Hōmai te Waiora ki Ahau: A tool for the measurement of wellbeing among Māori – the evidence of construct validity

Stephanie Palmer
Massey University, Wellington

This paper describes some initial steps in the development of Hōmai te Waiora ki Ahau, a tool for the measurement of wellbeing among Māori. Pictures were used to present the twelve items and participants were asked to rate the extent to which each item was a source of waiora, or wellbeing, in their lives at present. Self-ratings for the twelve items were summed to give an overall waiora score. Thirty-one women from Hauraki, aged 16-34 years, took part in the pilot during the 3rd trimester of pregnancy. The tool has also been administered to two smaller groups. Collectively, the data obtained from these small studies provide some evidence of construct validity. The tool is clearly workable, respondents are willing to take part in the process and participants with little or no fluency in te reo Māori are able to think about their feelings of waiora within a Māori conceptual paradigm. Future directions and strategies to improve the psychometric properties of this instrument are discussed.

In 1947 the World Health Organization proclaimed that good health was not only the absence of disease but also the presence of positive health and wellbeing. Two decades later, Bradburn (1969) put forward the idea that wellbeing could be measured by the predominance of good feelings over bad. His Affect Balance Theory has had a powerful influence on the development of wellbeing measures throughout the world (Acocck & Hurlbert, 1993; Black-Olten, 1993; Bowling, 1991; Connidis & McMullin, 1993; Flett 1986; Friedman, 1993; Hunt, 1993; Jacobson, 1993; Jenney & Campbell, 1997; Judge & Hulin, 1993; Kerr & Vos, 1993; Muller et al, 1993; Rodd, 1994; Sistler & Blanchard-Fields, 1993; Zika & Chamberlain, 1982). In Aotearoa, for example, the Affectormeter was developed in the early eighties and used in a range of contexts (Barry, 1989; Colhoun & Maxwell, 1986; Flett & Biggs, 1993; Kamman, 1983; Kamman & Flett, 1983, 1986; McIntosh, 1985; Maxwell, Flett & Colhoun 1990).

By the nineties, however, worldwide attention had shifted away from a purely affect-based definition of wellbeing towards the more comprehensive notions of general health or quality of life. Within these approaches, psychological wellbeing is one of several variables that contribute to an individual's overall sense of wellbeing. For example, the General Health Questionnaire (GHQ) set out to measure physical functioning, role limitations, social activities, mental health, health perceptions and experience of pain or problems (Nathawat, 1993; Shams & Jackson, 1993; Stewart et al, 1993). Similarly, the well-known Short Form Health Survey, initially marketed as SF-36, was developed around six dimensions of wellbeing: physical functioning, role functioning, social functioning, mental health, health perceptions and pain (McCallum et al, 1993; Solomon et al, 1993; Ware et al, 1993, 1998). Such approaches face a number of methodological challenges (Christie, 1995; Fayers, 1995; Glaser & Walker, 1995; Price & Harding, 1995; Scott, 1999; Toffler, 1995). Although the concept of psychological wellbeing appears to have been superceded by one of mental health, techniques for the measurement of mental health have continued to rely on Affect Balance Theory.

Within Aotearoa, an opportunity for public debate about the meaning of wellbeing arose in 1988 with the Royal Commission on Social Policy. From the 6000 submissions it was clear this concept had a plethora of meanings. Among non-Māori, wellbeing was said to involve a quality of life, a state of contentment or happiness, a sense of dignity and choice, genuine opportunity for personal control and self-determination, freedom from oppression and the right to participate in a safe, positive and functional society. An S-Factor or intangible quality was also identified. Among Māori, however, the concept of wellbeing was irrevocably tied to the resolution of land and sovereignty issues, balance between the physical and spiritual realms and protection of Māori identity. Indeed, the Commission was drawn to the notion of ngā pou mana, that four supports may be pre-requisites for wellbeing: whanaungatanga (family), ngā taonga
tuku iho (ancestral treasures), te ao tūroa (Māori estates) and turangawaewae (ancestral land) (Royal Commission on Social Policy, 1988). The significance of Ngā Pou Mana was readily accepted by Māori.

During the 1980's, other models had been developed to conceptualise Māori wellbeing and these had generated similar excitement. In particular, the Whare Tapa Whā had symbolically linked Māori wellbeing to a house which needs four walls to stand: taha wairua (the spiritual realm), taha hinengaro (the mental realm), taha tinana (the physical realm) and taha whānau (family) (Durie, 1998). Rose Pere also introduced the idea of Te Whake, the octopus, to symbolize the dynamic, multi-faceted and inter-related nature of Māori wellbeing (1982, 1991). Four tentacles embraced the cornerstones of Whare Tapa Whā but the other four gave substance to the importance of mana (status), mauri (life force), ngā taonga tuku iho and whatumana (emotional life). Waiaora, or complete wellbeing, is said to be found when each tentacle, or dimension of wellbeing, receives sufficient sustenance. Te Roopū Āwhina o Tokānui (1986) also put forward a model which likened the components of Māori wellbeing to a gallery of life. This gallery linked the cornerstones of Whare Tapa Whā with concepts of taia Māori (particularly Māori ways of life), taha Pakēhā (non-Māori ways of life), taha tangata (the social realm), taha whenua (land) and taha tikanga (protocols). Others have contributed much to the debate on Māori wellbeing (Barlow, 1991; Barrett-Aranui, 1981; Durie et al, 2002; Pohatu & Pohatu, 2003; Rangihau, 1977; Walker, 1990).

Table 1 presents the components of Whare Tapa Whā, Te Whake, Ngā Pou Mana and the gallery of life. Commonalities between these models are clearly evident. There is consensus on the importance of whānau, wairua, hinengaro and tinana. In addition, each model has been built around the principle of inter-relatedness, co-existence and symbiosis. It seems the components of Māori wellbeing are not independent, separate or distinct entities but the product of co-relationships, connectedness and mutually beneficial dependencies. Perhaps this is why Rose Pere invoked the notion of waiaora as a river of life-giving forces (1991). As a metaphor, the river implies images of ebb and flow, infinite capacity and interaction between multiple complementary processes.

Notwithstanding the importance of such models for Māori, Aoteraroa has invested heavily in the use of SF-36 and GHQ techniques to measure wellbeing (Brazier et al, 1992; Ministry of Health, 1999, 2002; Relvin 1999; Taylor & McPherson, 1999; Wright & Doughty, 1999). In the absence of viable alternative tools, Māori themselves have also embraced such methodologies (Coupe et al, 1997; Reid & Keefe, 1997; Te Puni Kokiri, 2000). Although convergence between some SF-36 and Whare Tapa Whā items has been raised, it is the differences which are more obvious (Scott et al, 2000). SF-36 is written in the English language, developed around non-Māori concepts and has not captured taha wairua at all (Kokaua et al, 1995). The ongoing use of affect balance techniques to measure and conceptualise the wellbeing of Māori is of considerable concern (Alpass, Neville & Flett, 2000; Brough & Kelling, 2002; Brown et al, 2002; Oliver & Brough, 2002). Under Article Two of the Treaty of Waitangi, Māori are entitled to participation in the Māori world and psychology has an obligation to develop appropriate tools and methodologies for the measurement of psychological attributes.

This paper outlines some fledgling steps towards the development of Hōmāi te Waiora ki Ahau as a tool for the measurement of Māori wellbeing within a Māori conceptual paradigm. The methodologies for item selection and data collection, preliminary indicators of construct validity and directions for improvement of this instrument's psychometric properties are briefly outlined. Three main objectives under-pin the design of Hōmāi te Waiora ki Ahau:

1. the items will comprise all components of the above Māori wellbeing models, ie: Whare Tapa Whā, Te Whake, Ngā Pou Mana and the gallery of life model by Te Roopū Āwhina o Tokānui;

2. the methodological approach will allow each item to be understood, and measured, even when respondents have little, or no, understanding of the Māori language;

3. the measurement tool will be consistent with Māori world views and will seek to maintain the wholistic character of waiaora.

Such objectives were addressed in the following ways:

Selection of items

The conceptual framework for Hōmāi te Waiora ki Ahau comprised twelve items: whanaungatanga, tinana, hinengaro, wairuatanga, mauri, whenua, mana, whatumana, tikanga Māori, tikanga Pakēhā, te ao tawhito and te ao hou. Each concept was defined by the literature and discussed with various mentors in a range of academic and community settings (Palmer, 2002a). Table 2 presents the components of Māori wellbeing that each item was intended to represent and broadly describes its meaning. Conceptually, it was felt that tikanga Māori, tikanga Pakēhā, te ao tawhito and te ao hou had the capacity to capture all that was meant by Māoritanga, Pakēhātanga, tangata,
tikanga, há, taonga tuku iho and te ao tūroa. The validity of this assumption and the content domain covered by each item in Hōmai te Waiora ki Ahau is the subject-matter of ongoing study (Palmer, in press).

**Technique for measurement**
A major challenge, in the development of this instrument, was the need for a technique which allowed the items to be understood and measured when respondents had little or no understanding of te reo Māori, the Māori language. To this end, the value of visual, aural and tactile mediums to portray concepts that have meaning within indigenous epistemologies has been well demonstrated (Jahoda & Lewis, 1988). In an attempt to bridge any gap which may have existed because of perceived inadequacies in te reo Māori, two artists were asked to draw pictures of the waiora items\(^1\). These were added to a series of pictures that had been commissioned by the Department of Health to depict the cornerstones of Whare Tapa Whā\(^2\). The illustrations in Hōmai te Waiora ki Ahau have been presented elsewhere (Palmer 2002a, 2002b).

**The measurement scale**
A further challenge, in the development of this tool, was the need for a technique which is responsive to Māori world views and does not fragment the concept of waiora. Implicitly, the notion of waiora is unidimensional. In other words, waiora is made up of multiple components which collectively contribute to the quality or intensity of experience. Within any one individual, it would seem, the experience of waiora may range through a series of states. Psychometrically, therefore, the Likert method of summated ratings, in particular the summed, bipolar interval scale, seemed a feasible technique for measuring the intensity of waiora gained from each item in Hōmai te Waiora ki Ahau (Hills, 2000; Robson, 1993; Trochim, 2003).

With regard to the lower endpoint of this bipolar continuum, the concept of Te Kore, within Māori cosmogony, is presented as a realm of latent energy, unrealised potential and the state of being from which all things proceed (Buck, 1949). As an endpoint, the symbolism of Te Kore is consistent with the idea that each component of Māori wellbeing is a potential source of waiora that may, or may not, be active. Visually, Te Kore is associated with the colour black, the night before light and the darkness before potential is realised. Māori cosmogony also provided an appropriate symbol for the upper end of this bipolar continuum. Indeed, Uenuku-a-Rangi, the rainbow or aniwaitiwa, is a universal symbol of complete and utter wellbeing (Elliot, 1988). Conceptually, therefore, the Te Kore-Aniwaitiwa continuum provided an appropriate scale to measure the intensity of waiora associated with each item in Hōmai te Waiora ki Ahau.

Theoretically, the number of response options for this type of scale is usually determined by the purposes of study (Trochim, 2003; Worchel, Cooper & Goethals, 1991). The literature has shown that reliability drops sharply when the number of response options is below seven (Chang, 1994; Nummally 1978, Russel & Bokbo, 1992). Others have suggested the optimum scale should have 8 to 20 options and be presented as a continuous measure which does not attach numerical values to sentiments. (Franklin & Jordan, 1995; Rodwell 1987; Scott 1989, Trochim, 2003). A middle value, or odd-numbered scale, is required if respondents are expected to identify a half-way point or neutral position.

In this case, a 13 point scale was chosen. Each response scale was 12cm long and presented as a Te Kore/Aniwaitiwa continuum with no numerical values. When the twelve item ratings were summated, the aggregated waiora score for each individual could, therefore, range from 0 to 144. As a concurrent validity check of this approach to the measurement of waiora, a thirteenth item was added to the measurement scale. This item asked respondents to self-rate their feelings of overall waiora on the same Te Kore/Aniwaitiwa continuum (Palmer, 2002a, 2002b). Statistically, it was hypothesised, there would be no difference between the sum of scores for the twelve

---

<table>
<thead>
<tr>
<th>Hōmai te Waiora ki Ahau item</th>
<th>components of the four Māori wellbeing models that each item represents</th>
<th>broad definition or meaning of each concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>whanaungatanga</td>
<td>taha whānau/whānau/whanaungatanga</td>
<td>family/social relationships</td>
</tr>
<tr>
<td>tinana</td>
<td>taha tinana/tinana</td>
<td>the physical, bodily realm</td>
</tr>
<tr>
<td>hinangaro</td>
<td>taha hinangaro/ hinangaro</td>
<td>the mind and intellect</td>
</tr>
<tr>
<td>wairuatanga</td>
<td>taha wairua/wairua/wairuatanga</td>
<td>the spiritual realm, metaphysical phenomena</td>
</tr>
<tr>
<td>mauri</td>
<td>mauri, há</td>
<td>life force, energy, interconnectedness</td>
</tr>
<tr>
<td>whenua</td>
<td>whenua/turanga/waewae</td>
<td>land, placenta, standing place</td>
</tr>
<tr>
<td>mana</td>
<td>mana/mana ake</td>
<td>uniqueness, divine essence, status, dignity</td>
</tr>
<tr>
<td>whatumanawa</td>
<td>whatumanawa</td>
<td>emotions and emotional life</td>
</tr>
<tr>
<td>tikanga Māori</td>
<td>Māoritanga/tikanga/taonga tuku iho/tangata</td>
<td>Māori protocols and ways of doing things</td>
</tr>
<tr>
<td>tikanga Pākehā</td>
<td>Pākehātanga/tikanga/tangata</td>
<td>Pākeha protocols and ways of doing things</td>
</tr>
<tr>
<td>te ao tawhito</td>
<td>há/taonga tuku iho/Māoritanga/tikanga</td>
<td>ancestral treasures, gifts from the old world</td>
</tr>
<tr>
<td>te ao hou</td>
<td>te ao tūroa/tangata/Māoritanga/Pākehātanga/tikanga</td>
<td>Māori estates, opportunities and prospects</td>
</tr>
</tbody>
</table>

---

* New Zealand Journal of Psychology Vol. 33, No. 2, July 2004
component items of Hōmai te Waiora ki Ahau (the aggregate waiora score) and the thirteenth item, a self-rating of overall waiora (the self-rated waiora score).

Research Design

This paper draws on three sets of data taken from three distinctly different groups. Two of these data sets were collected for the purposes of PhD study (Palmer, 2002a). In the first instance, Hōmai te Waiora ki Ahau was pre-tested prior to implementation in the pilot study. The pre-test mainly aimed to see whether the tool could be administered. Secondly, the tool was piloted with a larger group of participants. The third set of data was gathered when Hōmai te Waiora ki Ahau was administered to a group of summer school students.

Although the sample sizes involved in this research design are too small for definitive statistical analysis, it is clear that the data can still provide information about construct validity. Among contemporary psychometricians, the term construct validity encapsulates all that has historically been meant by validity and its many qualities or criteria (Embretson, 1983; Murphy & Davidshofer, 2001, Trochin, 2003). For any one tool, the process of establishing construct validity relies on the systematic and gradual accumulation of data from a variety of sources (Anastasi & Urbina, 1997; Embretson, 1983; Murphy & Davidshofer, 2001). An understanding of construct validity is gained from information about the behaviour of a tool in a range of contexts and numerous psychometric techniques can assist the accumulation of such knowledge.

In this case, for example, it was reasonable to expect some difference in the pattern of responses for each group. Furthermore, if the notion of unidimensionality held true, there should be no difference between the aggregate and self-rated waiora scores. It was also reasonable to expect that a Māori measure of psychological wellbeing would produce a different response than a non-Māori measure.

The minimum sample size for multivariate analysis is clearly the subject of ongoing debate (Alinga & Olejnik, 2000; Cicchetti, 1999).

However, the absolute minimum sample size for inferential statistics of any kind is generally said to be twenty-five (Hills, 2000; Wonnacott & Wonacott, 1982). Although inadequate for definitive analysis, a sample of twenty-five will allow the use of multivariate statistics, the development of benchmarks for later studies and an opportunity to explore reliability theory (Anastasi & Urbina, 1997; Embretson, 1983; Murphy & Davidshofer, 2001). A number of techniques have, therefore, been used to examine the reliability of pilot study data. In particular, this data has been screened for normality, difference between mean item scores and internal consistency using Chronbach’s α and r̄ internal coefficients (Hills, 2000; Graham & Lilly, 1984; Tabachnick & Fidell, 1996).

Method

Participants

Ten rural Māori women, aged 16-65 years took part in the pre-test of Hōmai te Waiora ki Ahau. Māori women from Hauraki were recruited for the pre-test because this was the target group for doctoral study. (Palmer, 2002a). Recruitment took place in Harataunga, a small Māori settlement in Hauraki. All but one of the women in this group were beneficiaries. Individual interviews were held in the participant’s own home.

Thirty-one Māori women, aged 16-34 years, took part in the pilot of this tool during the third trimester of pregnancy. All of the women were booked to give birth at Thames Hospital in Hauraki. A third were teenagers and the vast majority were low income sole parents with two or more children at home. Individual interviews were held in each participant’s home.

Thirteen summer school participants aged 25-50 years completed Hōmai te Waiora ki Ahau as part of a Māori health session. Two in this group were male, three were non-Māori and all were employed. In this instance, the tool was administered to the group, there were no individual interviews and data collection took place in Wellington.

Measures

Participants in these three groups completed Hōmai te Waiora ki Ahau and Affectometer 2, a non-Māori ten-item tool for the measurement of psychological wellbeing (Kammann & Flett 1983, 1986).

Procedure

In all cases, Hōmai te Waiora ki Ahau was presented in the following manner:

- the concept of waiora was discussed using the Te Kore-Uenuku/Awaiti measurement scale to demonstrate how the intensity of experience may range from unrealised potential or not yet active to complete and utter waiora;
- respondents were shown how the Te Kore-Uenuku response scales could be used to measure the degree to which each component was a source of waiora;
- the illustrations for each of the twelve items were presented respectively and used as a visual aide to assist discussion about the meaning of each concept;
- participants were asked to mark the pathway between Te Kore and Uenuku at a place which they felt represented the extent to which each item was a source of waiora in their lives at present;
- participants were asked to self-rate their overall sense of waiora.

Participants were then asked to complete the ten items for Affectometer 2, along a -2 to +2 response scale. Interviews took thirty to fifty minutes.

Results

Descriptive statistics were used to look at broad differences in the means, range and distribution of responses for each group.

Figure 1 displays Hōmai te Waiora ki Ahau item means and standard deviations along with the self-rated and aggregate scores for the three groups of participants. Between group differences in the mean scores are clearly evident. Most notable are the higher means for all items except mauri, tikanga Māori and te ao tawhito among summer school students. Whanaungatanga, tinana, hinengaro and mana were the strongest sources of waiora for this group. In contrast, ngā wāhine hapū had the lowest mean scores for all items except tikanga Māori, te ao tawhito and mauri. These three items provided the most waiora for ngā wāhine hapū but tikanga
Pākehā and te ao hou yielded the least. Among pre-test participants, mean scores were highest for mauri and te ao tawhito but lowest for tikanga Pākehā and whatumanawa. For all groups, little difference was found between aggregated and self-rated waiora mean scores. Among pilot study participants, the relationship between aggregate and self-rated waiora was linear ($r = .478$, $p<.01$) and no correlation was found between the aggregate waiora and Affectometer 2 scores.

Between group differences in the variability of responses are evident. Among pilot study participants, standard deviations ranged from 2.39 to 4.07 being lowest for tinana, mauri and te ao tawhito but highest for whenua, mana, tikanga Pākehā and te ao hou. Among pre-test participants, deviations ranged from 1.76 to 3.78 being most similar for mauri and tikanga Māori and most variable for whanaungatanga, whenua and tikanga Pākehā. Among summer school participants, standard deviations ranged from 1.29 to 4.15 with little variability in responses for whanaungatanga, tinana, hinengaro, mana and whatumanawa but considerable difference in the scores for tikanga Māori and te ao tawhito.

Table 3 presents the range of scores obtained for items and self-rated waiora by each group. In comparison with the pretest and summer school groups, pilot participants were considerably more likely to utilize the full range of response options. Nevertheless, this group did not utilize the lowest quartile of response options for wairuatanga, mauri, te ao tawhito and self-rated waiora. In contrast, the pretest and summer school participants almost never utilized the full range of response options and the scores for most items never fell into the lowest quartile. This finding explains the low variability for most items among pretest and summer school participants. However, some other explanation is needed to explain the low variability on some pilot items.

By central tendency, nine of the twelve pilot means fell within the middle zone of the rating scale. The exceptions were mauri, te ao tawhito and tikanga Māori. Furthermore, differences between the mean and median were >1 for whanaungatanga, hinengaro, mauri, mana, te ao tawhito and tikanga Māori. Such findings suggest these items may have irregular distributions.

Skewness and kurtosis coefficients were used to determine the significance of irregular distributions within the pilot data. Skewness was significant for whanaungatanga ($s = -1.194$, $p<.01$); hinengaro ($s = -0.833$, $p<.05$); mauri ($s = -1.223$, $p<.01$); whenua ($s = -0.915$, $p<.05$); mana ($s = -0.877$, $p<.05$); te ao tawhito ($s = -1.139$, $p<.01$) and tikanga Māori ($s = -2.490$, $p<.01$). In all cases, the skewness was negative and indicates a tendency towards the upper end of the scale. A positive kurtosis, evident in the distribution of tikanga Māori scores ($k = 6.752$, $p<.01$), suggests responses were clustered closely together.

Screening for central tendency also provides information about the distribution of scores for the tool as a whole. For the pilot data in general, the mean of all scores, at 95.64, was within one standard deviation of the total mid-score. This suggests borderline robustness and again highlights the predominance of higher scores. At 16.98 the standard deviation of scores was relatively low and suggests the measure could have been more sensitive to individual differences. Nevertheless, the distribution of pilot scores did not show significant skewness or kurtosis and there was no disparity between the mean and median. Furthermore, both ANOVA, [$TF(30,11) 9.4559$, $p<.001$] and Hotelling's T² [$TF(11,20) 4.3069$, $p<.001$]

Table 3: Range of scores for Hōmai te Waiora ki Ahau items by pretest, pilot and summer school groups

<table>
<thead>
<tr>
<th></th>
<th>pretest min</th>
<th>pretest max</th>
<th>pilot min</th>
<th>pilot max</th>
<th>summer school min</th>
<th>summer school max</th>
</tr>
</thead>
<tbody>
<tr>
<td>whanaungatanga</td>
<td>1</td>
<td>11</td>
<td>0</td>
<td>12</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>tinana</td>
<td>4</td>
<td>12</td>
<td>1</td>
<td>11</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>hinengaro</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>wairuatanga</td>
<td>6</td>
<td>12</td>
<td>2</td>
<td>12</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>mauri</td>
<td>6</td>
<td>12</td>
<td>3</td>
<td>12</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>whenua</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>mana</td>
<td>5</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>whatumanawa</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>7</td>
<td>12</td>
</tr>
<tr>
<td>tikanga Māori</td>
<td>7</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>tikanga Pākehā</td>
<td>1</td>
<td>12</td>
<td>0</td>
<td>11</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>te ao tawhito</td>
<td>3</td>
<td>12</td>
<td>3</td>
<td>12</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>te ao hou</td>
<td>4</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>self rated waiora</td>
<td>6</td>
<td>12</td>
<td>2</td>
<td>12</td>
<td>4</td>
<td>12</td>
</tr>
</tbody>
</table>
were significant. Such findings suggests the measure was sensitive to individual differences.

Homogeneity, or internal consistency, of the pilot data was assessed through $r_{nu}$ coefficients and Chronbach's coefficient alpha ($\alpha$). In both approaches, internal consistency is estimated from average inter-correlations among the items and reliability is demonstrated when items inter-correlate (Tabachnick & Fidell, 1996).

As a general convention, an $r_{nu} > .7$ indicates an item that may need to be discarded or changed, because it is too similar to other components, whereas a significant $r_{nu}$ flags an item which is different but consistent with other items in the scale (Hills, 2000). Chronbach's $\alpha$ provides a single indicator of internal consistency. Ideally, this should be $\geq .7$ but .6 is acceptable for a pilot of this kind. This coefficient can be used to check the contribution each item makes to the total score. If coefficient $\alpha$ increases when the item is deleted it is assumed the item is not making a significant contribution to the reliability of the scale.

Table 4 presents the pilot data for Hōmai te Waiora ki Ahau by $r_{nu}$ and $\alpha$ coefficient if item deleted. None of the items had an $r_{nu} > .7$ and six items reached significance ($r_{nu} = .3494$, $p < .05$). Chronbach's $\alpha$ for the tool was .6846 and this figure would only have improved slightly with the removal of items. In general, such findings suggest the items were making their own contribution to the total score but there is certainly room for improvement in the internal consistency and reliability of this measure.

**Discussion**

Despite the small sample sizes involved in these studies, the data gathered for Hōmai te Waiora ki Ahau thus far provides several indicators of construct validity. It is clear, for example, that respondents understood its purpose, were willing to engage in the process and were able to respond. As a measurement tool and data gathering technique, it would appear that Hōmai te Waiora ki Ahau is certainly workable.

Furthermore, between group differences in the mean item scores and the direction of such differences seemed to make intuitive sense and were consistent with theoretical expectations about the likely direction of responses. In particular, the summer school group had considerably higher mean scores for whanaungatanga, tinana, hinengaro, mana, tikanga Pākehā and te ao hou. This could be explained by two characteristics of the summer school group. Firstly, these participants were all employed and, secondly, this group comprised a small number of non-Māori members. It is certainly reasonable to expect that feelings of self-esteem, social support, physical and mental wellbeing and attitudes towards non-Māori ways of doing things would be more favourable among people who are employed, able to engage in career development opportunities and more representative of non-Māori.

The variability in summer school responses to te ao tawhito and tikanga Māori also made sense given the presence of non-Māori participants.

At first glance, it was of concern to find that mean scores for most waiora items were lowest for ngā wāhine hapū. Given, however, that these women were in their third trimester of pregnancy such a finding could be expected and would make sense to any woman who has experienced the physical, emotional and spiritual changes of late pregnancy. Reduced activity and withdrawal from participation in te ao hou is a natural part of maternal birth preparation. In stark contrast, however, ngā wāhine hapū reported the highest mean scores for tikanga Māori, te ao tawhito and mauri. Such findings also make tremendous sense given the powerful lifeforce these women were carrying and the imminence of childbirth. It would appear that opportunities to experience, think about and reflect on traditional Māori concepts with its holistic rituals and emphasis on protection, sacredness and the joining of past and future generations may offer peace of mind and wellbeing to Māori women as they prepare for the journey of childbirth. At the very least, such findings suggest these items are strong sources of waiora for Māori women during pregnancy and the implications for development of Māori maternity services are clear.

Other indicators of construct validity are evident in the pilot data. No evidence of a linear relationship was found between Hōmai te Waiora ki Ahau and Affectometer 2 responses. Such a finding suggests that the tools are indeed measuring theoretically distinct constructs. This is an important finding because it leans towards the view that non-Māori measures of wellbeing may be inappropriate for some Māori. Most notable, however, is the linear relationship between aggregated and self-rated waiora scores. This finding supports the notion of unidimensionality and suggests the aggregate score was a good predictor of self-rated waiora. The alpha coefficient is also supportive of unidimensionality, that each of the twelve items did indeed contribute to the aggregate waiora score. Although borderline, the alpha coefficient was acceptable for a pilot study (Hills, 2000).
Ideally, an alpha co-efficient of .8 to .9 is required to demonstrate unidimensionality and a range of techniques may help to improve this co-efficient if necessary. One such technique is to identify redundant or badly performing items through item-total correlations and/or screening for irregularities in the range and distribution of scores. Once an irregular item is identified, methodologies to improve the distribution of scores can be implemented or the item can be removed. In the case of this pilot, the alpha coefficient would not have improved with the removal of items but a number of irregular distributions were identified.

It is clear that participants did not always utilize the full range of response options and ceiling effects were evident in the distribution of some scores. In addition to the tendency towards higher scores, responses were sometimes clustered too closely together. By ANOVA and Hotellings T², however, there is reason to feel confident that the tool was sensitive to individual differences. This means response patterns differed from person to person and participants were thinking about their responses rather than simply going through the motions or doing the same thing. Collectively, such findings raise a number of issues.

Firstly, the response patterns may simply be an artefact of the small sample size. Secondly, it is possible that responses to some of the items in this tool are naturally high and the scale itself or the approach to measurement of these items is inappropriate. Thirdly, the meaning of some items may be unclear and participants may be unsure how to respond. And fourthly, perhaps the concept of waiora, as presented in this tool, is not unidimensional at all and a multi-dimensional approach to the measurement of this construct may be more suitable.

There are a number of ways in which such issues could be addressed but the most sensible place to start would be with content domain, or the definition of items. The implementation of methodologies to ensure that content domain has been appropriately, adequately and accurately defined is clearly a fundamental and critical first step in the development of any psychological measurement tool (Murphy & Davidshöfer 2001, Trochin, 2003). The items in this tool were largely defined by the literature and consensus in expert opinion. This approach was appropriate at the time but it would be prudent to revisit the meaning of items among contemporary Māori. Once the boundaries for each item have been defined and described then attention can shift to other matters, particularly the issues around generalisability, whether the tool is uni- or multidimensional and continuing to accumulate knowledge about construct validity (Embreton, 1983; Embreton & Reise, 2000; van der Linden & Hambleton, 1996).

It is likely that the development of Hōmai te Waiora ki Ahau will hold a number of advantages for Māori (Palmer, 2003). In particular, it has the capacity to:

- foster a transformation of consciousness towards psychological constructs and concepts which have relevance for Māori;
- provide a culturally responsive tool for the measurement of wellbeing which can be administered when respondents have little or no understanding of the Māori language;
- be used as a health outcome measure or technique for performance review and monitoring change;
- generate opportunities for Māori to debate and discuss the manner in which their lives may be influenced by the components of waiora;
- provide a unidimensional score and/or a multi-dimensional profile of scores that may identify pathways for individual or collective development and change;
- ensure the development of Māori capacity, capability and expertise in the use of psychometric theory and techniques.

Although Hōmai te Waiora ki Ahau is not yet ready for use by clinicians, provider organisations or the general public, this tool clearly warrants ongoing study and development.

References


University of Waikato, Hamilton.


Reid, P. & Keeve, V. (1997). Health service utilisation by Māori women and their whānau. Te Roopū Rangahau a Eru Pomare, Wellington School of Medicine, Wellington


Notes
1. Tumohe Clarke (Ngāti Koroki, Ngāti Tipa, Ngāti Hauna) and Elizabeth Anderson (Ngāti Hauna-Wihangaunui River, Tōwharetoa, Ngāti Maru-Tarannaki).

2. Drawn by Robyn Kahuikwa (Ngāti Porou) during the early 90s.

3. Skewness and kurtosis coefficients will equal zero if a distribution is normal, when divided by its standard error an outcome of 1.98 or 2.58 indicates significant difference from a normal distribution at the p<0.05 and p<0.01 levels respectively.

Author Note:
Stephanie Palmer, Ngāti Porou, Erithapeti Murchie Postdoctoral Research Fellow, Te Pūmanawa Haurua, Te Pūtahi-a-Toi

Address for correspondence:
Dr Stephanie Palmer
Massey University
PO Box 756
Wellington

Email: S.K.Palmer@massey.ac.nz