

## The Psychometric Properties of the GLWS

When evaluating a questionnaire or survey tool for use within organisations, it is typical for potential users to investigate the psychometric properties of the tool to establish that it has been well constructed and is a robust measurement of the thing it purports to measure. The features in question are usually the reliability and validity of the measurement tool.

This document has been written to provide an insight into the GLWS in relation to its reliability and validity, as well as answer the question of norming.

### RELIABILITY

Reliability is the overall consistency of a measure. A measure is said to have a high reliability if it “produces similar results under consistent conditions”. The reliability of a questionnaire scale is normally expressed as the correlation between two or more sets of scores on the same scale for the same group of individuals.

There are three main types of reliability relevant to tests/measures:

1. *Test – Retest Reliability* assesses the degree to which test scores are consistent from one test administration to the next. Measurements are gathered from a single rater who uses the same methods or instruments and the same testing conditions.
2. *Internal Consistency Reliability* assesses the “consistency of results across items within a test”. It is based on the correlations between different items on the same test, or the same subscale in a larger/longer test. It measures whether several items that propose to measure the same general construct produce similar scores. Internal consistency is usually measured with Cronbach’s alpha, a statistic calculated from the pair wise correlations between items. Scale reliabilities higher than 0.90 suggest some item redundancy, with too many items being basically the same question in a slightly different guise. Scale reliabilities falling below 0.70 suggest some mixed item content indicative of possible multi-dimensionality. Best practice test construction is therefore to achieve an ideal range of between 0.70 to 0.90 in order to claim the test has sound internal consistency.
3. *Alternate/Parallel Forms Reliability* is a measure of reliability obtained by administering different versions of an assessment tool or test to the same group of individuals (both versions must contain items that probe the same construct, skill, knowledge base etc). The scores from the two versions can then be correlated to evaluate the consistency of results across alternative versions.

## **The GLWS & Test-Retest**

The GLWS measures wellbeing, and follows the premise that wellbeing is a state of being that is dynamic, rather than fixed: this suggests that being well is about how you feel rather than how you live and that this is expected to change depending on circumstances and events and your response to these.

Hence for GLWS, wellbeing is defined as a “delicate balancing act between an individual’s social, emotional, psychological and physical assets (resources) and the particular social, emotional, psychological and physical liabilities (challenges) they are facing in life and work at any one time. When individuals have more challenges than resources, their seesaw dips, along with their wellbeing, and vice-versa” (Dodge, Daly, Huyton, & Sanders, 2012).

It is hoped that by completing the GLWS, respondents will learn about what enhances and what detracts from their wellbeing and that this knowledge will assist them in making any necessary changes to their thoughts, feelings and actions in order to experience a higher state of wellbeing more of the time (i.e. achieve a good balance between their challenges and resources). We are aiming that their wellbeing will change as a result of the GLWS experience, and for some respondents, changes may even occur during the completion process. This being the case, we would not expect to see a high degree of consistency between one administration of GLWS and the next, making Test – Retest reliability a redundant concept for the GLWS.

A more relevant and interesting use of the Test-Retest approach would be a measure of the extent to which respondents have in fact made positive changes in the aspects of their GLWS profiles that were represented as red flags (detractors of wellbeing). We have positive anecdotal evidence about change at an individual level and are looking for further opportunities to extend this investigation to a larger group of GLWS respondents.

## **The GLWS & Internal Consistency**

In early 2018 we conducted this analysis on a sample of just over 1000 GLWS respondents and achieved Cronbach’s alphas of between 0.74 and 0.88 for all domains – a psychometrically sound result, in line with best practice.

Individual Domain/Scale internal consistencies are shown in the table below.

Working Well – 11 items per Domain		Living Well – 11 items per Domain	
Authentic Relationships	.812	Authentic Relationships	.828
Meaning, Purpose & Direction	.794	Meaning, Purpose & Direction	.792
Resilience & Equanimity	.794	Resilience & Equanimity	.840
Vitality & Energy	.735	Vitality & Energy	.798
Balance & Boundaries	.788	Balance & Boundaries	.875
Intellectual Engagement	.803		

### The GLWS & Alternate/Parallel Forms

This is clearly not a practical means of establishing reliability for a tool such as the GLWS as no alternate form exists and it is not practical to create one. This method of examining reliability is of most use with aptitude tests where it is easier and more desirable to have a variety of versions of the same test.

### VALIDITY

Validity refers to how well a test measures what is it purported to measure. Hence a wellbeing survey designed to measure wellbeing and used to measure wellbeing, should be measuring wellbeing!

There are four main types of validity relevant to tests/measures:

1. *Face Validity* is concerned with whether the test appears to measure what it was designed to measure. Although this kind of validity has no technical or statistical basis, its vitally important if the test is to be accepted by our senior leader target audience who are renowned for having a critical or even cynical disposition.
2. *Content Validity* ensures that the test items cover the broad range of areas within the concept being measured.
3. *Construct Validity* is used to ensure that the test is actually measuring what it is intended to measure (i.e. the construct) and not other variables. Typically, this means that the test measures some theoretical construct such as wellbeing. Building up a picture of the construct validity of a measure involves any information which throws some light on the nature of the construct under investigation.

For example, correlations with other scales/questionnaires will provide useful information on a test’s construct validity. Factor analysis is often used to investigate the construct validity of measures with multiple factors.

- 4. *Predictive Validity* is the extent to which a test predicts some future outcome or ‘criterion’. It correlates test results with another criterion of interest.

**The GLWS & Face Validity**

GLWS has strong face validity as demonstrated by blind data from an adjective checklist collected from 106 respondents during an early trial, showing:

Thought provoking	76.4%	Uplifting	3.8%
Interesting	48.1%	Fun	2.8%
Beneficial	31.1%	Repetitive	2.8%
Helpful	28.3%	Significant	2.8%
Meaningful	21.7%	Distressing	1.9%
Enlightening	20.8%	Boring	0.9%
Enjoyable	15.7%	Annoying	0%
Too Long	14.2%	De-motivating	0%
Motivating	11.3%	Frustrating	0%
Stimulating	10.4%	Irritating	0%
Neutral	6.6%	Anxiety-provoking	0%
Energising	5.7%		
Depressing	4.7%		

The five most frequently cited adjectives upon completion of the survey were:

1. Thought Provoking
2. Interesting
3. Beneficial
4. Helpful
5. Meaningful

In its original form, 14.2% of the respondents described the survey as ‘Too Long’ but since this time, the 150-item trial version has been whittled down by approx. 20% to its current number of only 121 questions.

There is a pleasing lack of endorsement for the negative adjectives despite being actively and anonymously sought.

Similar research on a smaller sample of 37 managers from one large Australian corporate showed similarly positive reactions:

Thought provoking	67.57%	Anxiety Provoking	5.41%
Interesting	54.05%	Irrelevant	2.70%
Constructive	54.05%	Frustrating	0.00%
Beneficial	48.65%	Irritating	0.00%
Motivating	32.43%	De-motivating	0.00%
Enjoyable	32.43%		
Meaningful	29.73%		

The insight from this research is that respondents find their experiences of the survey as being constructive and valuable.

### **The GLWS & Content Validity**

GLWS has been designed with strong content validity, drawing on a thorough range of existing definitions of wellness and wellbeing to inform the GLWS Framework and a wide body of research on what factors impact wellbeing in the workplace and beyond. All items in the GLWS were selected on basis of relevance to wellbeing either from research evidence, published theories and/or clinical experience.

### **The GLWS & Construct Validity**

Construct validity for the GLWS will be examined in a variety of ways and in more depth when there are large enough sample sizes to achieve meaningful, statistically significant results.

Having established the internal consistencies are satisfactory, we have conducted 3 studies examining the relationships between and within the GLWS domains, and within and across the Working Well and Living Well sections.

This first set of data is summarised in a paper presented to the International Congress of Psychology in July 2016 and is available for download from:

<https://www.glswellbeing.com/wp-content/uploads/2018/02/McGibbon-Kurz-Gillespie-2016-ICP-Yokohama-MEWS-Presentation.pdf>.

The second set of data is based on a much bigger sample of 1000+ respondents to the 2016/17 release of the survey and repeats the above analysis. This data is currently being written into a paper which will be available for download at:

<https://www.glswellbeing.com/research/when-completed>.

The third set of construct related data examining the link between wellbeing and the Hogan Assessment System scales was presented at the Australian Psychological Society's 12th Industrial and Organisational Psychology Conference on 13-15th July 2017 and is summarised at:

<https://www.glswellbeing.com/wp-content/uploads/2018/02/Exploring-the-link-between-subjective-wellbeing-and-resilience-.pdf>

and at:

<https://www.glswellbeing.com/wp-content/uploads/2018/02/Exploring-the-link-between-dark-side-derailers-and-subjective-wellbeing.pdf>.

Finally, ongoing research correlating the GLWS with other tools will be undertaken as the opportunity presents itself (for example, with the Hogan's Assessment Suite, Happiness Scales, Life Satisfaction Scales, other wellbeing measures).

### **The GLWS & Predictive Validity**

In the case of a wellbeing measure, it is plausible that an organisation might want to use the results to predict those individuals who will prove more resilient to pressure, be less prone to stress related ill health and absence and/or more likely to stay in the organisation. It could be hypothesised that this data would be used to provide additional support for those individuals whose wellbeing profile was predicting a less 'well' future, however the ethics involved in this hypothetical scenario are complex and not attractive.

The GLWS has been designed and is used for developmental purposes alone – under the Terms & Conditions of usage, GLWS is not to be used for selection, promotion or any other decision an organisation makes about its senior people. Indeed, the individual GLWS data is not accessible by the organisation under any circumstances.

The GLWS is designed to enable conversations about wellbeing, with the goal of building an individual's understanding about their own drivers and detractors of wellbeing and enabling them to take actions to maximise their wellbeing throughout their life.

With this clear vision for the GLWS, establishing any form of predictive capacity for the GLWS is not a priority. However, we do keenly embrace the already well-established link between wellbeing and sustainable high performance, and we will be alert to opportunities for the GLWS to contribute further evidence in this regard i.e. to show that individuals with strongly positive GLWS results will also be strong and sustainable high performers in their professional lives, correlating GLWS data with performance outcomes.

## **NORMING**

EEK & SENSE has made a deliberate decision NOT to norm the GLWS, i.e. we don't provide insight for respondents on how their responses compare with other executives. This is for the following three reasons:

1. Wellbeing is a subjective and dynamic aspect of being. How "well" we feel is dependent on our interpretation of the events/life we experience and the internal and external resources that we each bring to bear to respond to challenges along the way.
2. Understanding what impacts one's wellbeing at work and out of work is key to ensuring optimal and sustainable performance and to living the best life we can. Understanding this in the context of other people's wellbeing seems irrelevant and potentially distracting.
3. We could norm the GLWS and present how most senior executives tend to respond to each of the items, however, there is a risk that even if MOST people respond this way, it doesn't make it the optimal way to behave! For example, we know from solid research that wellbeing across numerous domains is enhanced by regularly getting 7-9 hours of sleep per night. Our 'norms' could show that most GLWS respondents report rarely or never getting this much sleep. This could provide encouragement that 'it must be ok if everyone else is the same' as oppose to using the evidence base to change maladaptive behaviours.

As a more meaningful alternative to norming the GLWS, we have determined numerical cut-offs using a red/green/amber flag system – based on prior evidence, theory or clinical experience that more or less of particular factors are associated with greater or lesser wellbeing – to represent if a respondent’s mean score on both a domain and individual item basis is likely to be strongly enhancing through to strongly detracting from wellbeing. We also provide a detailed indication on each of the domains of wellbeing showing where they are currently experiencing more enhancers, more detractors or a combination of both.

Having said all that, with the GLWS Group Report option we have the capacity to create norms and can undertake an organisational or team analysis showing the norm for the organisation or team, should this prove helpful in informing future activities/initiatives.