean score for the SCWBS was M = 44.14 with a standard deviation of SD = 7.55. A Pearson correlation was run to assess the Construct Validity of the scale and a significant positive correlation was found between the WEMWBS and the SCWBS above the $r = 0.7$ benchmark. A strong significant positive correlation of $r = .647$ was found between the SCWBS and the DuBois Self-Esteem scale indicating a strongly related construct. The lower correlation between the SCWBS and the DuBois Self-Esteem scale was expected as the DuBois scale measured a related aspect of wellbeing as opposed to an actual measure of wellbeing. Overall this showed the SCWBS to have good construct validity.

The initial Principal Components analysis showed the scale to be uni-dimensional with three sub-components, however one of these components was comprised of negatively worded items that could be collectively described as indications of depression or anxiety. It was decided to remove these negative items and to include a social desirability sub-scale in the final scale. In light of this the resulting scale consisted of two sub-components that could be described as Positive Emotional State and Positive Outlook. All remaining items loaded onto the two components of

the scale with values greater than 0.4 indicating robust internal reliability. This was confirmed using Factor Analysis where a Cronbach's Alpha of 0.823 was recorded, which was above the conservative threshold of 0.8. Item Analysis showed that there would be a decrease in the Cronbach's Alpha should any of the items be removed. The scale showed little evidence of redundancy of items with no excessively high inter-item correlations.

Positive Emotional State and Positive Outlook can easily be interpreted as Subjective (Hedonic) Wellbeing and Psychological (Eudaimonic) Wellbeing. This initially appears to represent a strong correspondence between an adult model and a children's model; however, from a closer inspection of the items that were omitted from the scale it appears that this interpretation may be problematic. The sub-components autonomy, personal development, energy and self-acceptance of the theorised component of Positive Functioning were all omitted from the end scale. An inspection of these components showed that the items measuring personal development, and those measuring energy and self-acceptance had very strong ceiling effects in the primary school participants. These ceiling effects were less severe for secondary participants, especially for the energy-related items. There may have been issues with the statements construction that was not picked up in the initial vetting. This can be seen in the typical response for the statement of self-acceptance "I am happy to be me" which was "who else am I going to be?" The items may require a level of abstract introspection that is not established at the younger ages. Self-acceptance may be largely based on external rather than internal factors such as the desire for qualities and possessions that others may have rather than an integral change in the make up of the self. It may also be that as children are forming their identities and sense of self, and working through the transition of attachment to their families and the formation of social ties, the qualities of autonomy, personal development and self acceptance are intrinsically different to those of adults. Certainly with the increase in age the sub-components became more clearly defined with fewer ceiling effects and a stronger loading onto the scale. It could be the case that there is a more general Positive Outlook in young children, which develops into more specific sub-components with age. This is clearly an area of wellbeing development in children that needs further research; for the purposes of this study in providing a general measure of wellbeing for educational professionals, a more general component of Positive Outlook will be accepted.

In the final scale it was decided to omit negatively worded statements and to include a social desirability sub-scale. The decision to exclude negative items was due to their forming a separate sub-component in the Principal Components analysis that had indications of symptoms of depression or anxiety such "I get tired easily" or "I get confused easily". The benefits of having negatively worded items could clearly be seen in the results where negative items scored 1 while positive items scored 5. In most cases after the first negative item, participants scored more evenly without a continuous pressing of a single key. For this reason the social desirability sub-scale was included to ensure that the participants were engaging in the items on the scale and to provide a helpful measure of response set /socially desirable answers. The social desirability sub-scale included 3 items such as "I always tell the truth" and "I always share my sweets". The understanding here is that most participants would not always tell the truth and would not always share their sweets and therefore a score of 15 on the sub-scale would be a clear indication of response set/socially desirable response. Scores for these items would of course not be counted towards the overall score on the SCWBS.

The resulting scale (see Appendix C) that was used in the secondary study consisted of 12 items covering the originally hypothesised components of Positive Affect

including optimism, cheerfulness and relaxation; Satisfying Interpersonal Relationships; and Positive Functioning including clear thinking and competence. These components loaded onto a single component with two sub-components best described as Positive Emotional State and Positive Outlook with 6 items loading onto each of the sub-components (see table 2). A 3-item stand-alone social desirability sub-scale was included in the overall SCWBS in order to assess response set/socially desirable responses and to encourage active engagement with the items. Overall the SCWBS looked to be a robust scale of acceptable internal reliability and good initial external validity that is easy for children and young people to complete.

Table 2.
Sub-Components and Items on the SCWBS

Wellbeing Sub-Component	SCWB Item
Positive Emotional State	I've been feeling calm
	I've been feeling cheerful about things
	I've been feeling relaxed
	I've been in a good mood
	I've been getting on well with people
	I enjoy what each new day brings
Positive Outlook	I think there are many things that I can be proud of.
	I feel that I am good at some things
	I think good things will happen in my life
	I can find lots of fun things to do
	I think lots of people care for me
	I've been able to make choices easily

Secondary Study

The second phase of the study set out to confirm the results from the initial study and to further test for external reliability. External reliability was assessed by conducting a test-retest analysis using a Pearson correlation with a benchmark correlation greater than $r = 0.7$. The re-test was administered 1 week after the initial test was undertaken. The secondary study also set out to assess whether there were any effects in using PowerVote as opposed to the traditional paper and pencil method. The construct validity of the SCWBS was further tested for the primary aged school children using the WHO (five) and the secondary children using the WHO (five) and the DuBois Self-Esteem scale.

Participants

For the secondary study the SCWBS including the social desirability sub-scale was administered to 701 children, again from Primary 4 to Secondary 4, across 6 schools within the Stirling Council area, 5 primary and 1 secondary. The participant details were kept to age and gender; however in order to ensure that the initial and re-test results were matched a record was kept by the class teachers of each child's participant number based on the keypad that they used. The participant numbers could then be matched while still keeping the names of the children anonymous to the researchers. All the primary school children between the years of P4 and P7 were administered the scale unless they had opted out. In the secondary school the participants were additionally chosen by convenience of their respective class timetables. The schools again ranged from rural to urban and from affluent to deprived catchment areas. (see Appendix B for the list of schools and sample sizes).

Procedure and Materials

All three scales (SCWBS, WHO (five), DuBois Self-Esteem Scale) were administered to the participants using PowerVote and followed the same procedure as in the initial study. Where the effects of using PowerVote were assessed the study followed a mixed design of within subjects and between groups. This consisted of two groups, the control group and the experimental group. All groups would be administered the SCWBS twice – an initial test and after a one-week gap, a retest. The control group solely used PowerVote for both the initial test and the retest. The experimental group initially completed the SCWBS using paper and pencil and then used PowerVote in the retest. A 2X2 mixed ANOVA analysis was carried out to assess if there were any main effects between the control and the PowerVote group. In all cases the class teacher kept a register of the children's names and keypad numbers. The researchers kept details of the school and class in order to match the keypad numbers to those kept by the class teachers for the retest. The data was entered into the SPSS database where the analysis was carried out.

Results

The scale was administered to 701 children of whom 13 were excluded due to response set/socially desirable responses determined by the scores on the Social Desirability scale and a visual inspection of the responses. The distribution of scores again showed a fairly normal distribution curve with some negative skewing. The mean score for the SCWBS was M = 43.51 with a standard deviation of SD = 6.66. There were no indications of ceiling effects amongst the distribution of scores for each of the items.

Through Principal Components Analysis the scale again proved to be uni-dimensional with two subcomponents of Positive Emotional State and Positive Outlook. The scale scored a strong Cronbach's alpha over the 0.8 benchmark. The scale showed good split half reliability with Cronbach alphas above 0.7 and a medium Pearson's correlation of r = 0.5 between the halves.

Construct validity was assessed by calculating Pearson Correlations with the WHO (five) and the DuBois Self-Esteem Scale. The SCWBS had a strong positive correlation with the WHO (five) above r = 0.7 and a strong positive correlation with the DuBois Self-Esteem scale (see Table 3). This confirmed the findings of the initial study, of good external validity.

Table 3.
Pearson Correlations for Construct Validity

|		SCWBS	WHO (five)	DuBois Self-Esteem Scale
SCWBS	Pearson Correlation	1.000	.742	.692
	Sig. (2-tailed)		.001	.001
	N	492	467	274
WHO (five)	Pearson Correlation		1.000	569
	Sig. (2-tailed)		.001	.001
	N		562	274
DuBois Self-Esteem Scale	Pearson Correlation			1.000
	Sig. (2-tailed)			.001
	N			274

External Reliability

The external reliability of the scale was tested using the test /re-test method. A Pearson correlation was run between the initial test scores and the re-test scores taken a week later. The analysis showed a strong correlation between the initial scores and the retest scores ($r = 0.752$, $N = 232$, $p < .01$) showing that the scale had good external reliability.

The Effects of Using PowerVote

An initial inspection of the mean scores showed that in the control group the initial test scores using PowerVote (M = 44.57; SD = 6.86) were higher than that of the retest scores using PowerVote (M = 40.88; SD = 6.99) while the initial paper test scores in the experimental group (M = 45.24; SD = 7.03) were only marginally higher than the PowerVote re-test scores (M = 44.70; SD = 9.09). The initial indications were that there was a main interaction between the methods of administering the SCWBS (see Figure 4).

Figure 4
Estimated Marginal Means of SCWBS
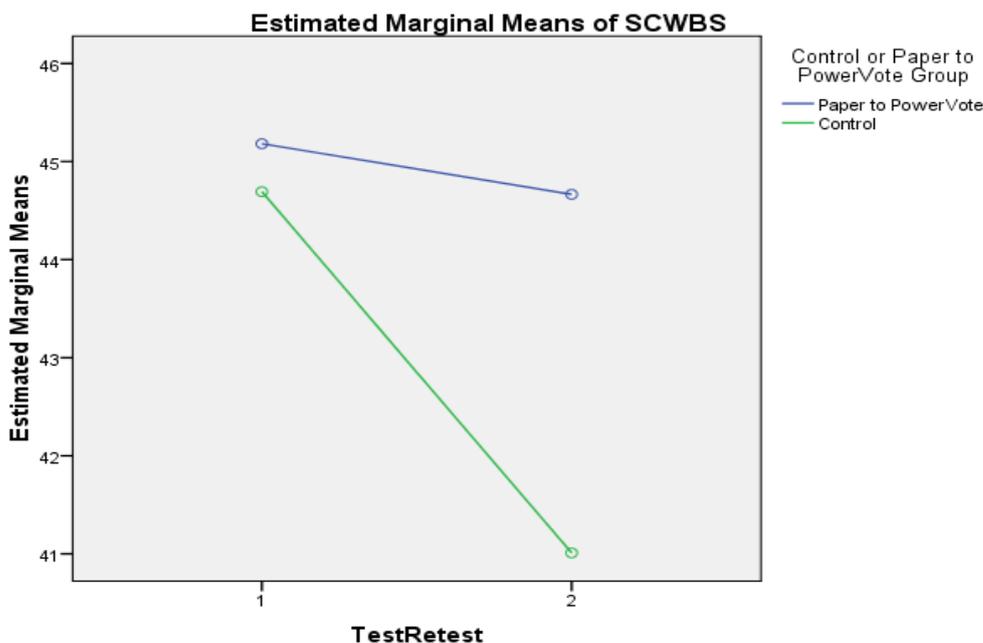
The mixed ANOVA was run resulting in a main interaction within the initial test and re-test between the control and experimental group: $F (1,284) = 15.016$, $p < 0.01$, $\eta^2 = 0.05$. Although the main interaction was of high significance the effect size was negligible ($\eta^2 = 0.05$) indicating that the use of PowerVote had a very small effect on participant scores.

The internal reliability results for the paper test scores recorded a Cronbach's alpha of 0.795 which is above the 0.7 benchmark and just shy of the conservative benchmark of 0.8. Interestingly Principal Component analysis showed a third sub-component that could be described as satisfying personal relationships. The test re-test Pearson correlation gave a strong positive correlation which was however below the 0.7 benchmark ($r = 0.603$, $N = 186$, $P < .01$). There may have been a confounding variable in that a number of the registers matching participants' names

to keypads were misplaced and the researchers had to rely on the participants having to remember their keypad numbers. This may well have reduced the overall correlation due to mismatched participant identifiers.

Discussion

The analysis of the secondary study confirmed the results found in the initial study. The resulting scale (see Appendix C) had good internal reliability (Cronbach's Alpha exceeding the 0.8 benchmark) and good external validity (Pearson's correlation exceeding the $r = 0.7$ benchmark). The scale was further found to have good external reliability with a strong positive Pearson correlation above the benchmark of $r = 0.7$. This showed the scale as having good stability and reliability over time. Overall the scale was robust with good internal and external reliability and validity. The mean score for the scale was $M = 43.95$ with a standard deviation of $SD = 7.29$ and an interquartile range of 9. There was a good distribution of scores producing a bell shaped curve: however the scale was positively weighted with some negative skewing. This can be expected as the majority of the population has a good positive wellbeing with a small percentage having a low sense of wellbeing. There are indications of a clustering of scores around and below the 30 mark which may be indicative of poor mental health. Compared to the distribution on the other scales and in other studies, this positive weighting and negative skewing is not unusual. The distribution showed little in the way of ceiling effects and has good potential to measure any changes in overall wellbeing.

Sensitivity to change is an area that unfortunately could not be assessed in this study; however indications were good that the scale will be sensitive and have a discriminant quality. This can be seen in the dispersion of scores across age groups and the clustering of scores at the 30 value. Scores showed a decline in wellbeing with age, which has an intuitive logic considering the concerns of approaching adulthood, school expectations, transitional periods and pressures for performance facing teenagers. There could well be another consideration in that the decrease in wellbeing scores with age may simply reflect an increase in comprehension of the statements. It could be conjectured that the clustering of scores at the bottom end of the scale may be indicating a group of participants who suffer from poor mental health again showing good signs of discriminant validity. A possible alternate explanation however is that it can be a group of contrary school children giving low ratings for fun. The sensitivity of the SCWBS would best be tested through the assessments of various interventions targeting children's wellbeing and additionally seeing if it can discriminate between known groups of differing wellbeing levels. Over a period of time and with a consolidation of research the SCWBS could well have a diagnostic role in identifying children that potentially have mental health difficulties, however this is not the intention of the scale. Children scoring low on the SCWBS should have their results treated with some caution. It would be advisable to undertake further assessments for mental health issues and to discuss the results with those children.

Administering and Scoring the SCWBS

The SCWBS consists of 12 items measuring emotional wellbeing and psychological wellbeing. The scale can be found in Appendix C and the key in Appendix D. The scale consists of two sub-components consisting of 6 items each relating to emotional and psychological wellbeing, namely Positive Emotional State and Positive Outlook. The scale additionally includes a social desirability sub-scale in order to determine whether any participant's scores have response set or predominantly

socially desirable answers. The social desirability sub-scale consists of three items. All items on the scale are rated on a 5-point Likert-based scale.

The SCWBS was developed using both an electronic voting system, *PowerVote*, and the traditional paper and pencil method. The scale can be administered electronically or by paper and pencil methods. The scale performed better when using an electronic voting system due to the increased accessibility to younger children and children with reading difficulties. It is the authors' intention to make the scale available online to allow individual electronic access with auditory support. The items on the scale are designed to measure any effects projects and interventions are having on wellbeing, and as such relate to how the participants have been feeling and acting 'over the last couple of weeks'. The items relate to life both in school and outwith it. The minimum score for the SCWBS scale is 12 and the maximum 60 with a mean score of 44 having been recorded in this study. 50% of all scores fell between 39 and 48. It is recommended that for children scoring low on the scale that the results be explored further by chatting with the children and potentially administering further assessments for mental health difficulties. Any participants measuring 3 or 14-15 on the social desirability sub-scale are likely to be answering the questions with a response set or giving socially desirable (or undesirable) answers, and their scores on the wellbeing scale should therefore be treated with caution. Any measure of effectiveness in projects and interventions can be analysed by using a student's t-test between the pre-intervention test scores and the post-intervention test scores.

Conclusion

The scale was founded on a positive psychological perspective measuring positive aspects of wellbeing as opposed to a deficit-based mental health model. The scale proved to be a reliable and valid measure of wellbeing meeting the benchmark criteria set out for standardising measures. The resulting SCWBS consisted of 12 items measuring emotional and psychological wellbeing and 3 items forming a social desirability sub-scale.   With further research the scale's sensitivity and discriminant validity can be established, and from which some diagnostic features may emerge. The scale is suitable for paper and pencil testing, though it is recommended that it be used electronically to increase accessibility for young children and children with reading difficulties.   Overall the scale should provide educational professionals with a concise and robust measure of wellbeing that can assess the effectiveness of projects and interventions for children aged from 8 to 15 years.

Appendices

Appendix A

Schools and Sample Sizes for the Pilot Study

School	Sample Size	Sample Size by Age
Balfron HS, Balfron	N = 204	≤ 12 = 30
		13 = 58
		14 = 45
		≥ 15 = 71
Our Lady's PS Raploch	N = 25	≤ 8 = 8
		9 = 3
		10 = 9
		≥ 11 = 5
Borestone PS, Borestone	N = 74	≤ 8 = 11
		9 = 18
		10 = 18
		≥ 11 = 27
Kincardine in Menteith PS, Kincardine	N = 26	≤ 8 = 4
		9 = 10
		10 = 5
		≥ 11 = 7
Thornhill PS, Thornhill	N = 37	≤ 8 = 9
		9 = 12
		10 = 7
		≥ 11 = 9
Deanston PS, Deanston	N = 11	≤ 8 = 1
		9 = 5
		10 = 3
		≥ 11 = 2
St Ninians PS, Stirling	N = 187	≤ 8 = 41
		9 = 44
		10 = 45
		≥ 11 = 53
St Mary's Bannockburn PS, Stirling	N = 129	≤ 8 = 12
		9 = 28
		10 = 44
		≥ 11 = 45
East Plean PS Plean	N = 71	≤ 8 = 18
		9 = 20
		10 = 18
		≥ 11 = 15
Bridge of Allan PS, Bridge of Allan	N = 190	≤ 8 = 36
		9 = 51
		10 = 38
		≥ 11 = 65
Buchlyvie PS, Buchlyvie	N = 22	≤ 8 = 6
		9 = 6
		10 = 8
		≥ 11 = 2
Newton PS, Dunblane	N = 175	≤ 8 = 40
		9 = 41
		10 = 48
		≥11 = 46

Appendix B

Schools and Sample Sizes for the Secondary Study

School	Sample Size	Sample Size by Age
McLaren HS, Callander	N =274	$\leq 12 = 86$
		13 = 73
		14 = 71
		$\geq 15 = 44$
St Mary's PS, Dunblane	N = 38	$\leq\ \ 8 = 11$
		9 = 9
		10 = 9
		$\geq 11 = 9$
Fallin PS, Fallin	N = 124	$\leq\ \ 8 = 22$
		9 = 29
		10 = 31
		$\geq 11 = 42$
Doune PS, Doune	N = 70	$\leq\ \ 8 = 21$
		9 = 12
		10 = 14
		$\geq 11 = 23$
Strathyre PS, Strathyre	N = 22	$\leq\ \ 8 = 4$
		9 = 5
		10 = 5
		$\geq 11 = 8$
Braehead PS, Stirling	N = 88	$\leq\ \ 8 = 20$
		9 = 21
		10 = 25
		$\geq 11 = 22$

Appendix C

The Stirling Children's Wellbeing Scale

Here are some statements or descriptions about how you might have been feeling or thinking about things over the past couple of weeks.

For each one please put a tick in the box which best describes your thoughts and feelings; there are not right or wrong answers.

	Statements	Never	Not much of the time	Some of the time	Quite a lot of the time	All of the time
1	I think good things will happen in my life	1	2	3	4	5
2	I have always told the truth	1	2	3	4	5
3	I've been able to make choices easily	1	2	3	4	5
4	I can find lots of fun things to do	1	2	3	4	5
5	I feel that I am good at some things	1	2	3	4	5
6	I think lots of people care about me	1	2	3	4	5
7	I like everyone I have met	1	2	3	4	5
8	I think there are many things I can be proud of	1	2	3	4	5
9	I've been feeling calm	1	2	3	4	5
10	I've been in a good mood	1	2	3	4	5
11	I enjoy what each new day brings	1	2	3	4	5
12	I've been getting on well with people	1	2	3	4	5
13	I always share my sweets	1	2	3	4	5
14	I've been cheerful about things	1	2	3	4	5
15	I've been feeling relaxed	1	2	3	4	5

Appendix D

SCWBS Key

Wellbeing Sub-components and Related Items

Wellbeing Sub-Component	Item	Related Item on the SCWBS
Positive Emotional State	9	I've been feeling calm
	14	I've been feeling cheerful about things
	15	I've been feeling relaxed
	10	I've been in a good mood
	12	I've been getting on well with people
	11	I enjoy what each new day brings
Positive Outlook	8	I think there are many things that I can be proud of.
	5	I feel that I am good at some things
	1	I think good things will happen in my life
	4	I can find lots of fun things to do
	6	I think lots of people care for me
	3	I've been able to make choices easily

Each item is scored 1 to 5.
The minimum for the scale is 12 and the maximum 60.
Currently the mean average score is 44 with 50% of all scores within the range of 39 and 48.

Social Desirability Sub-Scale

	Item	Related Item on the SCWBS
	2	I have always told the truth
	7	I like everyone I have met
	13	I always share my sweets

Each Item is scored 1 to 5.
Overall scores of 3 or 14/15 on this sub-scale would indicate that the participant's wellbeing scores should be treated with caution.

References

Bech, P. (2004). Measuring the dimensions of psychological general well-being by the WHO-5. *QoL Newsletter* 2004; 32: 15-16.

Benson, J., & Hocevar, D. (1985). The impact of item phrasing on the validity of attitude scale for elementary school children. *Journal of Educational Measurement, 22,* 231-240.

Chambers, C., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Paediatric Psychology, 27*, 27 – 36.

Coolican, H. (2009). *Research Methods and Statistics in Psychology (5th Edition).* UK: Hodder Education.

Davis-Kean, P. E., & Howard, S. M. (2001). A Meta-analysis of measures of self-esteem for young children: A Framework for future measures. *Child Development, 72*, 887 – 906.

Diener, E., & Suh, M. E. (1997). Subjective well-being and age: An International analysis. *Annual Review of Gerontology and Geriatrics, 17, 304* - 324.

van Dierendonck, D. (2003). The construct validity of Ryff's Scales of Psychological Well-being and its extension with spiritual well-being. *Personality and Individual Differences, 36*, 629 – 643.

Keyes, C. L. M. (2002). The mental health continuum: From languishing to flourishing in life. *Journal of Health and Behavior Research, 43*, 207 – 222.

Long, B. H. (1972). A Second plea for position response set controls. *American Educational Research Journal, 9*, 326 – 329.

Lucas, R. E., Diener, E., & Suh. (1996). Discriminant validity of well-being measures. *Journal of Personality and Social Psychology, 71*, 616 – 628.

Maltby, J., Day, L., & Macaskill, A. (Eds.). (2007). *Personality, Individual Differences and Intelligence.* Bilbao, Spain: Pearson Education Ltd.

Marsh, H. W. (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-development-mental phenomenon. *Developmental Psychology*, 22, 37 – 49.

McDowell, I. (2009). Measures of self-perceived well-being. *Journal of Psychosomatic Research,* Article in press.

Merrell, K. W. (2007). *Behavioral, Social, and Emotional Assessment of Children and Adolescents (3rd Edition).* US: Lawrence Erlbaum Associates Inc.

Ryan, R. M., & Deci, E. L. (2001). On Happiness and human potentials: A Review of research on Hedonic and Eudaimonic Well-Being. *Annual Review of Psychology, 52*, 141 – 166.

Ryff, C. D., & C. L. Keyes. (1995). The Structure of Psychological Well-Being revisited. *Journal of Personality and Social Psychology, 69*, 719 - 727.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive Psychology: An Introduction. *American Psychologist, 55*, 5 – 14.

Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive Psychology progress: Empirical validation of interventions. *American Psychologist, 60*, 410 – 421.

Springer, K. W., and Hauser, R. M. (2006). An Assessment of the construct validity of Ryff's Scales of Psychological Well-being: Method, mode, and measurement effects. *Social Science Research, 35*, 1080 – 1102.

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warkwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes,* 5:63.