- Project Report-
A comparison of invariance across online and postal survey modes: a pilot study conducted ahead of the 2018 Health, Work and Retirement survey

Report prepared by
Joanne Allen
Health and Ageing Research Team
Massey University

Summary

This pilot study assessed the potential for psychometric differences to arise between online and postal survey administrations of the core measures of the New Zealand Health, Work and Retirement study. These measures include demographic indices, as well as multi-item measures, namely the SF-12, the CES-D-10, the De Jong Gierveld and Tilburg scale for social and emotional loneliness, the CASP-12, the Social Provisions Scale, the Economic Living Standards Index and, the LS-CAPE. Participants in the pilot (aged 55+) were recruited via advertisement, with participation promoted through Twitter and routine newsletters released by Grey Power New Zealand (https://greypower.co.nz/) and Age Concern New Zealand (https://www.ageconcern.org.nz/). Participants were invited to complete the survey online ($n = 252$ completed) and then again approximately 3 weeks later in a paper booklet format delivered to their postal address ($n = 123$ returned). Due to a low rate of paper surveys returned, an age, gender and education matched sample of responded from persons who returned the paper booklet administered in the 2016 Health, Work and Retirement survey ($n = 504$) were extracted for model comparison. Results indicate that measurement models for core measures of the Health, Work and Retirement survey were maintained across survey modes. A stepped approach to adopting an online survey mode is recommended to mitigate risks to the HWR longitudinal study.

Contents

Project rationale

The New Zealand Health, Work and Retirement study (HWR) is a longitudinal study of persons aged 55 years and over conducted by the Health and Ageing Research Team at Massey University. The study commenced in 2006 as a biennial postal survey assessing experiences of health and wellbeing among older New Zealanders. The study broadly aims to track and identify factors associated with positive and negative health outcomes in this population over time. New participant cohorts are regularly recruited to the study to maintain its representation of persons aged 55+ years of age and participants have been invited to complete postal surveys every two years. To date, almost 11,000 participants have returned over 25,000 surveys.

While postal surveys conducted in New Zealand may take advantage of an excellent sampling frame (the New Zealand electoral roll), paper-based surveys are resource-intensive in terms of materials, postage, data entry, data cleaning, physical storage, and require participants to return-mail the completed survey. In contrast, online survey administration requires relatively minor material and time costs and can incorporate in-built data validation to prevent entry of invalid responses, thus reducing missing data and use of resources associated with data cleaning.

For these reasons, the Health and Ageing Research Team have considered offering participants the option to complete the 2018 wave of the New Zealand Health, Work and Retirement study online. However, the potential for responses to the online survey to not be comparable to those to the postal survey is recognised. This poses a significant risk to not only the continuity of the longitudinal research project, but in the potential collection of data that would not be usable as designed, wasting participant time and effort. A review of the literature assessing the measurement invariance of multi-mode surveys suggest that self-administered survey modes (i.e., paper and online) generally display a high level of equivalence and that they display lower error variance than that observed between self- and interviewer-administered (i.e., face to face) survey modes (Hox, De Leeuw, & Zijlmans, 2015). However, it is possible that this consensus in part represents a publication bias in favour of models that confirm the comparability of results across survey modes. Factors such as variations in sample source and demographic composition also appear to influence findings of measure equivalence across different survey modes.

This pilot study aimed to compare the psychometric consistency of data collected via online and postal survey administration of core measures of the HWR survey. In light of the potential impacts of sample effects on psychometric consistency of responses, both online and paper survey formats were administered to a single sample of persons aged 55+ to evaluate differences in responses attributable to survey mode.

Protocol

Participants were invited complete both the online and paper versions of the survey. Study materials (information statement, letters and survey form) are provided in Appendix 1. A low-risk notification for the project (Ethics Notification Number: 4000018541) was made to Massey University Human Ethics Committee. The below flow chart summarises the participation/data collection procedure.
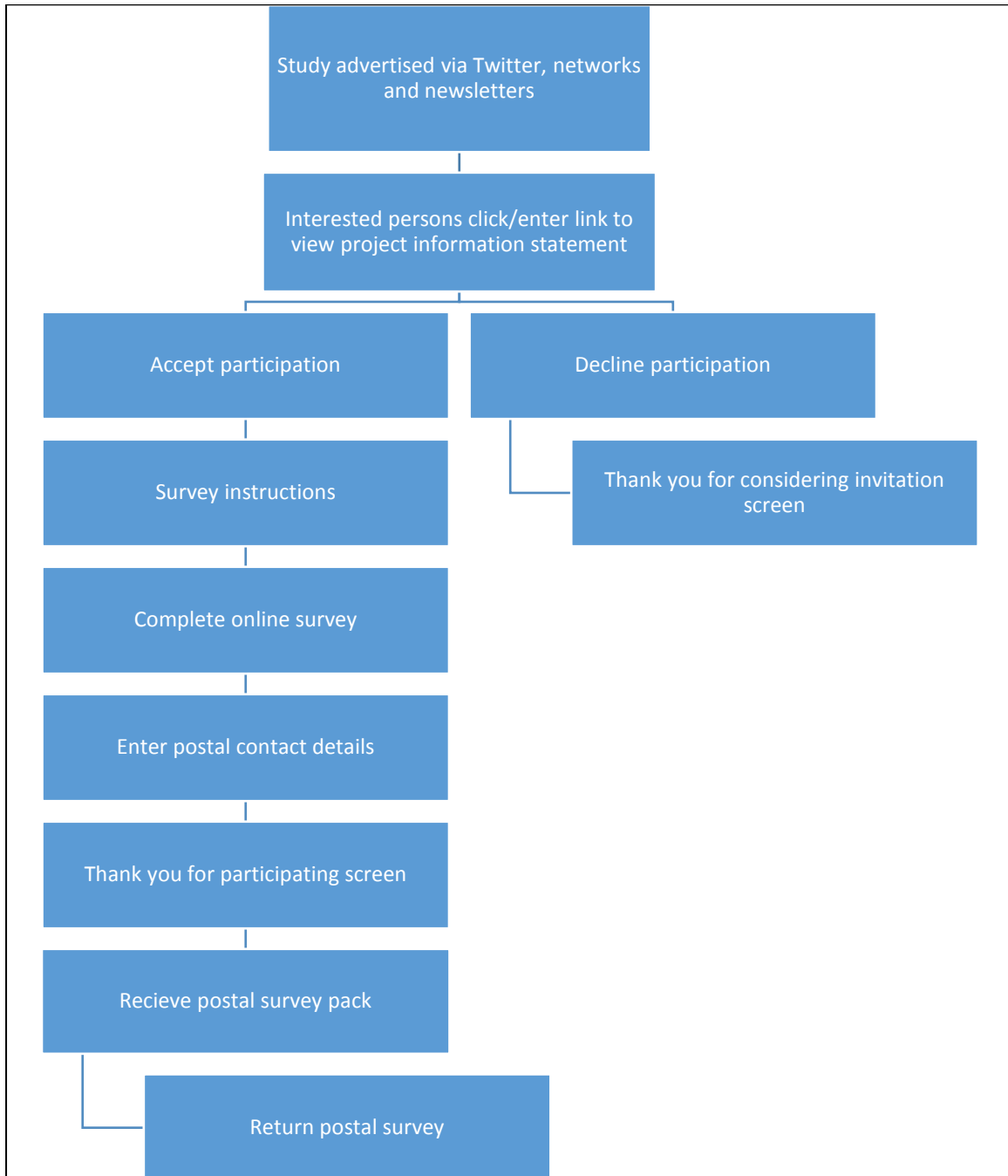
Figure. Flow chart illustrating data collection procedure.

The study was promoted via Twitter and in newsletters from Age Concern and Grey Power and included a URL (https://psylab.massey.ac.nz/hart) to the study information sheet:

Figure. Example of study advertisement.

From the study information sheet, potential participants could either accept or decline the invitation to participate by clinking the relevant button at the bottom of the page. If participants declined, they were shown a message thanking for their time and did not progress to the survey:

Figure. Screen thanking participants for considering study invitation

If they accepted, then they saw the instructions on how to complete the survey, including bowser requirements, before beginning the online survey:
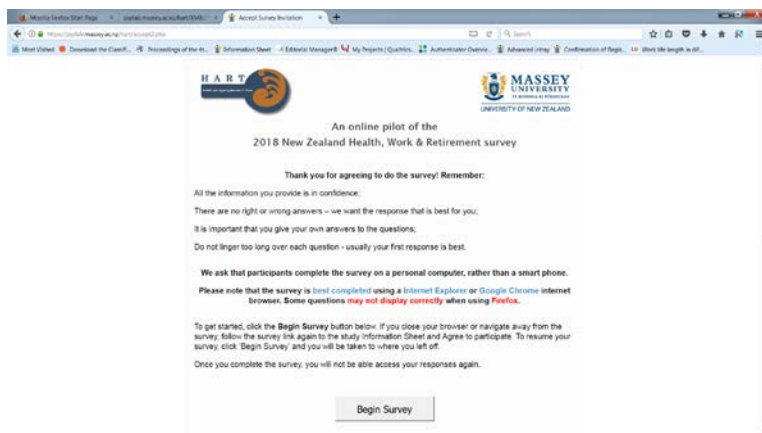
Figure. Instructions for completing the online survey

Online survey

Once participants clicked the 'Begin Survey' button they were sent to the online survey hosted by Qualtrics. Online and paper surveys were formatted to closely match that of past NZ Health, Work and Retirement surveys and each other. Figures below provide examples of measure layout by mode.
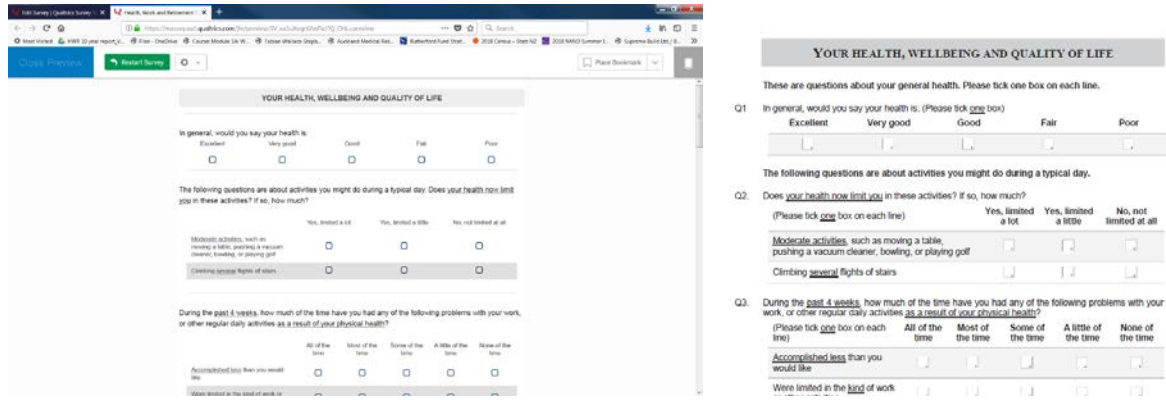
Figure. Layout of SF-12 v2 in the online (left) and paper (right) surveys.
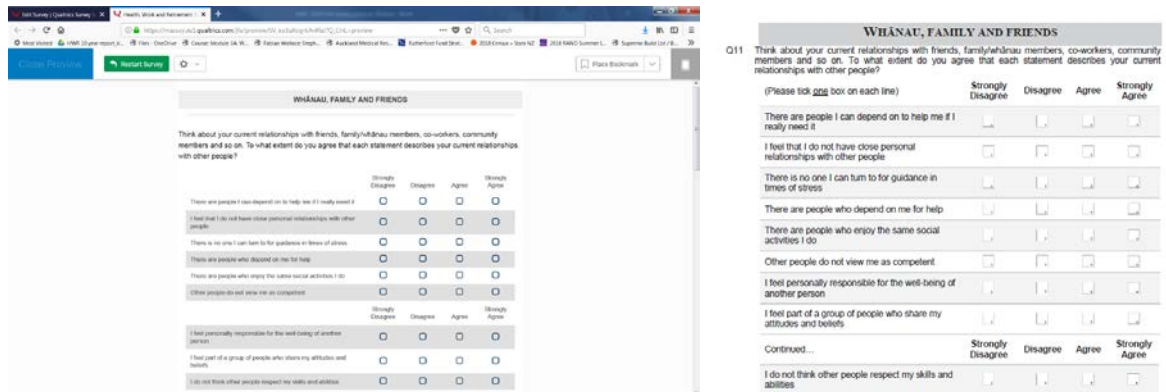
Figure. Layout of Social Provisions Scale in the online (left) and paper (right) surveys.
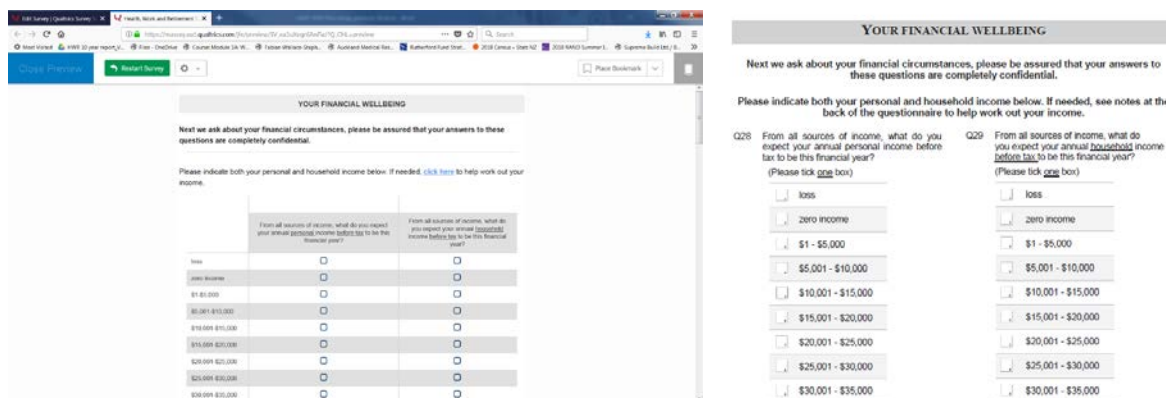
Figure. Layout of self-reported income brackets in the online (left) and paper (right) surveys.

Participant privacy online

Online survey data were collected via the Qualtrics online survey platform. In assessing the security of this platform, the relevant technical documentation regarding secure encryption in the collection and storage of data were consulted: 'Security White Paper Lite: Information Security – a brief overview of privacy, compliance, and operational policies and procedures' (February 6, 2017: version 5.01). All data collected for the study via Qualtrics are owned and controlled by Massey University and cannot be accessed by any other party or for any other purposes.

As respondents may wish to complete the online survey over multiple sittings, the survey could be revisited and accessed by anyone who signed onto the same computer with the same computer login details and accessed the survey link until the survey had been completed (as indicated by having clicked through the entire survey). Once the participant has completed the survey, their responses could not be accessed again. On completing the online survey, a random number was assigned to each case response and respondents were directed to a separate survey hosted on a secure Massey server to input their postal contact information. The random number was added to the URL link directing the participant to the contact information survey and recorded with the participants contact information.

Paper survey

Participants were posted paper survey packs 1-2 weeks following their online survey response. Packs included a window-faced envelope, a 20-page survey booklet, an addressed letter, and a return-addressed reply-paid envelope. The random number generated at the time of the respondent's online survey completion was written at the top left-hand corner of the survey booklet to anonymously link online and paper survey responses.

Data from returned paper surveys were entered into a dummy version of the Qualtrics online pilot questionnaire. This was done to reduce data entry errors and ensure that variable coding remained consistent. Data exported from Qualtrics were stored on a secure Massey server. Participant identifying and contact data were stored in a separate database on a separate secure drive and deleted at the conclusion of the project (February 26, 2018). Paper surveys booklets are stored in a locked storeroom in the School of Psychology to be destroyed five years after the completion of the study.

Project close

The online survey opened on November 1, 2017 and closed on January 18, 2018. Data included in the current report were obtained from hardcopy surveys returned by February 20, 2018. Between Nov 3 - Dec 21 survey was promoted through:

- Twitter:
  - **Chris Stephens** @tokeawa 3 Nov 2017
    - 10 Retweets, 1 like
  - **Massey HART** @MasseyHART 22 Nov 2017
    - 3 Retweets, 2 likes
  - **Andy Towers @AndyTowersNZ** 22 Nov 2017
    - 5 Retweets, 4 likes
  - **Massey Uni Health** @MasseyUniHealth 28 Nov 2017
    - 11 Retweets, 3 likes
  - **Massey University** @MasseyUni 4 Dec 2017
    - 8 Retweets, 3 likes
  - **Age Concern** @AgeConcernNZ 20 Dec 2017
    - 3 Retweets, 5 likes

- o **Massey University @MasseyUni** 21 Dec 2017
  - 4 Retweets, 4 likes

- Word of mouth/email to private networks

- November subscriber e-newsletter released by Grey Power New Zealand (https://greypower.co.nz/) and an internal staff newsletter from Age Concern New Zealand (https://www.ageconcern.org.nz/).

Recruitment to the project closed January 18, 2018 to allow final hard copy surveys to be mailed out and data received ahead of the preparation of the project report. After this date, the information statement accessed via the online survey link was replaced by a message that the survey was now closed.
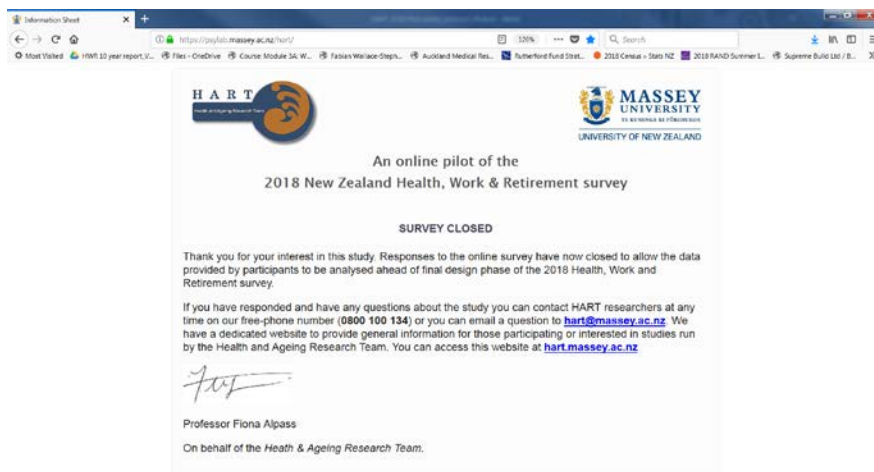
Figure. Message displayed via the survey link after pilot project was closed to responses.

Reported issues and feedback from participants

- A participant reported an issue with not being linked to the instruction form from the information/consent form when clicking 'accept'. When provided with the link to the instruction form, the participant was not linked to the survey from the instruction form when clicking 'begin survey'. The participant was subsequently provided a link directly to the survey.

- A participant reported that they could not change their response to the income question and thus multiple responses were entered.

- Some participants reported that they had entered their address into the address form but never received the paper survey. Inspection of the address database showed that no addresses were recorded on the system from these participants. It is unclear why some address records were not recorded.

Analytic approach and sample

Initial model selection

Measures were modelled and assessed for fit to data using Confirmatory Factor Analyses (CFA). In addition to survey mode (online vs postal), researchers were mindful of other factors which may influence model fit and confidence in findings related to model invariance across measure administrations. These included the appropriateness of the model specified, differing demographic characteristics of respondents (e.g., age, gender and education) between groups (see Hox et al., 2015), and the statistical power of the models.

Initial models for assessment of invariance were specified with consideration to the original theoretical construction of the measure, existing literature on the measure's factor structure and, initial model fit. Adequate initial model fit is required to be able to proceed with assessment of differential model fit in nested invariance models. Consideration was also given to how measures are conceptualised and historically scored in the Health, Work and Retirement study – as such, extensive model re-specification and item elimination was not undertaken to achieve model fit and paired comparison of summated scale scores were conducted across the online and paper survey measures.

Sample size calculation

Simulation studies indicate that the sample size required to achieve adequate statistical power in CFA varies with the number of latent variables, the number of indicators, the magnitude of the factor loadings and the correlation between latent factors (Wolf, Harrington, Clark, & Miller, 2013). The minimum required sample size increases with the number of latent variables in a model, is greater for factors with fewer indicators, and decreases with stronger factor loadings. For models with more than one latent factor, greater factor correlations also decreases the minimum sample size compared to lower factor correlations. The power of invariance tests are similarly influenced by the number of factor indicators and their loadings (i.e., precision of model estimation) in addition to sample size. For example, models with three indicators and high factor loadings may have more accurately estimated parameters, and thus tests of metric invariance across samples are more likely to fail. Recommended minimum sample sizes for CFAs with different number of factors, indicators, loading magnitudes and correlations are reproduced from Wolf et al. (2013) in the table below. The sample size corresponding to each of the seven initial CFA models tested are indicated in the right hand column.

Table. Minimum sample size requirements to achieve adequate statistical power in CFA models by number of factors, indicators and magnitude of factor loadings (Wolf et al., 2013).

N factors	N indicators per factor	Magnitude of factor loadings	Min N to achieve 80% power, alpha .05		Model
			Factor correlation 0.3	Factor correlation 0.5	
1	4	0.50	190	.	
	4	0.65	90	.	SPS subscales[1]
	4	0.80	60	.	ELSI-SF
	6	0.50	90	.	
	6	0.65	60	.	LSCAPE-6
	6	0.80	40	.	
	8	0.50	90	.	
	8	0.65	50	.	CESD
	8	0.80	30	.	
2	3	0.50	460	310	Loneliness
	3	0.65	200	170	
	3	0.80	120	100	
	6	0.50	190	100	
	6	0.65	120	60	
	6	0.80	100	40	
	8	0.50	160	100	SF-12
	8	0.65	120	60	
	8	0.80	90	60	
3	3	0.50	420	320	
	3	0.65	220	150	CASP-12
	3	0.80	150	100	
	6	0.50	190	160	
	6	0.65	130	80	
	6	0.80	100	60	
	8	0.50	160	120	
	8	0.65	110	80	
	8	0.80	100	80	

Following initial model specification, the aim of our planned analytic approach was to assess the invariance of measures across the online and paper survey modes for the same participants. For longitudinal models (within the same sample), the number of factors and variables for each model essentially double, necessitating a larger sample size. To assess the invariance of our CFA models and the consistency of item-weighted factor scores (as assessed by paired-samples t-test across online and paper pilot responses), we aimed to obtain $n = 400$ responses to both the online and paper surveys.

Respondent sample and revised analytic plan

The achieved sample size was lower than that required for our initial analytic strategy to be applied to all models. There were $n = 252$ valid responses to the online survey. Of these, $n = 160$ (63.5%) provided postal address details at which to receive the paper survey and $n = 123$ (48.4%) returned a paper survey. As such, the sample of participants who returned both online and paper mode surveys was too small to achieve adequate statistical power and to represent low-frequency item responses. In light of this, the analytic plan for assessment of model invariance across survey modes was revised. In the revised plan, two samples were used to assess the invariance of measures across postal and survey modes using multi-group confirmatory factor analyses (MGCFA): online pilot responses ($n = 252$) were compared with an age, gender and education matched sample of responders to the

---

[1] To assess the invariance of the Social Provisions Scale, the invariance of each first-order subscale (6 subscales each comprised of 4 items) was assessed independently. This was due to the limited literature on the factor structure of the Social Provisions Scale and very poor fit of the 24-item 6 factor model to data collected in the 2016 Health, Work and Retirement survey.

2016 HWR survey (n = 504), which was administered via paper survey booklet. The matched sample was drawn in a 2:1 ratio to online responders to increase variability in the set without unduly effecting the chi-square values and associated model fit statistics.

Sample characteristics of the online sample, the subsample who completed both the online and paper surveys and the matched sample of HWR respondents are presented in the table below. The initial model fit for each measure was assessed in both the matched sample and online pilot data (n = 252) prior to assessment of multi-group invariance across survey modes. Even with the revised analytic plan, there were inadequate observations to assess the invariance of the De Jong Gierveld and Tilburg scale for social and emotional loneliness in the online pilot sample.

Table. Demographic characteristics of all online respondents, the subsample of respondents completing the online and paper surveys, and data from the matched sample of respondents to the 2016 HWR survey.

	Online	Online + Paper subsample	HWR matched sample
N	252	123	504
Age (M, SD)	65.30 (7.47)	66.11 (7.85)	66.15 (6.95)
Female%	70.6	73.1	70.6
Education			
No qualifications%	3.6	4.2	3.6
Secondary school%	8.3	10.8	8.3
Post-secondary school%	25.8	22.5	25.8
University%	62.3	62.5	62.3
NZ European%	89.4	90.7	78.6
Fully retired, no paid work%	26.9	33.6	25.6
Mean personal income band *(11 = $40k-$50k)*	11.04 (3.79)	10.60 (3.73)	10.67 (3.59)
Mean household income band *(13 = $60k-$70k)*	12.77 (3.48)	12.26 (3.43)	12.56 (3.34)
Self-rated overall health *(item range 1-5)*	2.23 (.89)	2.13 (0.82)	2.24 (.90)

Survey time and lag duration

Excluding one respondents who took 46 hours and 16 minutes to complete the online survey, respondents took an average of 19 minutes (SD = 13m, range 00h07m - 02h33m) to complete the online survey. There was an average lag of approximately 3.5 weeks between submission of the completed online survey and return of the paper survey (range: 6 - 73 days; see Figure below). No systematic differences between the online and paper survey responses to single item ratings (i.e., income, qualifications etc.) were expected to be attributable to this lag.
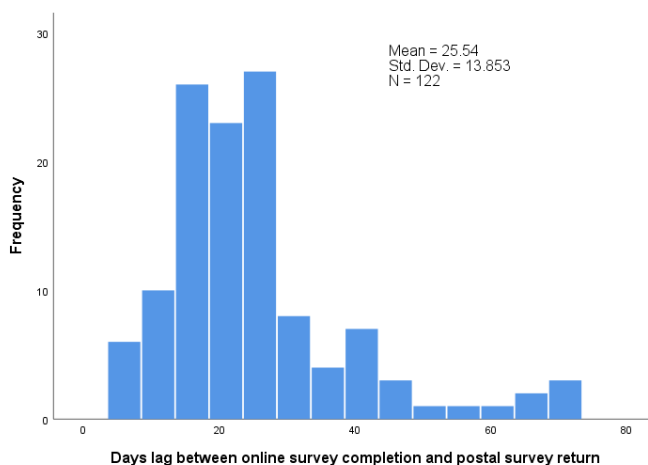

Figure. Histogram plotting the number of days lag between completion of online survey and return of postal survey.

Invariance of multi-item scales across survey modes

Following appropriate recoding of variables, missing item data were imputed within each sample using the SPSS v24 inbuilt EM algorithms. Summated cores were calculated and paired samples t-tests conducted to compare scores from respondents to the online and paper version of the pilot survey. Prior to assessing confirmatory models, initial checks were conducted to confirm that items displayed variability across participants within each sample – if only a small number ($n < 5$) participants endorsed an item value, these cases were removed from consideration in the analysis related to the corresponding measure.

Initial model fit to data was established serially using data from the HWR matched sample and online survey datasets using CFA before invariance across these datasets was assessed using Multi-Group Confirmatory Factor Analyses (MGCFA). Initial confirmatory models displaying unacceptable fit were not considered for invariance testing. For models in which items with less than 5 response options were employed as indicators of latent factors, an ordered categorical variable methodology and weighted least squares (WLSMV) estimation, as implemented in MPLUS 7.4, were used. Where subscale scores were employed as indicators (i.e., ELSI-SF), maximum likelihood estimation, as implemented in AMOS 25, was used.

Model fit was assessed against multiple fit indices, namely CFI, Gamma Hat [Gamma hat = $n$vars / (($n$vars) + (2* df *(RMSEA^2))], and McDondalds NCI [Mc = exp(-1/2(($X^2 – df$)/(n-1)))]. Although reported, given the small n of the online sample and small df of many confirmatory models, RMSEA was not considered in assessing model fit as artificially large and small values of RMSEA may result due to sampling error and computational limits of the estimate (i.e. when X^2 is lower than the model df). This limitation was also considered when assessing Gamma Hat values, which are calculated using values of RMSEA. For the remaining fit indices, values $\geq$ .90 were considered to indicate acceptable fit and values $\geq$ .95 to indicate good fit.

Measurement invariance was assessed using nested MGCFA. These models successively assessed whether the factor structure (*Configural invariance*), factor loadings (*Metric invariance*), and item intercepts (*Scalar invariance*) differ significantly across the datasets (Steinmetz, Schmidt, Tina-Booh, Wieczorek, & Schwartz, 2008; Steenkamp, Benedict, & Baumgartner, 1998). Scalar invariance of all indicators is generally required for the comparison of latent and composite (factor) means across groups (Steenkamp et al., 1998; Meredith & Teresi, 2006; Steinmetz, 2013). Residual invariance (invariance of error terms) is often advised against as a criteria for concluding model invariance as it is not often met in practice and as such was not assessed in the current project.

As recommended by Cheung and Rensvold (2002) and Meade, Johnson, and Braddy (2006), nested measurement model invariance was assessed against changes in goodness-of-fit indices ΔCFI, ΔGamma Hat, and ΔMc. These indices are considered to be robust statistics for testing the between-group invariance of CFA models, and provide non-redundant assessments of model fit. Critical values of ΔCFI |0.01|, ΔGamma hat |.001|, and ΔMc |.02| were used to indicate that the null hypothesis of invariance should be rejected. Chi-square difference tests were conducted using the DIFFTEST option in MPLUS and are also reported.

Open data

All data used to assess the invariance of multi-item scales are archived and available on figshare:
https://figshare.com/s/56faaeae4703c0fc8cb0

Consistency of single item ratings and reports

Responses from participants completing both the online and postal pilot surveys were used to assess the consistency with which individuals reported demographic information across modes (i.e. date of birth, annual income, employment status, marital status, highest level of education, ethnic identity). While some variation was expected for some indices, we were interested in whether online and paper survey modes elicited systematic differences in reporting. Participant responses to both surveys were compared using cross-tab tables.

Results

Invariance of the Short-Form Health Survey (SF-12v2)

The SF-12 is a shortened version of the SF-36. The SF-12 is designed to represent the higher-order physical and mental health components of health, rather than to produce scores holistic representation of each of the eight subscales (general health, physical functioning, role physical, vitality, social functioning, mental health and role emotional) assessed by the SF-36. The eight subscales are conceptualised as assessing two orthogonal health factors: three summed subscales load positively on a physical component [physical functioning (2 items), role physical (2 items) and, bodily pain (1 item)] and three load positively on a mental component score [social functioning (1 item), role emotional (2 items) and, mental health (2 items)]. Two subscales correlate moderately with both components [general health (1 item) and vitality (1 item)], although most strongly with physical and mental health, respectively (Ware, Kosinski, & Keller, 1996; Ware, Kosinski, Turner-Bowker, & Gandek, 2002). Items of the SF-12 each have 3-5 response options.

Following the recommended approach of Ware et al. (1998), the two-factor structure of the SF-12 in new samples is typically confirmed using exploratory principal component analyses of the eight summated subscale scores, with several countries adopting the resulting country-specific factor weights and normative scores in the calculation of overall physical and mental component scores. Specifically, the orthogonal factor weights from exploratory analyses are applied to the relevant standardized subscale score (standardized to a given population/sample) to calculate the overall physical and mental component scores. Factor weights and population normative values for subscales have been assessed in New Zealand (Frieling, Davis, & Chiang, 2013; Table 2). In representative Australia and New Zealand samples, general health and vitality items have been observed to load near equally across the orthogonal physical and mental health component scores (Frieling et al., 2013; Bartsch et al., 2011).

Table. Orthogonal factor weights for SF-12 subscale scores in New Zealand population - reproduced from Frieling et al. (2013).

Subscale (composite)	Physical factor weight	Mental factor weight
Physical functioning	0.397	-0.160
Role physical	0.367	-0.097
Bodily pain	0.340	-0.123
General health	0.150	0.110
Social functioning	0.050	0.212
Vitality	0.028	0.257
Role emotional	-0.131	0.390
Mental health	-0.225	0.491

Given the conceptualisation and scoring of the physical and mental component scores of the SF-12, it is not surprising that there is a lack of consensus regarding appropriate confirmatory models of its factor structure. The orthogonal conceptualisation and associated scoring structure are controversial (e.g., Anagnostopoulos, Niakas, & Tountas, 2009; Farivar, Cunningham, & Hays, 2007) and studies using SEM to assess the construct validity of the MOS forms suggested that the proposed second-order factor structure provide a poor fit to the item data (e.g., Güthlin & Walach, 2007; Hann & Reeves, 2008). The table below provides a brief review of some of the structural models of the SF-12 tested in the CFA literature and associated fit indices. Some have modelled the item data and others have modelled the subscale scores. Models of the factors as uncorrelated do not provide a good fit to the data. Models of the items representing two correlated subscale scores have provided adequate fit. Others have modelled the established cross-loading of items across factors. Finally, several studies modelling the two-factor structure of the items have included error correlations for items within the same subscale, which use the same question stem and response scale options.

Table. Items of the SF-12, associated component factor, and confirmatory models tested in the literature

Item	Item wording	Structural CFA models tested in the literature					
		2 uncorrelated factors	2 correlated factors	2 correlated factors (based on subscale scores)	2 correlated factors with 3 items loading on both factors (correlated errors for PF items)	2 correlated factors with correlated error terms within the RP and RE subscale items	2 correlated factors with correlated errors within all subscales
ge1	In general, would you say your health is:	Physical	Physical	Physical	Physical/Mental	Physical	Physical
pfb	(Health limits) Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	Physical	Physical	Physical	Physical	Physical	Physical
pfd	(Health limits) Climbing several flights of stairs	Physical	Physical	Physical	Physical	Physical	Physical
rpb	(Last 4 weeks - physical) Accomplished less than you would like	Physical	Physical	Physical	Physical	Physical	Physical
rpc	(Last 4 weeks - physical) Were limited in the kind of work or other activities	Physical	Physical	Physical	Physical	Physical	Physical
reb	(Last 4 weeks - emotional) Accomplished less than you would like	Mental	Mental	Mental	Mental	Mental	Mental
rec	(Last 4 weeks - emotional) Did work or activities less carefully than usual	Mental	Mental	Mental	Mental	Mental	Mental
bpb	During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?	Physical	Physical	Physical	Physical	Physical	Physical
tmhc	Have you felt calm and peaceful?	Mental	Mental	Mental	Mental	Mental	Mental
mhd	Have you felt downhearted and depressed?	Mental	Mental	Mental	Mental	Mental	Mental
vtb	Did you have a lot of energy?	Mental	Mental	Mental	Physical/Mental	Mental	Mental
sfb	(Last 4 weeks - social) how much of the time has your physical health or emotional problems interfered with your social activities	Mental	Mental	Mental	Physical/Mental	Mental	Mental
	References and reported fit indices	Jakobsson, Westergren, Lindskov, and Hagell (2012); *n* = 4278, GFI = .70, NFI = .865, CFI = .866, RMSEA = .171 (.214-.221).	Christensen, Ehlers, Larsen, and Jensen (2013): n = 26,397, GFI = .945, CFI = .939, SRMA = .0495, RMSEA = .08	Montazeri et al. (2011): n = 3685, GFI = 0.93, AGFI = 0.87, RMSE = 0.10, (0.10 - 0.11), NFI = 0.96, and CFI = 0.96. Montazeri, Vahdaninia, Mousavi, and Omidvari (2009): n = 5587, GFI= 0.96, AGFI = 0.93, RMSE = 0.090, 95% CI RMSE = 0.085- 0.095, NFI = 0.93, and CFI = 0.93. Fleishman, Selim, and Kazis (2010): n = 53,399 CFI = 0.952, TLI = 0.922, and RMSEA = 0.078. SRMSR = 0.037.	Okonkwo, Roth, Pulley, and Howard (2010): N = 14640; X2 = 4,625.76, df = 32, P <0.0001, CFI = 0.976, RMSEA = 0.059 Fisher and Newbold (2014): X2 = 294.46, CFI .94, SRMR .06, RMSEA .06.	Damásio, Andrade, and Koller (2015): X2 187.483(51), CFI = .937, TLI = .918, RMSEA = 0.66 (.056-.076)	Wilson, Tucker, and Chittleborough (2002): n = 3007, RMSEA = 0.07 (no other fit stats reported) Tucker, Adams, and Wilson (2010): n = 17,479, X2 = 2919.8 df = 49, RMSEA = 0.058 (0.056 - 0.060), p < 0.000, NNFI = 0.9679, CFI = 0.9762.

CFA and MGCFA

As two SF-12 items have only 3 response options, analyses were conducted using MPLUS. The fit of items to the two-factor correlated model was assessed using CFA in both the HWR matched sample and online respondent sample. Model fit was poor (table below) and modification indices indicated acceptable model fit could only be obtained by including the established cross-loadings of items. As item cross-loadings cannot be modelled in categorical MGCFA as implemented by MPLUS, the physical and mental models were assessed separately.

SF-12					**Model fit indices**							
		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI
CFA												
2 correlated factors	HWR	504	468.034	53	0.88	.	0.66	.	0.948	.	0.125	.114-.135
2 correlated factors	Online pilot	252	291.837	53	0.86	.	0.62	.	0.963	.	0.134	.119-.149

The figures below present the standardized regression weights for the unconstrained models by sample and survey mode.

Figure. Standardized weights for initial unconstrained models by sample and survey mode.

Physical: HWR matched sample   Physical: Online pilot sample

Mental: HWR matched sample   Mental: Online pilot sample
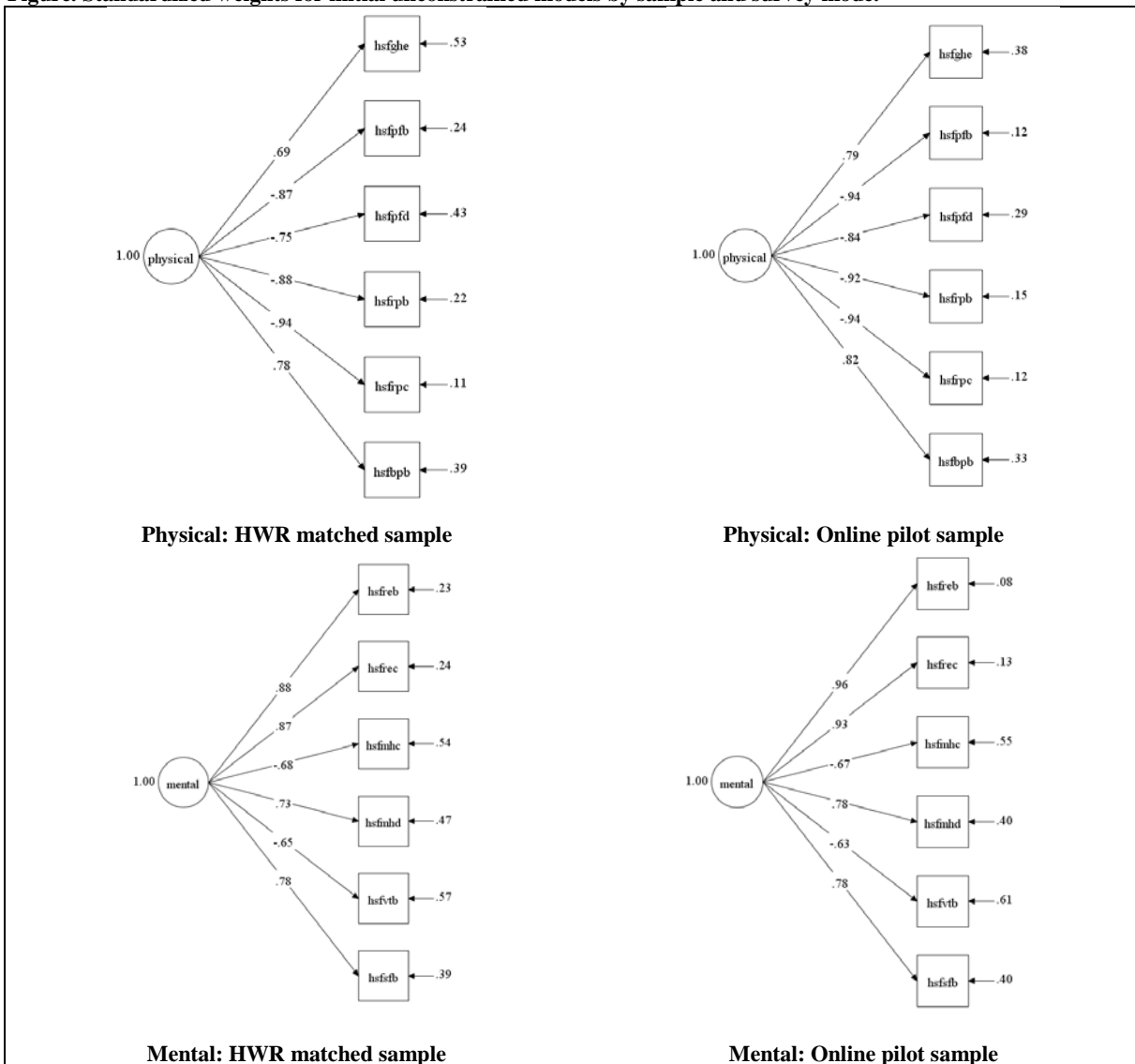
The table below presents the fit indices for the single factor CFA and MGCFA models. Assessment of initial model fit within the matched HWR sample and online survey sample indicated that the physical and mental health models displayed adequate fit in both survey modes on indices of Gamma Hat, Mc and CFI. In inspecting response frequencies prior to conducting MGCFA, it was observed that no values of item MHd 1 = 'all of the time' or item MHc 5 = 'none of the time' were observed in the online survey responses. N = 6 cases in the matched HWR sample who endorsed these responses were excluded from the MGCFA. As such, the thresholds associated with these low frequency responses were not assessed in the current multi-group models of the mental health factor. Multi-group assessment of nested model invariance indicated that the models displayed scalar invariance in respect to change in Mc and CFI. Change in Gamma Hat indicated that models did not display metric invariance.

SF-12													Model fit indices
Physical Factor		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	X^2Δ test
CFA													
6-item Physical factor	HWR	504	83.995	9	0.952	.	0.928	.	0.985	.	0.129	.104-.154	
6-item Physical factor	Online	252	29.899	9	0.973	.	0.959	.	0.996	.	0.096	.059-.135	
MGCFA													
	Configural	748	116.849	21	0.922	.	0.938	.	0.990	.	0.110	.091-.130	
	Metric	748	108.164	26	0.932	-0.010	0.946	-0.009	0.992	-0.002	0.092	.074-.110	Not computed
	Scalar	748	114.366	36	0.941	.	0.949	-0.002	0.992	0.000	0.072	.061-.092	$X^2(10) = 13.833, p = .187$
Mental Factor		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	X^2Δ test
CFA													
6-item Mental factor	HWR	504	118.611	9	0.933	.	0.897	.	0.966	.	0.155	.131-.181	
6-item Mental factor	Online	252	47.256	9	0.952	.	0.927	.	0.988	.	0.130	.095-.169	
MGCFA													
	Configural	748	161.304	18	0.887	.	0.909	.	0.977	.	0.146	.126-.167	
	Metric	748	139.654	23	0.906	-0.020	0.925	-0.016	0.982	-0.005	0.116	.098-.135	$X^2(5) = 4.844, p = .435$
	Scalar	748	150.22	37	0.909	.	0.927	-0.002	0.982	0.000	0.090	.076-.106	$X^2(14) = 24.078, p = .045$

Note critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|; red text highlights values indicating inadequate differential model fit.

Invariance of the Center for Epidemiological Studies Depression scale (CESD-10)

The 10 item CESD-10 was derived from the 20-item version, which was designed to represent four factors: Depressed affect, Positive affect, Somatic and retarded activity, and Interpersonal issues (Radloff, 1977; Shafer, 2006). Responses were provided on a four-point scale ranging from 0 (none of the time) to 3 (all of the time). The 10 item brief version has been represented by 2 factors (general depressed affect and positive affect) in older adult populations (Lee & Chokkanathan, 2008) and as a single factor with correlated error variance for the positive affect items (González et al., 2017; Mohebbi et al., 2017) and additionally for two items associated with energy (Mohebbi et al., 2017).

Table. Items of the CESD-10, associated factors, and confirmatory models tested in the literature

|			Structural CFA models tested in the literature		
Variable	Question	Factor in 20-item version	2 correlated factor model (depressed affect and positive affect)	Single factor with correlated errors for positive affect items (González et al., 2017)	Single factor with correlated errors for positive affect items and 'energy' items (Mohebbi et al., 2017)
tHCESDa	I was bothered by things that usually don't bother me	Somatic and retarded	General depression	General depression	General depression
tHCESDb	I had trouble keeping my mind on what I was doing	Somatic and retarded	General depression	General depression	General depression
tHCESDc	I felt depressed	Depressed affect	General depression	General depression	General depression
tHCESDd	I felt that everything I did was an effort	Somatic and retarded	General depression	General depression	Energy
tHCESDe	I felt hopeful about the future	Positive affect (R)	Positive affect (R)	Positive affect (R)	Positive affect (R)
tHCESDf	I felt fearful	Somatic and retarded	General depression	General depression	General depression
tHCESDg	My sleep was restless	Somatic and retarded	General depression	General depression	General depression
tHCESDh	I was happy	Positive affect (R)	Positive affect (R)	Positive affect (R)	Positive affect (R)
tHCESDi	I felt lonely	Depressed affect	General depression	General depression	General depression
tHCESDj	I could not "get going"	Somatic and retarded	General depression	General depression	Energy

CFA and MGCFA

As CESD items have only 4 response options, analyses were conducted using MPLUS The table below presents CFA and MGCFA models assessed. Results suggest a single factor model of the CESD provided adequate fit to the HWR matched sample across two indices of model fit. The addition of the previously observed correlation between the two positively-worded items, assessing positive affect, provided good model fit across all indices for both survey modes.

CESD-10					Model fit indices								
		N	X²	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
1 factor	HWR	504	217.41	35	0.932	.	0.834	.	0.943	.	0.102	.089-.115	
+ e5 <-> e8	HWR	504	110.17	34	0.970	.	0.927	.	0.976	.	0.067	.053-.081	
+ e5 <-> e8	Online	252	84.42	34	0.961	.	0.904	.	0.972		0.077	.056-.097	
MGCFA													
	Configural	756	201.93	78	0.938		0.921		0.975		0.065	.054-.076	
	Metric	756	197.04	87	0.945	0.007	0.930	0.009	0.978	0.003	0.058	.047-.069	X²(9) = 19.099, p = 0.024
	Scalar	756	217.85	106	0.944	.	0.929	-0.001	0.978	0.000	0.053	.043-.063	.

Note critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

The figure presents the standardized regression weights for the unconstrained models by survey mode. Multi-group assessment of invariance of this model across the HWR matched and online pilot samples indicated that the model displayed scalar invariance across survey modes in relation to change in Mc, CFI. Change in Gamma Hat indicated that the model did not display metric invariance.

Figure. Standardized weights for initial unconstrained models by sample and survey mode.

| HWR matched sample | Online pilot sample |
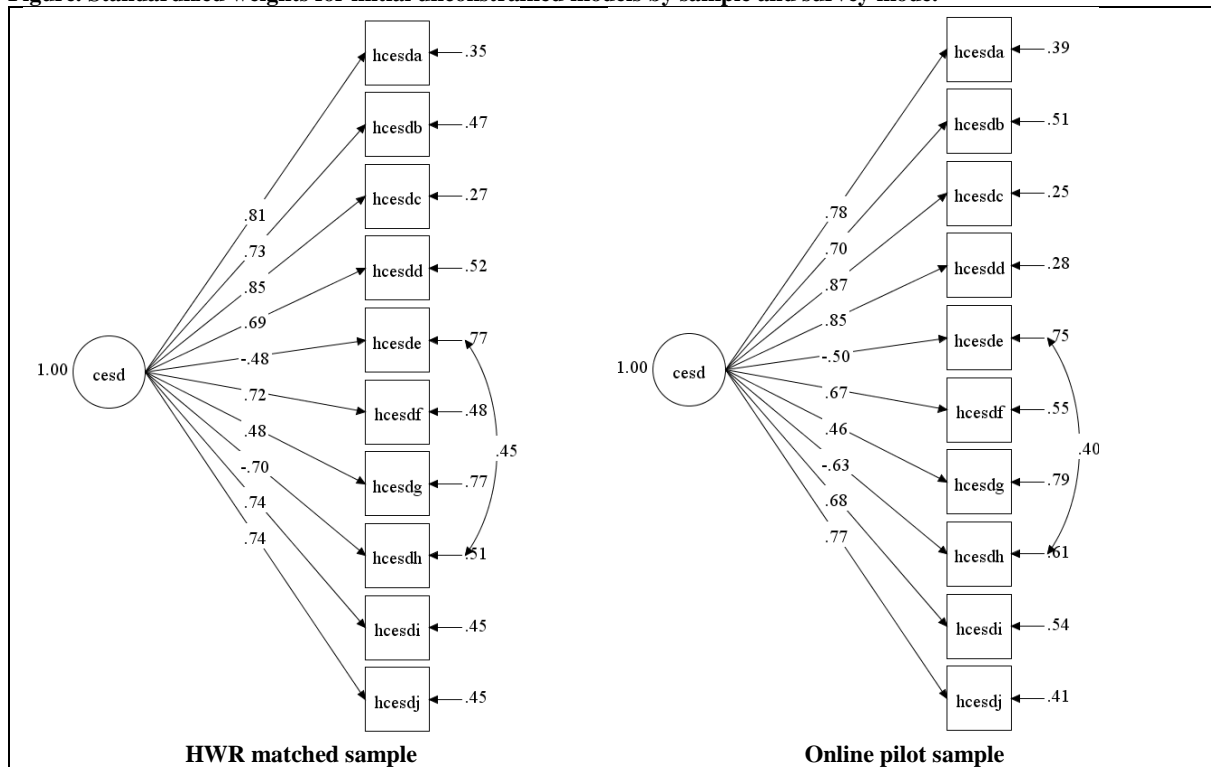
## Invariance of the De Jong Gierveld and Tilburg scale for loneliness

This is a well-established measure of loneliness, which is conceptualised as two distinct but correlated factors - social and emotional loneliness. The sample size was considered inadequate to generate robust initial CFA model estimates for each group and results are reported here for descriptive purposes.

Variable	Question	Response range	Factor
FLONEc	I experience a general sense of emptiness	1 - 3 (Yes - No)	Emotional
FLONEd	There are plenty of people I can rely on when I have problems	1 - 3 (Yes - No)	Social (R)
FLONEg	There are many people I can trust completely	1 - 3 (Yes - No)	Social (R)
FLONEh	There are enough people I feel close to	1 - 3 (Yes - No)	Social (R)
FLONEi	I miss having people around	1 - 3 (Yes - No)	Emotional
FLONEj	I often feel rejected	1 - 3 (Yes - No)	Emotional

CFA and MGCFA

As loneliness scale items have only 3 response options, analyses were conducted using MPLUS. The table below presents the CFA and MGCFA models assessed. Results suggest that the two-factor model of the loneliness scale provided good fit to the HWR matched and online pilot sample data.

Figure. Standardized weights for initial unconstrained models by sample and survey mode.

HWR matched sample   **Online pilot sample**
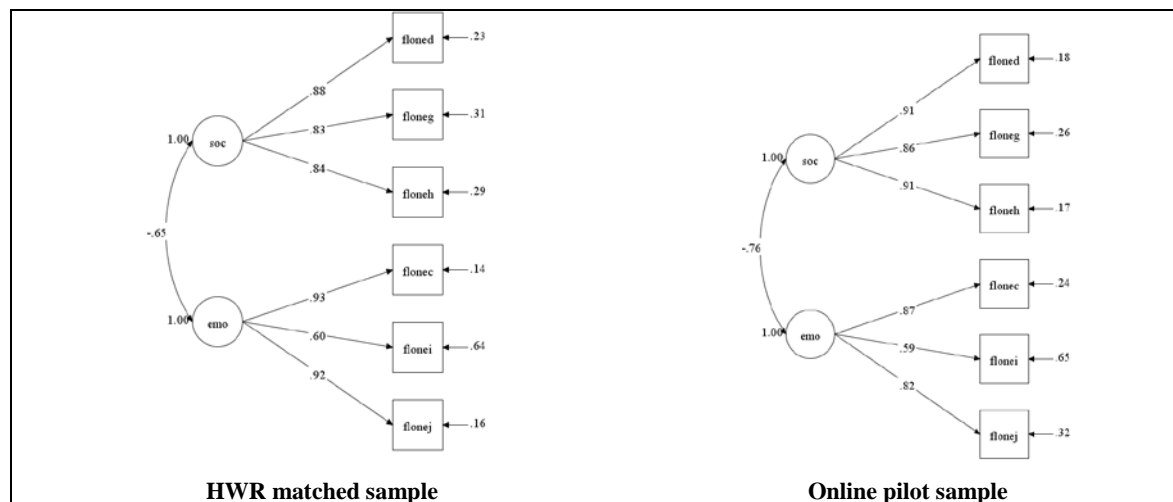
Loneliness				Model fit indices									
		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
2 factors	HWR	504	24.229	8	0.990	.	0.984	.	0.990	.	0.063	.035-.093	
2 factors	Online	252	5.244	8	1.000	.	1.006	.	1.000		0.000	.000-.054	
MGCFA													
	Configural invariance	756	31.701	16	0.986	.	0.990	.	0.994	.	0.051	.024-.077	.
	Metric invariance	756	28.161	20	0.993	0.006	0.995	0.005	0.997	0.003	0.033	.000-.059	$X^2(4) = 0.972$, p = .914
	Scalar invariance	756	36.972	24	0.989	.	0.991	-0.003	0.995	-0.002	0.038	.006-.061	$X^2(4) = 9.443$, p = .051

*critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

Multi group assessment of invariance of this model across the HWR matched and online pilot samples suggests that the measurement model displayed metric and scalar invariance across survey modes in relation to change in McDonald's NCI, CFI. The model did not display metric invariance in relation to change in Gamma Hat.

Invariance of the Social Provisions Scale

The Social Provisions Scale is a 24 item scale with each item rated on the same four response options. These items are designed to assess 6 distinct aspects of social provisions, each indicated by 4 items: Reliable Alliance, Attachment, Guidance, Opportunity for Nurturance, Social Integration, and Reassurance of Worth. No exploratory or confirmatory models of its factor structure were identified in the literature. The 6 factor structure was not evidence in exploratory analyses of the full 2016 HWR survey dataset ($n = 4028$). As such, the current models confirm the factor structure of each of subscales and assess their invariance across the HWR postal survey and online pilot survey datasets.

Item	Variable	Item wording
1	SSPRAa16	There are people I can depend on to help me if I really need it
2	SSPATa16	I feel that I do not have close personal relationships with other people (R)
3	SSPGDa16	There is no one I can turn to for guidance in times of stress (R)
4	SSPONa16	There are people who depend on me for help
5	SSPSIa16	There are people who enjoy the same social activities I do
6	SSPRWa16	Other people do not view me as competent (R)
7	SSPONb16	I feel personally responsible for the well-being of another person
8	SSPSIb16	I feel part of a group of people who share my attitudes and beliefs
9	SSPRWb16	I do not think other people respect my skills and abilities (R)
10	SSPRAb16	If something went wrong, no one would come to my assistance
11	SSPATb16	I have close relationships that provide me with a sense of emotional security and well-being
12	SSPGDb16	There is someone I could talk to about important decisions in my life
13	SSPRWc16	I have relationships where my competence and skills are recognized
14	SSPSIc16	There is no one who shares my interests and concerns (R)
15	SSPONc16	There is no one who really relies on me for their well-being (R)
16	SSPGDc16	There is a trustworthy person I could turn to for advice if I were having problems
17	SSPATc16	I feel a strong emotional bond with at least one other person
18	SSPRAc16	There is no one I can depend on for aid if I really need it (R)
19	SSPGDd16	There is no one I feel comfortable talking about problems with (R)
20	SSPRWd16	There are people who admire my talents and abilities
21	SSPATd16	I lack a feeling of intimacy with another person (R)
22	SSPSId16	There is no one who likes to do the things I do (R)
23	SSPRAd16	There are people I can count on in an emergency
24	SSPONd16	No one needs me to care for them (R)

CFA and MGCFA

As Social Provisions scale items have only43 response options, analyses were conducted using MPLUS. An example of a single factor 4-indicator model applied to each of the SPS subscales is presented below.

Figure. Example single factor model for SPS subscale - standardized weights for initial unconstrained models by sample and survey mode for Reassurance of Worth.
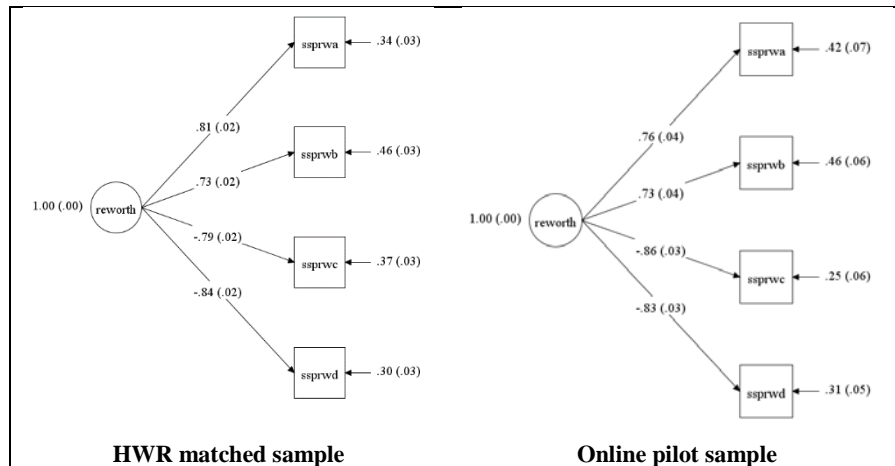
Attachment

Initial model fit was good across all indicators in both the HWR and online pilot samples. Assessment of model invariance indicated that the Attachment scale displayed scalar invariance with reference to change in Mc and CFI, however it failed to display metric invariance with reference to Gamma Hat and failed to display scalar invariance on the Chi-square difference test.

Attachment					**Model fit indices**								
		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
1 factor	HWR	504	8.674	2	0.993	.	0.993	.	0.996	.	0.081	.032-.140	
1 factor	Online	252	6.11	2	0.992	.	0.992	.	0.995	.	0.090	.009-.176	
MGCFA													
	Configural	756	14.893	4	0.986	.	0.993	.	0.996	.	0.085	.042-.133	
	Metric	756	20.2	7	0.983	-0.003	0.991	-0.002	0.995	-0.001	0.071	.036-.108	X2(3) = 7.605, p = .055
	Scalar	756	36.463	13	0.970	.	0.985	-0.007	0.991	-0.004	0.069	.043-.096	X2(6) = 17.381, p = .008

*critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

Social Integration

Initial model fit was good across all indicators in both the HWR and online pilot samples. Assessment of model invariance indicated that the Social Integration scale displayed scalar invariance with reference to change in Mc and CFI, however it failed to display metric invariance with reference to Gamma Hat and failed to display scalar invariance on the Chi-square difference test.

Social Integration		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
								Model fit indices					
CFA													
1 factor	HWR	504	21.34	2	0.981	.	0.981	.	0.994	.	0.139	.089-.194	
1 factor	Online pilot	252	4.492	2	0.995	.	0.995	.	0.998	.	0.070	.000-.159	
MGCFA													
	Configural	756	31.201	4	0.965	.	0.982	.	0.994	.	0.134	.093-.180	
	Metric	756	30.100	7	0.971	0.005	0.985	0.003	0.995	0.001	0.093	.061-.129	X2(3) = 6.001, p = .112
	Scalar	756	57.583	13	0.945	.	0.971	-0.014	0.989	-0.006	0.095	.071-.121	X^2(6) = 29.217, p < .000

*critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

Opportunity for Nurturance

Initial model fit was good across all indicators in both the HWR and online pilot samples. Assessment of model invariance indicated that the Opportunity for Nurturance scale displayed scalar invariance with reference to change in Mc and CFI, however it failed to display metric invariance with reference to Gamma Hat and failed to display scalar invariance on the Chi-square difference test.

Opportunity for Nurturance		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
								Model fit indices					
CFA													
1 factor	HWR	504	26.99	2	0.976	.	0.975	.	0.985	.	0.157	.108-.213	
1 factor	Online	252	1.907	2	1.000	.	1.000	.	1.000	.	0.000	.000-.123	
MGCFA													
	Configural	756	30.822	4	0.966	.	0.982	.	0.993	.	0.133	.092-.179	
	Metric	756	28.436	7	0.972	0.007	0.986	0.004	0.994	0.001	0.090	.057-.126	X2(3) = 5.652, p = .130
	Scalar	756	60.398	13	0.941	.	0.969	-0.017	0.987	-0.007	0.098	.074-.124	X2(6) = 32.096, p < .001

*critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

Reliable Alliance

While *n* = 3 cases in the online pilot indicated that they 'strongly disagree' with the statement '*There are people I can depend on to help me if I really need it*' (SSPRAa), no cases nominated this option in the HWR matched sample. These three cases were excluded from the analysis and the model re-specified, such that this variable only had two response thresholds. Initial model fit was good across all indicators in both the HWR and online pilot samples. Assessment of model invariance indicated that the Reliable Alliance scale displayed metric invariance with reference to change in Gamma hat, Mc and CFI and Chi-square difference test, however it failed to display scalar invariance in relation to these indices.

Reliable Alliance					Model fit indices								
		N	X²	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
1 factor	HWR	504	3.27	2	0.999	.	0.999	.	1.000		0.035	.000-.102	
1 factor	Online	252	5.84	2	0.992	.	0.992	.	0.998		0.087	.000-.174	
MGCFA													
	Configural	753	7.025	4	0.996	.	0.998	.	0.999		0.045	.000-.098	
	Metric	753	4.714	7	1.000	0.004	1.002	0.004	1.000	0.001	0.000	.000-.049	X2(3) = 0.292, p = .09616
	Scalar	753	124.234	12	0.870	-0.130	0.928	-0.073	0.978	-0.022	0.158	.133-.183	X2(5) = 110.060, p < .001

*critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

Guidance

Initial model fit was good across all indicators in both the HWR and online pilot samples. Assessment of model invariance indicated that the Guidance scale displayed scalar invariance with reference to change in Mc and CFI, however it failed to display metric invariance with reference to Gamma Hat and the Chi-square difference test.

Guidance					Model fit indices								
		N	X²	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
1 factor	HWR	504	26.548	2	0.976	.	0.976	.	0.994		0.156	.107-.211	
1 factor	Online	252	2.678	2	0.999	.	0.999	.	1.000		0.037	.000-.136	
MGCFA													
	Configural	756	28.477	4	0.969	.	0.984	.	0.996	.	0.127	.086-.173	
	Metric	756	25.224	7	0.976	0.008	0.988	0.004	0.997	0.001	0.083	.050-.119	X2(3) = 8.223, p = .042
	Scalar	756	46.802	13	0.957	.	0.978	-0.010	0.994	-0.003	0.083	.058-.109	.

*critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

Reassurance of Worth

Initial model fit was good across all indicators in both the HWR and online pilot samples. Assessment of model invariance indicated that the Reassurance of Worth scale displayed scalar invariance with reference to change in Mc and CFI, however it failed to display metric invariance with reference to Gamma Hat and the Chi-square difference test.

Reassurance of Worth		N	X^2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
1 factor	HWR	504	42.192	2	0.962	.	0.961	.	0.984	.	0.200	.150-.254	
1 factor	Online pilot	252	88.427	2	0.946	.	0.842	.	0.959	.	0.239	.198-.283	
MGCFA													
	Configural	751	71.289	4	0.918	.	0.956	.	0.981	.	0.212	.170-.256	
	Metric	751	55.165	7	0.940	0.023	0.968	0.012	0.987	0.006	0.135	.103-.170	$X^2(3) = 8.49$, p = .037
	Scalar	751	58.754	13	0.942	.	0.970	0.002	0.987	0.000	0.097	.072-.123	.

*critical values: ΔGamma hat |.001|; ΔMcDonald's NCI |.02|; ΔCFI |0.01|

CASP-12 (NZCASP-11)

The CASP-12 developed as a non-health-related assessment of quality of life for older persons. Originally developed as the CASP-19 (Hyde, Wiggins, Higgs, & Blane, 2003) with separate subscales of Control, Autonomy, Self-realisation and Pleasure, with consideration to item functioning and perceived lack of conceptual divergence between the Control and Autonomy items, the measure and associated model was simplified to the CASP-12 (Wiggins, Netuveli, Hyde, Higgs, & Blane, 2008). Composite scoring of the CASP-12 provides three factor scores (Control & Autonomy, Self-realisation and Pleasure). Several studies have attempted to define the factor studier of the CASP12 items through exploratory an confirmatory methods (Kim et al., 2015; Sexton, King-Kallimanis, Conroy, & Hickey, 2013; Pérez-Rojo, Martín, Noriega, & López, 2017; Towers, Yeung, Stevenson, Stephens, & Alpass, 2015), however none have provided a definitive or replicable model. Using data from the 2012 wave of the New Zealand Health, Work and Retirement study, Towers et al. (2015) determined an 11-item, 3-factor solution with two correlated error terms provided acceptable fit to the data. The measure was called the NZCASP-11. Given the high level of relevance to the current New Zealand population and study mode, this model was used to assess the invariance of the CASP across paper and online administrations for older persons in New Zealand.

Table. Items of the CASP12 and associated subscales

Variable	Question	Range	Original factor	NZ CASP 11
tHCASPa16	My age prevents me from doing the things I would like to (R)	1 – 4 (never - often)	Control	Control
tHCASPb16	I feel that what happens to me is out of my control (R)	1 – 4 (never - often)	Control	Control
tHCASPc16	I feel left out of things (R)	1 – 4 (never - often)	Control	Control
HCASPd16	I can do the things that I want to do	1 – 4 (often – never)	Autonomy	Independence
HCASPe16	I feel that I can please myself what I do	1 – 4 (often – never)	Autonomy	Independence
tHCASPf16	Shortage of money stops me from doing things I want to do (R)	1 – 4 (never - often)	Autonomy	(Dropped)
HCASPg16	I look forward to each day	1 – 4 (often – never)	Pleasure	Independence
HCASPh16	I feel that my life has meaning	1 – 4 (often – never)	Pleasure	Independence
HCASPi16	I enjoy the things that I do	1 – 4 (often – never)	Pleasure	Independence
HCASPj16	I feel full of energy these days	1 – 4 (often – never)	Self-realisation	CASP-3
HCASPk16	I feel that life is full of opportunities	1 – 4 (often – never)	Self-realisation	CASP-3
HCASPl16	I feel that the future looks good for me	1 – 4 (often – never)	Self-realisation	CASP-3

CFA and MGCFA

Confirmatory analyses for the NZCASP-11 were conducted in AMOS. The NZCASP-11 model, originally established in the 2012 HWR survey factor, displayed good fit to the data in both groups. MGCFA indicated that the measure displayed scalar invariance across all indices.

Figure. Standardized weights for initial unconstrained models by sample and survey mode.

HWR matched sample   Online pilot sample
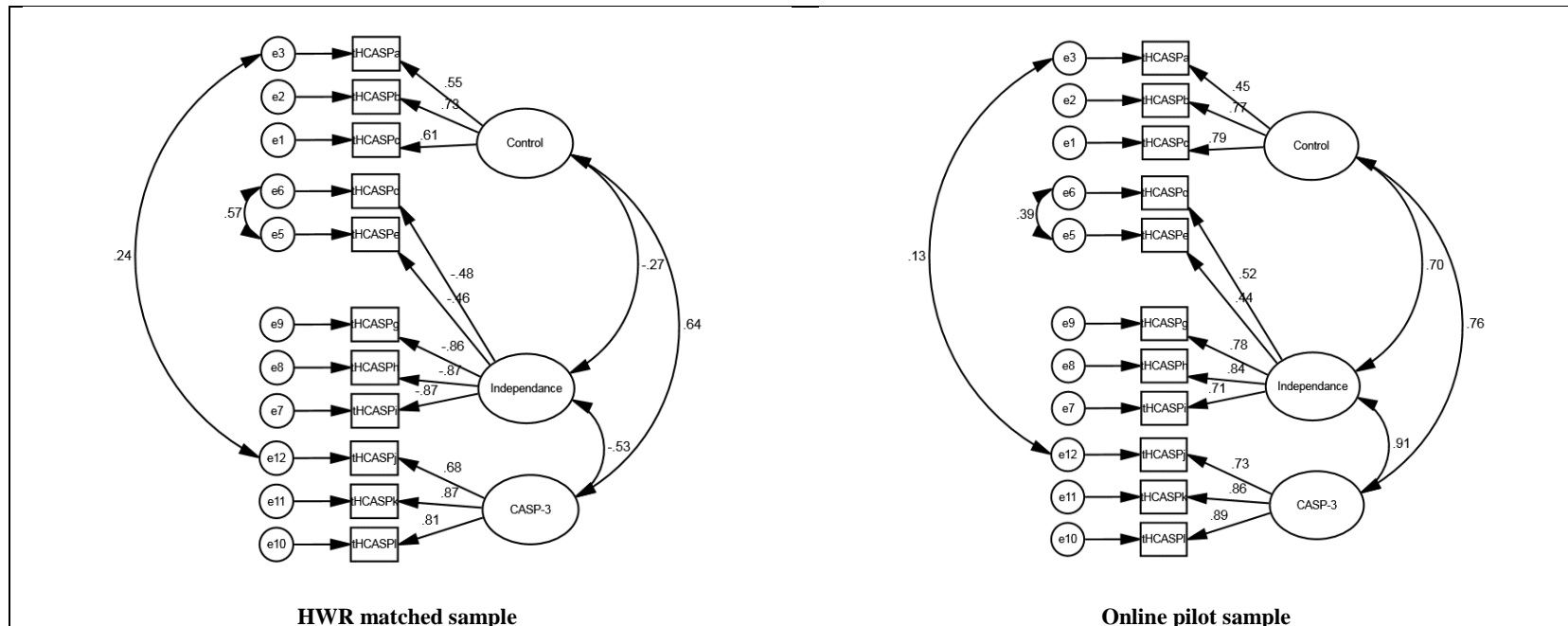
NZCASP-11		N	X²	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
	HWR	504	74.967	39	0.987	.	0.965	.	0.986	.	0.043	.028-.057	.
	Online pilot	252	80.62	39	0.971	.	0.920	.	0.969	.	0.065	.045-.086	.
MGCFA													
	Configural	756	155.643	78	0.982	.	0.950	.	0.980	.	0.036	.028-.045	.
	Metric	756	161.274	86	0.982	0.000	0.951	0.001	0.980	0.000	0.034	.026-.042	X2(8) = 5.632, p = .688
	Scalar	756	178.252	97	0.981	-0.001	0.948	-0.004	0.979	-0.001	0.033	.026-.041	X2(19) = 22.61, p = .255

*critical values: ΔGamma hat |.001|; ΔMc |.02|; ΔCFI |0.01|

Invariance of the Economic Living Standards Index short form (ELSI-SF)

The ELSI-SF (Jensen et al., 2002) comprises 25 questions which assess four economic domains: restrictions in the ownership of household items; restrictions in social participation; the extent to which respondents economised to keep living costs down, and; self-rated satisfaction with living standards. Items were re-coded as per the scoring instructions and summary scores calculated. As evidenced by high mean scores, several items displayed no variation in the HWR matched and pilot survey samples (i.e., telephone ownership). The current report adopts the approach to CFA of the original ELSI (Jensen, Spittal, Crichton, Sathiyandra, & Krishnan, 2002), in which the summered subscale scores were modelled as indicators of an overall ELSI dimension.

Item	Variable
Ownership restrictions (item range 0-1)	
tDELOWa	Telephone
tDELOWb	Washing machine
tDELOWc	Keep the main rooms of your home adequately heated
tDELOWd	At least two pair of good shoes
tDELOWe	Suitable clothes for important or special occasions
tDELOWf	Personal computer
tDELOWg	Home contents insurance
Social participation (item range 0-1)	
tDELSPd	Give presents to family/whanau or friends on birthdays, Christmas or other special occasions
tDELSPa	Visit the hairdresser at least once every three months
tDELSPb	Have holidays away from home for at least a week every year
tDELSPc	Enough room for family/whanau to stay the night
tDELSPe	Have a holiday overseas at least every three years
tDELSPf	Have a night out for entertainment or socialising at least once a fortnight
tDELSPg	Have family/whanau or friends over for a meal at least once every few months
Economising (item range 0-2)	
tDELECa	Gone without or cut back on fresh fruit and vegetables to help keep down costs
tDELECb	Continued wearing clothing that was worn out because you couldn't afford a replacement
tDELECc	Put off buying clothes for as long as possible to help keep down costs
tDELECd	Stayed in bed longer to save on heating costs
tDELECe	Postponed or put off visits to the doctor to help keep down costs
tDELECf	NOT picked up a prescription to help keep down costs
tDELECg	Spent less time on hobbies than you would like to help keep down costs
tDELECh	Gone without or cut back on trips to the shops or other local places to help keep down costs
Self-rated living standard (item range 0-4^, and *0-3)	
tDELMSa^	Generally, how would you rate your material standard of living?
tDELMSb^	Generally, how satisfied are you with your current material standard of living?
tDELMSc*	How well does your total income meet your everyday needs for such things as accommodation, food, clothing and other necessities?

CFA and MGCFA

As the ELSI-SF subscale scores are continuous variables, analyses were conducted in AMOS. The single factor model of the ELSI-SF subscale scores displayed good fit to the data in both groups. MGCFA indicated that the measure displayed scalar invariance with reference to ΔMc and ΔCFI however, it did not display metric invariance with reference to Δ Gamma hat and the Chi-square difference test.

Figure. Standardized weights for initial unconstrained models by sample and survey mode.

HWR matched sample   Online pilot sample
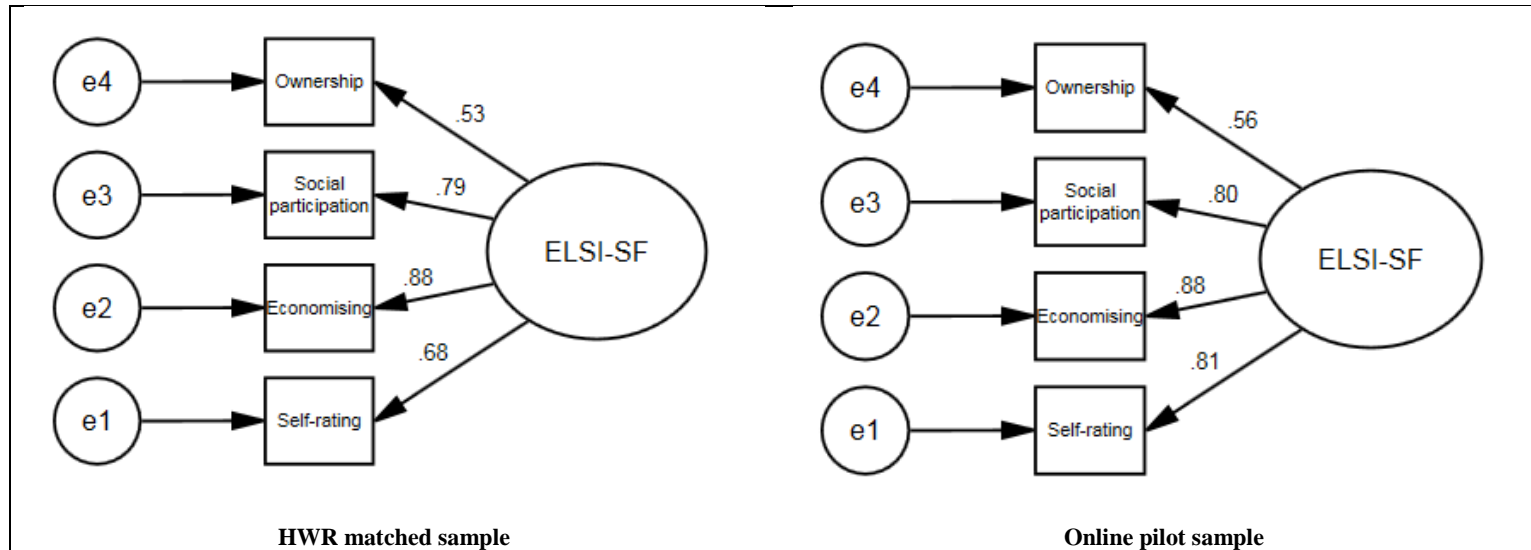
ELSI-SF		**Model fit indices**											
		N	X²	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
1 factor	HWR	504	5.616	2	0.996		0.996		0.995		0.060	.000 - .121	
1 factor	Online pilot	252	1.871	2	1.000		1.000		1.000		0.000	.000 - .123	
MGCFA													
	Configural invariance	756	7.485	4	0.998		0.998		0.997		0.034	.000-.071	
	Metric invariance	756	21.28	7	0.991	-0.007	0.991	-0.007	0.988	-0.009	0.052	.028-.078	X²(3) = 13.80, p = .003
	Scalar invariance	756	30.49	11	0.987	.	0.987	-0.003	0.983	-0.005	0.048	.028-.069	.

*critical values: ΔGamma hat |.001|; ΔMc |.02|; ΔCFI |0.01|

Invariance of the Living standards short-form (LSCAPE-6)

The LSCAPE-6 (Breheny et al., 2014) is a single-factor measure of economic circumstances of older persons, conceptualised as varying from constraint to freedom. The six items of the measure are each rated on a scale of 1-5. One item was reverse coded.

Variable	Item
DELELDERp	I can afford to go to a medical specialist if I need to
DELELDERw	I am able to visit people whenever I wish
DELELDERm	I am able to give to others as much as I want
DELELDERe	I am able to do all the things I love
DELELDERt	I expect a future without money problems
DELELDERae	My choices are limited by money (R)

CFA and MGCFA

As the LSCAPE items are each rated in a scale of 1-5, analyses were conducted in AMOS. Inspection of standardized weights in the HWR matched and online pilot sample datasets indicate that all items loaded strongly on the latent factor, although the reverse-coded item '*my choices are limited by money*' displayed the weakest loading.

Figure. Standardized weights for initial unconstrained models by sample and survey mode.

HWR matched sample   Online pilot sample
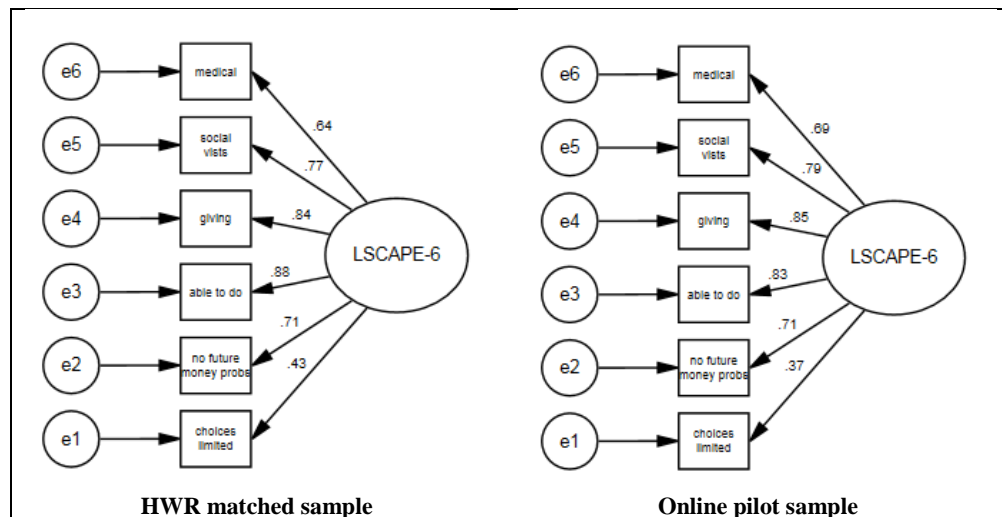
CFA indicated that the single factor model of the LSCAPE-6 items displayed acceptable fit to the data in both groups.

LSCAPE-6		N	X2	df	Gamma Hat	Δ	Mc	Δ	CFI	Δ	RMSEA	90% CI	DIFFTEST
CFA													
1 factor	HWR	504	68.1	9	0.962	.	0.943	.	0.958	.	0.114	.090-.140	
1 factor	Online pilot	252	32.45	9	0.970	.	0.954	.	0.966	.	0.102	.066-.141	
MGCFA													
	Configural	756	100.5	18	0.965	.	0.947	.	0.961	.	0.078	.064-.093	.
	Metric	756	105.8	23	0.965	0.000	0.947	0.000	0.961	0.000	0.069	.056-.083	X2(5) = 5.27, p = .384
	Scalar	756	118.2	29	0.962	-0.003	0.943	-0.004	0.958	-0.003	0.064	.052-.076	X2(11) = 17.684, p = .089

Model fit indices

*critical values: ΔGamma hat |.001|; ΔMc |.02|; ΔCFI |0.01|

MGCFA indicated that the measure displayed scalar invariance by ΔMc, ΔCFI and chi square difference test, however it only displayed metric invariance with reference to Δ Gamma hat.

Paired comparison of summated scores across modes

Paired samples t-tests generally indicated no difference across survey modes. Differences of small effect size were observed between the online and paper survey administrations for two measures of mental health – the Mental Component Score of the SF-12 (dz = 0.19) and the Centre for Epidemiology Scale for Depression (dz = 0.20) – such that respondents reported poorer mental health in the online form. However, in light of the sample and effect size, this comparison provided insufficient evidence to reject the null hypothesis of no difference between the means.

Table. Comparison of mean total summary scores utilised in the HWR study by survey mode for participants returning both the online and paper survey versions.

	Online			Paper			Paired samples t-test			
	N	Mean	SD	N	M	SD	Mean diff	t	df	p
Social Loneliness	122	1.04	1.19	122	0.93	1.12	0.11	1.36	121	0.175
Emotional Loneliness	122	0.67	0.93	122	0.59	0.90	0.08	1.17	121	0.245
Loneliness Total Score	122	1.71	1.77	122	1.52	1.71	0.19	1.62	121	0.109
Attachment	122	13.40	2.41	122	13.42	2.38	-0.02	-0.10	121	0.918
Social Integration	122	13.30	2.15	122	13.56	2.45	-0.25	-1.71	121	0.089
Opportunity for Nurturance	122	11.97	2.83	122	11.90	2.88	0.07	0.40	121	0.688
Reassurance of Worth	122	13.46	1.87	122	13.36	2.16	0.10	0.68	121	0.496
Reliable Alliance	122	14.21	2.10	122	14.22	2.08	-0.01	-0.06	121	0.952
Guidance	122	13.82	2.29	122	13.89	2.28	-0.07	-0.43	121	0.670
Social Provisions Total Score	122	80.16	10.87	122	80.34	11.33	-0.18	-0.32	121	0.750
ELSI Short Form Score	122	24.92	6.34	122	24.70	6.36	0.21	1.01	121	0.315
ELSI Short Form Category Score	122	5.64	1.56	122	5.53	1.52	0.11	1.77	121	0.080
SF-12 Physical Component Score	122	47.01	10.51	122	46.90	10.48	0.11	0.28	121	0.783
SF-12 Mental Component Score	122	49.71	9.55	122	50.99	9.97	-1.27	-2.07	121	0.041
CES-D total score	122	6.39	4.43	122	5.70	4.90	0.69	2.17	121	0.032
CASP Control & Autonomy Score	122	13.15	3.11	122	13.16	3.10	-0.02	-0.08	121	0.936
CASP Pleasure Score	122	8.38	1.19	122	8.30	1.43	0.08	0.71	121	0.480
CASP Self-realisation Score	122	7.20	1.99	122	7.30	1.98	-0.11	-0.76	121	0.450
CASP-12 total Score	122	28.72	5.55	122	28.77	5.72	-0.05	-0.13	121	0.897

Consistency of demographic characteristics reported across online and paper survey modes

While there were some differences in agreement across forms, agreement was high overall and there were no clear systematic defences across modes in the pilot sample.

Current work status

Which of the following best describes your current work status:	Full-time paid work, for an employer	Part-time paid work, for an employer	Full time self-employed paid employment	Part time self-employed paid employment	Flexible work schedule negotiated with employer	Project or contract work (short term and full time)	Project or contract work (short term and part time)	Fully retired, no paid work	Full time homemaker	Unable to work due to health or disability issue	Unemployed and seeking work	Other
Full-time paid work, for an employer	36	1	0	0	0	0	0	0	0	0	0	0
Part-time paid work, for an employer	1	6	0	0	0	0	0	0	0	0	0	0
Full time self-employed paid employment	0	0	3	0	0	0	0	0	0	0	0	0
Part time self-employed paid employment	0	0	0	7	0	0	0	0	0	0	0	0
Flexible work schedule negotiated with employer	0	0	0	0	0	0	0	0	1	0	0	0
Project or contract work (short term and full time)	0	0	0	0	0	2	0	0	0	0	0	0
Project or contract work (short term and part time)	0	0	0	1	1	0	8	0	0	0	0	0
Fully retired, no paid work	0	0	0	0	0	0	1	39	0	0	0	0
Full time homemaker	0	0	0	0	0	0	0	0	0	0	0	0
Unable to work due to health or disability issue	0	0	0	0	0	0	0	1	0	2	0	0
Unemployed and seeking work	0	0	0	0	0	0	0	1	0	0	1	0
Other	0	0	0	0	0	0	0	0	0	0	0	3

(Left axis label: Which of the following best describes your current work status:)

Income personal

$n = 13$ participants indicated more than one answer in the online survey. Issues reported by a participant indicated that she was unable to correct a given response and was able to entered two responses. Thus, we excluded cases who provided multiple responses from this analysis.

	Personal income (paper)														
	$1-$5,000	$5,001-$10,000	$10,001-$15,000	$15,001-$20,000	$20,001-$25,000	$25,001-$30,000	$30,001-$35,000	$35,001-$40,000	$40,001-$50,000	$50,001-$60,000	$60,001-$70,000	$70,001-$100,000	$100,001-$150,000	$150,001-$200,000	$200,001 or more
$1-$5,000	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$5,001-$10,000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$10,001-$15,000	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0
$15,001-$20,000	0	0	0	5	2	0	0	1	0	0	0	0	0	0	0
$20,001-$25,000	0	0	0	2	8	1	0	0	0	0	0	0	0	0	0
$25,001-$30,000	0	0	1	1	3	6	0	0	0	0	0	0	0	0	0
$30,001-$35,000	0	0	0	0	0	3	4	1	0	0	0	0	0	0	0
$35,001-$40,000	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
$40,001-$50,000	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
$50,001-$60,000	0	0	0	0	0	0	0	1	1	8	1	0	0	0	0
$60,001-$70,000	0	0	0	0	0	0	0	0	0	1	5	2	1	0	0
$70,001-$100,000	0	0	0	0	0	0	0	0	0	0	0	12	1	0	0
$100,001-$150,000	0	0	0	0	0	0	0	0	0	0	0	1	12	0	0
$150,001-$200,000	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0
$200,001 or more	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Personal income (online)

Income household

$n = 5$ participants indicated more than one answer in the online survey. Issues reported by a participant indicated that she was unable to correct a given response and was able to entered two responses. Thus, we excluded cases who provided multiple responses from this analysis.

		Household income (paper)														
	loss	zero income	$5,001-$10,000	$10,001-$15,000	$15,001-$20,000	$20,001-$25,000	$25,001-$30,000	$30,001-$35,000	$35,001-$40,000	$40,001-$50,000	$50,001-$60,000	$60,001-$70,000	$70,001-$100,000	$100,001-$150,000	$150,001-$200,000	$200,001 or more
loss	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
zero income	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
$5,001-$10,000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$10,001-$15,000	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
$15,001-$20,000	0	0	0	0	1	2	0	0	1	0	0	0	0	0	0	0
$20,001-$25,000	0	0	0	0	0	3	1	0	0	0	0	0	0	0	0	0
$25,001-$30,000	0	0	0	0	0	3	4	0	0	0	1	0	0	0	0	0
$30,001-$35,000	0	0	0	0	0	0	2	4	1	0	0	1	0	0	0	0
$35,001-$40,000	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0
$40,001-$50,000	0	0	0	0	0	0	0	0	1	1	2	1	0	0	0	0
$50,001-$60,000	0	0	0	0	0	0	0	0	0	2	4	0	0	0	0	0
$60,001-$70,000	0	0	0	0	0	0	0	0	0	0	2	2	3	0	0	0
$70,001-$100,000	0	0	0	0	0	0	0	0	0	0	1	1	12	1	0	0
$100,001-$150,000	0	0	0	0	0	0	0	0	0	0	0	0	1	11	3	0
$150,001-$200,000	0	0	0	0	0	0	0	0	0	0	0	1	0	4	7	0
$200,001 or more	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	6

Household income (online)

Marital status

	Which one of these statements is true about you? (paper)					
Which one of these statements is true about you? (online)	I am married	I am in a civil union/de facto/partnered relationship	I am divorced or permanently separated from my legal husband or wife	I am a widow or widower	I am single	Total
I am married	62	0	0	0	0	62
I am in a civil union/de facto/partnered relationship	1	9	0	0	0	10
I am divorced or permanently separated from my legal husband or wife	0	0	17	0	1	18
I am a widow or widower	0	0	0	12	0	12
I am single	0	0	0	0	13	13
Total	63	9	17	12	14	115

Qualifications

	What is your highest educational qualification? (paper)				Total
What is your highest educational qualification? (online)	No qualifications	Secondary school qualifications (e.g., School Certificate, University entrance, NCEA)	Post-secondary certificate, diploma, or trade diploma	University degree	
No qualifications	5	0	0	0	5
Secondary school qualifications (e.g., School Certificate, University entrance, NCEA)	0	10	3	0	13
Post-secondary certificate, diploma, or trade diploma	0	1	22	2	25
University degree	0	0	2	70	72
Total	5	11	27	72	115

Ethnic identity

	Ethnic group you identify with the most (paper)				
Ethnic group you identify with the most (online)	New Zealand European	Māori	Indian	Other	Total
New Zealand European	99	0	0	3	102
Māori	0	1	0	0	1
Indian	0	0	0	1	1
Other	3	0	0	4	7
Total	102	1	0	8	111

Conclusion

With the exception of the Reliable Alliance subscale of the Social Provisions Scale, core measures of the Health, Work and Retirement study displayed scalar invariance across the paper and online survey modes with reference to change in CFI and McDondalds NCI. The Reliable Alliance subscale displayed metric invariance across all indicators. In contrast, measures generally failed to display metric invariance across modes with reference to the recommended critical values for Gamma Hat, calculated using values of RMSEA.

Within the pilot study sample, total scores were generally consistent across surveys completed in the online and paper survey modes, although there was a small effect of order/mode such that mental health as assessed by the SF-12 and CESD rated more poorly when participants completed the initial online questionnaire. Mean differences in self-reported experiences of health and wellbeing over a 3-week period may occur, that these differences do not necessarily represent mode effects. Thus, while we would not expect great variation in current experiences of mental health over the time between the surveys, it is possible that these effects represent a small but measureable difference in the psychological experiences of respondents between their mental state at the time they decided to participate in the online survey (directed via social media or personal networks), to the time they were prompted to complete the paper survey a month later. As such, these findings of general consistent of participant's reported experiences between the test and re-test assessment are considered in light of the expected small effect size and associated lack of statistical power.

There was a good level of consistency in single-item reports of demographic characteristics, including employment, income, marital status, education, and ethnicity. The lowest level of consistency was observed in reported personal and household income, however the great majority of differences were within 1 income bracket and there appeared to be no systematic difference in reporting greater or lesser income across survey modes.

Overall, results of the online pilot study suggest that the psychometrics properties of the cores measures of the Health, Work and Retirement study are maintained across online and paper survey administrations. The primary focus of the current work was the assessment of the measurement invariance of models across these survey modes between two samples, matched on basic demographic characteristics (age, gender and education). These groups were also highly similar in terms of employment, income and health. However, the findings were limited but a small sample size and the use of two separate samples, recruited through different modes.

Recommendation

Based on the current findings and the opportunities presented by the use of online survey methods, we suggest that the Health, Work and Retirement study partially adopts an online survey method for 2018 survey. Specifically, the 2018 survey could maintain the paper survey for longitudinal participants invited to participate in the 2018 survey. However, the online survey mode may be offered to new participants, recruited as part of the steady state design (in which persons aged 55-57 are recruited every 2 years to maintain representation of persons aged 55+ in each survey wave). In this way, risks to the longitudinal continuity of measures is mitigated, while the acceptability of the online mode and its psychometric invariance may be further evaluated in a larger sample within the sampling frame of the HWR. This is similar to the strategy adopted by other longitudinal studies which have moved to take advantage of online survey modes while minimising risk and with respect to the preferred mode of responding by existing participants, such as that employed for the most recent cohorts recruited Australian Longitudinal Study on Women's Health (Loxton et al., 2018; Mishra et al., 2014). In 2018, the study will aim to

recruit $n = 1066$ new participants. It could be expected that a good portion of these would take up the opportunity to complete the survey online. This larger sample may be used to confirm the current pilot study results in a larger sample, allowing the online and paper survey data to be combined. If the psychometrics consistency of the online version is confirmed in 2018, a wider offering of the online survey mode could be made to longitudinal participants in 2020.

References

Hox, J. J., De Leeuw, E. D., & Zijlmans, E. A. O. (2015). Measurement equivalence in mixed mode surveys. *Frontiers in Psychology, 6*(87).

Wolf, E., Harrington, K., Clark, S., & Miller, M. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement, 73*(6), 913-934.

Steinmetz, H., Schmidt, P., Tina-Booh, A., Wieczorek, S., & Schwartz, S. H. (2008). Testing measurement invariance using multigroup CFA: differences between educational groups in human values measurement. *Quality & Quantity, 43*(4), 599.

Steenkamp, J., Benedict, E. M., & Baumgartner, H. (1998). Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research, 25*(1), 78-107.

Meredith, W., & Teresi, J. A. (2006). An Essay on Measurement and Factorial Invariance. *Medical Care, 44*(11), S69-S77.

Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology, 9*(1), 1-12.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*(2), 233-255.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2006). *The utility of alternative fit indices in tests of measurement invariance.* Paper presented at the Academy of management proceedings.

Ware, J., Kosinski, M., & Keller, S. (1996). A 12-Item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220-233.

Ware, J., Kosinski, M., Turner-Bowker, D., & Gandek, B. (2002). *How to Score Version 2 of the SF-12® Health Survey (With a Supplement Documenting Version 1).* Lincoln, RI: Quality Metric Incorporated.

Ware, J. E., Jr., Kosinski, M., Gandek, B., Aaronson, N. K., Apolone, G., Bech, P., . . . Sullivan, M. (1998). The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. International Quality of Life Assessment. *J Clin Epidemiol, 51*(11), 1159-1165.

Frieling, M. A., Davis, W. R., & Chiang, G. (2013). The SF-36v2 and SF-12v2 health surveys in New Zealand: Norms, scoring coefficients and cross-country comparisons. *Australian and New Zealand Journal of Public Health, 37*(1), 24-31.

Bartsch, L. J., Butterworth, P., Byles, J. E., Mitchell, P., Shaw, J., & Anstey, K. J. (2011). Examining the SF-36 in an older population: Analysis of data and presentation of Australian adult reference scores from the Dynamic Analyses to Optimise Ageing (DYNOPTA) project. *Quality of Life Research, 20*(8), 1227-1236.

Anagnostopoulos, F., Niakas, D., & Tountas, Y. (2009). Comparison between exploratory factor-analytic and SEM-based approaches to constructing SF-36 summary scores. *Quality of Life Research, 18*(1), 53-63.

Farivar, S. S., Cunningham, W. E., & Hays, R. D. (2007). Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.1. *Health and Quality of Life Outcomes, 5*.

Güthlin, C., & Walach, H. (2007). MOS-SF 36: Structural equation modeling to test the construct validity of the second-order factor structure. *European Journal of Psychological Assessment, 23*(1), 15-23.

Hann, M., & Reeves, D. (2008). The SF-36 scales are not accurately summarised by independent physical and mental component scores. *Quality of Life Research, 17*(3), 413-423.

Jakobsson, U., Westergren, A., Lindskov, S., & Hagell, P. (2012). Construct validity of the SF-12 in three different samples. *Journal of Evaluation in Clinical Practice, 18*(3), 560-566.

Christensen, L. N., Ehlers, L., Larsen, F. B., & Jensen, M. B. (2013). Validation of the 12 Item Short form Health Survey in a Sample from Region Central Jutland. *Social Indicators Research, 114*(2), 513-521.

Montazeri, A., Vahdaninia, M., Mousavi, S. J., Asadi-Lari, M., Omidvari, S., & Tavousi, M. (2011). The 12-item medical outcomes study short form health survey version 2.0 (SF-12v2): a population-based validation study from Tehran, Iran. *Health Qual Life Outcomes, 9*, 12.

Okonkwo, O. C., Roth, D. L., Pulley, L., & Howard, G. (2010). Confirmatory factor analysis of the validity of the SF-12 for persons with and without a history of stroke. *Qual Life Res, 19*(9), 1323-1331.

Damásio, B. F., Andrade, T. F., & Koller, S. H. (2015). Psychometric Properties of the Brazilian 12-Item Short-Form Health Survey Version 2 (SF-12v2). *Paidéia (Ribeirão Preto), 25*, 29-37.

Wilson, D., Tucker, G., & Chittleborough, C. (2002). Rethinking and rescoring the SF-12. *Sozial- und Praventivmedizin, 47*(3), 172-177.

Montazeri, A., Vahdaninia, M., Mousavi, S. J., & Omidvari, S. (2009). The Iranian version of 12-item Short Form Health Survey (SF-12): factor structure, internal consistency and construct validity. *BMC Public Health, 9*, 341.

Fisher, K., & Newbold, K. B. (2014). Validity of the SF-12 in a Canadian Old Order Mennonite Community. *Applied Research in Quality of Life, 9*(2), 429-448.

Tucker, G., Adams, R., & Wilson, D. (2010). New Australian population scoring coefficients for the old version of the SF-36 and SF-12 health status questionnaires. *Quality of Life Research, 19*(7), 1069-1076.

Fleishman, J. A., Selim, A. J., & Kazis, L. E. (2010). Deriving SF-12v2 physical and mental health summary scores: A comparison of different scoring algorithms. *Quality of Life Research, 19*(2), 231-241.

Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement, 1*(3), 385-401.

Shafer, A. B. (2006). Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology*(1), 123.

Lee, A. E. Y., & Chokkanathan, S. (2008). Factor structure of the 10-item CES-D scale among community dwelling older adults in Singapore. *International Journal of Geriatric Psychiatry, 23*(6), 592-597.

González, P., Nuñez, A., Merz, E., Brintz, C., Weitzman, O., Navas, E. L., . . . Gallo, L. C. (2017). Measurement properties of the Center for Epidemiologic Studies Depression Scale (CES-D 10): Findings from HCHS/SOL. *Psychological Assessment, 29*(4), 372-381.

Mohebbi, M., Nguyen, V., McNeil, J. J., Woods, R. L., Nelson, M. R., Shah, R. C., . . . Berk, M. (2017). Psychometric properties of a short form of the Center for Epidemiologic Studies Depression (CES-D-10) scale for screening depressive symptoms in healthy community dwelling older adults. *General Hospital Psychiatry*.

Hyde, M., Wiggins, R. D., Higgs, P., & Blane, D. B. (2003). A measure of quality of life in early old age: The theory, development and properties of a needs satisfaction model (CASP-19). *Aging & Mental Health, 7*(3), 186-194.

Wiggins, R. D., Netuveli, G., Hyde, M., Higgs, P., & Blane, D. (2008). The Evaluation of a Self-enumerated Scale of Quality of Life (CASP-19) in the Context of Research on Ageing: A Combination of Exploratory and Confirmatory Approaches. *Social Indicators Research, 89*(1), 61-77.

Kim, G. R., Netuveli, G., Blane, D., Peasey, A., Malyutina, S., Simonova, G., . . . Pikhart, H. (2015). Psychometric properties and confirmatory factor analysis of the CASP-19, a measure of quality of life in early old age: the HAPIEE study. *Aging & Mental Health, 19*(7), 595-609.

Sexton, E., King-Kallimanis, B. L., Conroy, R. M., & Hickey, A. (2013). Psychometric evaluation of the CASP-19 quality of life scale in an older Irish cohort. *Quality of Life Research, 22*(9), 2549-2559.

Pérez-Rojo, G., Martín, N., Noriega, C., & López, J. (2017). Psychometric properties of the CASP-12 in a Spanish older community dwelling sample. *Aging & Mental Health*, 1-9.

Towers, A., Yeung, P., Stevenson, B., Stephens, C., & Alpass, F. (2015). Quality of life in indigenous and non-indigenous older adults: assessing the CASP-12 factor structure and identifying a brief CASP-3. *Quality of Life Research, 24*(1), 193-203.

Jensen, J., Spittal, M., Crichton, S., Sathiyandra, S., & Krishnan, v. (2002). *Chapter 4: Development of the Generic Living Standard Scale*. Retrieved from Wellington https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/monitoring/living-standards/living-standards-elsi.html

Breheny, M., Stephens, C., Henricksen, A., Stevenson, B., Carter, K., & Alpass, F. (2014). Measuring living standards of older people using Sen's Capability Approach: development and validation of the LSCAPE-24 (Living Standards Capabilities for Elders) and LSCAPE-6. *Ageing and Society, 36*(2), 307-332.

Loxton, D., Tooth, L., Harris, M. L., Forder, P. M., Dobson, A., Powers, J., . . . Mishra, G. (2018). Cohort Profile: The Australian Longitudinal Study on Women's Health (ALSWH) 1989–95 cohort. *International Journal of Epidemiology, 47*(2), 391-392e.

Mishra, G. D., Hockey, R., Powers, J., Loxton, D., Tooth, L., Rowlands, I., . . . Dobson, A. (2014). Recruitment via the Internet and social networking sites: the 1989-1995 cohort of the Australian Longitudinal Study on Women's Health. *J Med Internet Res, 16*(12), e279.

Appendix 1: Study materials

Information statement

Letter to participants sent with paper version of the pilot survey

Paper version of the pilot survey

An online pilot of the
2018 New Zealand Health, Work & Retirement survey

INFORMATION SHEET (v C1.0)

What is the New Zealand Health, Work & Retirement longitudinal study?

The New Zealand Health, Work & Retirement survey is a longitudinal study of persons aged 55 years and over who are living in New Zealand. It aims to provide information on issues such as health, work, retirement and housing relevant to this age group. A core component of the study is a postal survey, sent to participants every two years.

The study is an initiative of Massey University's Health and Ageing Research Team and began in 2006. The study provides older New Zealand residents with the opportunity to share their experiences to help inform research and policy on these important issues. To date, over 10,000 people have been surveyed, with over 25,000 surveys returned.

What is the purpose of the 2018 pilot study?

In 2018 we aim to give participants in the New Zealand Health, Work & Retirement study an opportunity to complete the survey online. However, as the study has tracked a large number of people over a 10-year period, it is important that any change to our research method does not interfere with the conclusions made about how participants fare over time.

Although there are many benefits of making the survey online, a pilot study is needed to assess whether the questions in the online and postal surveys are answered in similar ways. The pilot study is being conducted by Professor Fiona Alpass, Professor Christine Stephens and Dr Joanne Allen, ahead of the 2018 Health, Work & Retirement survey to test whether responses to the online and postal surveys are comparable.

Who can participate in the pilot study?

We are seeking 400 people aged 55 and over to participate. If you have previously responded to the New Zealand Health, Work & Retirement survey, you are also welcome to respond to the current pilot study. However, your data for the pilot study will not be linked to any information previously provided to the New Zealand Health, Work & Retirement longitudinal study.

What is involved in participating?

Participation involves completing an online survey which contains questions about your physical health, mental health, social and economic wellbeing. At the end of the survey, participants are asked to provide a postal address so that they can be mailed a paper copy of the survey. The survey will be sent with a return addressed postage-paid envelope so that it can be returned once complete. Each survey will take about 10-15 minutes.

Does it matter if my responses to the online and postal surveys are not the same?

No. Given that the surveys will be completed 1-2 weeks apart and many of the questions are subjective, it is expected that some answers will change. Our analyses will examine whether association between responses to questions in the online and postal surveys are comparable.

Will my data remain secure and confidential?

All responses are completely confidential, and will be used only for the purposes of health research. It will not be possible to identify individuals in any dataset or report from the study. Paper copy surveys will be kept in a locked room and all surveys will be destroyed five years after the completion of the study.

What are my rights as a participant in this study?

If you decide to participate, you have the right to:

- Decline to answer any particular question;
- Ask any questions about the study or withdraw from the project by contacting HART researchers;
- Know that all information will be kept strictly confidential and will be used only for health research;
- Know that reports and publications from this study will not identify any individual taking part.

Who can I contact if I have further questions about the study?

You can contact HART researchers at any time on our free-phone number (**0800 100 134**) or you can email a question to **hart@massey.ac.nz**. We have a dedicated website to provide general information for those participating or interested in studies run by the Health and Ageing Research Team. You can access this website at **hart.massey.ac.nz**

This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researchers named in this document are responsible for the ethical conduct of this research. If you have any concerns about the conduct of this research that you want to raise with someone other than the researchers, please contact Dr Brian Finch, Director (Research Ethics), email humanethics@massey.ac.nz

Thank you!

We greatly appreciate your consideration of this invitation and we welcome your participation in the Health, Work & Retirement study. If you wish to participate you can click the Accept button below indicating that you have read this introductory information and the survey will begin.

Sincerely,

Professor Fiona Alpass

On behalf of the *Health & Ageing Research Team*.

<Todays_Date>

<Mailing_Name>
<Mailing_address_1>
<Mailing_address_2>
<Mailing_address_3>

Dear <Mailing_Name>

On behalf of Massey University's Health and Ageing Research Team (HART) I would like to thank you for completing the online pilot of the New Zealand Health, Work & Retirement survey. Launched in 2006, the survey provides older New Zealand residents with the opportunity to share their experiences to help inform national and international discussions on important issues. With your participation in this pilot study we will be able to evaluate whether the 2018 survey can be delivered online, without compromising its scientific validity, providing respondents with more options regarding how they want to participate and also conserving resources.

Enclosed you will find a blank hard copy of the survey which you recently completed online. So that we may compare responses to the online and hard copy versions of the survey, we ask that you complete and return the enclosed survey. This can be returned in the freepost envelope supplied. All information that you provide is completely confidential, and will be used only for the purposes of this research.

If you would like to learn more about the project and what is involved, please feel free to contact us at any time to discuss this project or ask any questions you may have. You can view the study information sheet again by visiting **https://psylab.massey.ac.nz/hart/** and you can contact us on the HART free-phone number (**0800 100 134**) or email at: **hart@massey.ac.nz**.

Thank you again for taking to participate. Your contribution to this national research study is appreciated and is vital to the success of this research initiative.

Yours sincerely

Professor Fiona Alpass, on behalf of the Health & Ageing Research Team (HART) at Massey University:

Professor Christine Stephens	Dr Mary Breheny
Professor Fiona Alpass	Dr Polly Yeung
Dr Joanne Taylor	Dr Andy Towers
Dr Rachael Pond	Mr Brendan Stevenson
Dr Joanne Allen	Dr Juliana Mansvelt
Ms Vicki Beagley	Dr Agnes Szabo

General instructions for completing the survey

Please read the following carefully

- All the information you give us is in confidence and will be used only for the purposes of the Health, Work and Retirement study.
- There are no right or wrong answers - we want the response that is best for you.
- It is important that you give your own answers to the questions.
- Do not linger too long over each question - usually your first response is best.

For each question in the survey you will be asked to provide either:

- <u>a single answer that is most appropriate</u>. These are the most common question types - for these items, please mark (e.g. ✔ or ✗) one box on each line in pen or pencil. If you make a mistake, simply scribble it out and mark the correct answer.
- <u>one or more responses, as appropriate.</u> For these items you will be instructed to *'Please tick all that apply'*.
- <u>a free text response</u>. To provide free text, please print your response as clearly as possible on the line provided.

Example question and response: Please tick 'Yes' to indicate if a health professional has told you that you have any of the following conditions:

(Please tick <u>one</u> box on each line)	No	Yes, in the last 12 months	Yes, prior to the last 12 months
Sleep disorder	✔ 1	2	3
Stroke	✔ 1	2	3
Cancer	1	✔ 2	3
Please specify cancer type:	*melanoma*		

- <u>a number</u>: where a number or date is required, print the figure in the box provided.

Example question and response: How many of the following people are you in regular contact with? Please place a zero or a number in the square as appropriate:

Adult child(ren) and/or grandchild(ren)/mokopuna	*5*

Thank you for taking the time to complete this questionnaire
If you need help to answer any questions please contact us either on the HART free-phone line <u>0800 100 134</u> or via email: <u>hart@massey.ac.nz</u>

YOUR HEALTH, WELLBEING AND QUALITY OF LIFE

These are questions about your general health. Please tick one box on each line.

Q1   In general, would you say your health is: (Please tick one box)

Excellent	Very good	Good	Fair	Poor
☐1	☐2	☐3	☐4	☐5

The following questions are about activities you might do during a typical day.

Q2.   Does your health now limit you in these activities? If so, how much?

(Please tick one box on each line)	Yes, limited a lot	Yes, limited a little	No, not limited at all
Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf	☐1	☐2	☐3
Climbing several flights of stairs	☐1	☐2	☐3

Q3.   During the past 4 weeks, how much of the time have you had any of the following problems with your work, or other regular daily activities as a result of your physical health?

(Please tick one box on each line)	All of the time	Most of the time	Some of the time	A little of the time	None of the time
Accomplished less than you would like	☐1	☐2	☐3	☐4	☐5
Were limited in the kind of work or other activities	☐1	☐2	☐3	☐4	☐5

Q4.   During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

(Please tick one box on each line)	All of the time	Most of the time	Some of the time	A little of the time	None of the time
Accomplished less than you would like	☐1	☐2	☐3	☐4	☐5
Didn't do work or other activities as carefully as usual	☐1	☐2	☐3	☐4	☐5

Q5.   During the <u>past 4 weeks</u>, how much did <u>pain</u> interfere with your normal work (including both work outside the home and housework)? (Please tick <u>one</u> box)

Not at all	A little bit	Moderately	Quite a bit	Extremely
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q6.   These questions are about how you feel and how things have been with you <u>during the past 4 weeks</u>. For each question, please give the one answer that comes closest to the way you have been feeling. How much time during the <u>past 4 weeks</u>:

(Please tick <u>one</u> box on each line)	All of the time	Most of the time	Some of the time	A little of the time	None of the time
Have you felt calm and peaceful?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Did you have a lot of energy?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
Have you felt downhearted and depressed?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q7.   During the <u>past 4 weeks</u>, how much of the time has your <u>physical health or emotional problems</u> interfered with your social activities (like visiting friends, relatives, whānau, etc.)?

(Please tick <u>one</u> box)

All of the time	Most of the time	Some of the time	A little of the time	None of the time
☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q8. Below is a list of some of the ways you may have felt or behaved. Please indicate how often you have felt this way during the past week (7 days).

(Please tick one box on each line)	Rarely or none of the time	Some or a little of the time	Occasionally or a moderate amount of the time	All of the time
I was bothered by things that usually don't bother me	☐ 1	☐ 2	☐ 3	☐ 4
I had trouble keeping my mind on what I was doing	☐ 1	☐ 2	☐ 3	☐ 4
I felt depressed	☐ 1	☐ 2	☐ 3	☐ 4
I felt that everything I did was an effort	☐ 1	☐ 2	☐ 3	☐ 4
I felt hopeful about the future	☐ 1	☐ 2	☐ 3	☐ 4
I felt fearful	☐ 1	☐ 2	☐ 3	☐ 4
My sleep was restless	☐ 1	☐ 2	☐ 3	☐ 4
I was happy	☐ 1	☐ 2	☐ 3	☐ 4
I felt lonely	☐ 1	☐ 2	☐ 3	☐ 4
I could not "get going"	☐ 1	☐ 2	☐ 3	☐ 4

Q9. Here is a list of statements that people have used to describe their lives or how they feel. We would like to know how often, if at all, you think the following applies to you.

(Please tick one box on each line)	Often	Sometimes	Not often	Never
My age prevents me from doing the things I would like to	1	2	3	4
I feel that what happens to me is out of my control	1	2	3	4
I feel left out of things	1	2	3	4
I can do the things that I want to do	1	2	3	4
I feel that I can please myself what I do	1	2	3	4
Shortage of money stops me from doing things I want to do	1	2	3	4
I look forward to each day	1	2	3	4
I feel that my life has meaning	1	2	3	4
I enjoy the things that I do	1	2	3	4
I feel full of energy these days	1	2	3	4
I feel that life is full of opportunities	1	2	3	4
I feel that the future looks good for me	1	2	3	4

The following questions are about your health and health related behaviours. Please tick the box that best answers each question.

Q10. How often do you take part in sports or activities that are:

(Please tick one box on each line)	More than once a week	Once a week	One to three times a month	Hardly ever or never
...vigorous (e.g., running or jogging, swimming, aerobics)	1	2	3	4
...moderately energetic (e.g., gardening, brisk walking)	1	2	3	4
...mildly energetic (e.g., vacuuming, laundry/washing)	1	2	3	4

WHĀNAU, FAMILY AND FRIENDS

Q11 Think about your current relationships with friends, family/whānau members, co-workers, community members and so on. To what extent do you agree that each statement describes your current relationships with other people?

(Please tick one box on each line)	Strongly Disagree	Disagree	Agree	Strongly Agree
There are people I can depend on to help me if I really need it	☐ 1	☐ 2	☐ 3	☐ 4
I feel that I do not have close personal relationships with other people	☐ 1	☐ 2	☐ 3	☐ 4
There is no one I can turn to for guidance in times of stress	☐ 1	☐ 2	☐ 3	☐ 4
There are people who depend on me for help	☐ 1	☐ 2	☐ 3	☐ 4
There are people who enjoy the same social activities I do	☐ 1	☐ 2	☐ 3	☐ 4
Other people do not view me as competent	☐ 1	☐ 2	☐ 3	☐ 4
I feel personally responsible for the well-being of another person	☐ 1	☐ 2	☐ 3	☐ 4
I feel part of a group of people who share my attitudes and beliefs	☐ 1	☐ 2	☐ 3	☐ 4

Continued…	Strongly Disagree	Disagree	Agree	Strongly Agree
I do not think other people respect my skills and abilities	☐ 1	☐ 2	☐ 3	☐ 4
If something went wrong, no one would come to my assistance	☐ 1	☐ 2	☐ 3	☐ 4
I have close relationships that provide me with a sense of emotional security and well-being	☐ 1	☐ 2	☐ 3	☐ 4
There is someone I could talk to about important decisions in my life	☐ 1	☐ 2	☐ 3	☐ 4
I have relationships where my competence and skills are recognized	☐ 1	☐ 2	☐ 3	☐ 4
There is no one who shares my interests and concerns	☐ 1	☐ 2	☐ 3	☐ 4
There is no one who really relies on me for their well-being	☐ 1	☐ 2	☐ 3	☐ 4
There is a trustworthy person I could turn to for advice if I were having problems	☐ 1	☐ 2	☐ 3	☐ 4

Continued…	Strongly Disagree	Disagree	Agree	Strongly Agree
I feel a strong emotional bond with at least one other person	□ 1	□ 2	□ 3	□ 4
There is no one I can depend on for aid if I really need it	□ 1	□ 2	□ 3	□ 4
There is no one I feel comfortable talking about problems with	□ 1	□ 2	□ 3	□ 4
There are people who admire my talents and abilities	□ 1	□ 2	□ 3	□ 4
I lack a feeling of intimacy with another person	□ 1	□ 2	□ 3	□ 4
There is no one who likes to do the things I do	□ 1	□ 2	□ 3	□ 4
There are people I can count on in an emergency	□ 1	□ 2	□ 3	□ 4
No one needs me to care for them	□ 1	□ 2	□ 3	□ 4

Q12 Please indicate for each of the statements below, the extent to which they apply to the way you feel now.

(Please tick one box on each line)	Yes	More or less	No
I experience a general sense of emptiness	□ 1	□ 2	□ 3
There are plenty of people I can rely on when I have problems	□ 1	□ 2	□ 3
There are many people I can trust completely	□ 1	□ 2	□ 3
There are enough people I feel close to	□ 1	□ 2	□ 3
I miss having people around	□ 1	□ 2	□ 3
I often feel rejected	□ 1	□ 2	□ 3

Q13. How satisfied are you with the support you receive from your personal relationships?
(Please tick one box)

Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
□ 1	□ 2	□ 3	□ 4	□ 5

Q14. How often do your personal relationships make you feel loved and wanted?
(Please tick one box)

None of the time	A little of the time	Some of the time	Most of the time	All of the time
□ 1	□ 2	□ 3	□ 4	□ 5

Q15.   Do you live with other people?

☐₁   Yes   ☐₂   No

Q16.   Do you have any <u>domestic pets</u> who live on the same property or in the same household as you?

☐₁   Yes   ☐₂   No (if not, please skip to Q22)

Q17.   Please tick as many options as you need to indicate all the <u>domestic pets</u> who live on the same property or in the same household as you. Please also indicate the number of each type of pet and the age of your oldest pet of each type.

(Please indicate <u>all</u> that apply)	Yes	Number of pets	Age of oldest pet
Dog	☐₁		
Cat	☐₁		
Bird	☐₁		
Fish	☐₁		
Reptile	☐₁		
Other(s)	☐₁		

Q18. Please indicate whether these items relate to your relationship with your pet(s)

(Please tick <u>one</u> box on each line)	Yes	No	Maybe	Don't know
Do you consider your pet a friend?	☐₁	☐₂	☐₃	☐₄
Do you talk to your pet?	☐₁	☐₂	☐₃	☐₄
Would you say that owning a pet adds to your happiness?	☐₁	☐₂	☐₃	☐₄
Do you talk with others about your pet?	☐₁	☐₂	☐₃	☐₄
Do you often play with your pet?	☐₁	☐₂	☐₃	☐₄
Does your pet know how you feel about things?	☐₁	☐₂	☐₃	☐₄

Q19.   How satisfied are you with the support you receive from your pet(s)? (Please tick <u>one</u> box)

Very dissatisfied	Dissatisfied	Neither satisfied nor dissatisfied	Satisfied	Very satisfied
☐₁	☐₂	☐₃	☐₄	☐₅

Q20.   How often do your pet(s) make you feel loved and wanted? (Please tick <u>one</u> box)

None of the time	A little of the time	Some of the time	Most of the time	All of the time
☐₁	☐₂	☐₃	☐₄	☐₅

Q21. I have close relationships with my pet(s) that provide me with a sense of emotional security and wellbeing (Please tick <u>one</u> box)

Strongly disagree	Disagree	Agree	Strongly Agree
☐ 1	☐ 2	☐ 3	☐ 4

WHERE YOU LIVE

Q22 Please rate your level of agreement to each of these statements in relation to your present home:

(Please tick <u>one</u> box on each line)	No, definitely not		Neutral		Yes, definitely
I am satisfied with my house	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
I am satisfied with my neighbourhood	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
I am unhappy with the living conditions of my house	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
My house enables me to see friends and family as often as I like	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
My house enables me to participate in community activities as often as I like	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
My house supports all my daily activities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
My home meets all my needs	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
I am able to keep my house warm	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
My house is difficult for me to clean	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
I have difficulties with house repairs	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Q23 Please rate your level of agreement to each of these statements in relation to your present neighbourhood:

(Please tick <u>one</u> box on each line)	No, definitely not		Neutral		Yes, definitely
I can get to shops easily	□1	□2	□3	□4	□5
I have access to transport	□1	□2	□3	□4	□5
I am close enough to any help I need	□1	□2	□3	□4	□5
I am close enough to important facilities	□1	□2	□3	□4	□5
I feel safe at home	□1	□2	□3	□4	□5
I feel safe in my neighbourhood	□1	□2	□3	□4	□5
The neighbourhood is peaceful	□1	□2	□3	□4	□5
I have peace of mind at home	□1	□2	□3	□4	□5
My neighbourhood is pleasant	□1	□2	□3	□4	□5

Q24. How would you describe the condition of your current residence? (Please tick <u>one</u> box)

No repairs or maintenance needed right now	Minor maintenance needed	Some repairs and maintenance needed	Immediate repairs and maintenance needed	Immediate and extensive repairs and maintenance needed
□1	□2	□3	□4	□5

Q25. Does your residence have a problem with dampness or mould? (Please tick <u>one</u> box)

No	Minor problem	Moderate problem	Major problem
□1	□2	□3	□4

Q26. In winter, is your current residence colder than you would like? (Please tick <u>one</u> box)

Yes - always	Yes - often	Yes - sometimes	No
□1	□2	□3	□4

YOUR WORK AND RETIREMENT STATUS

Q27   Which of the following best describes your **current** work status

(Please tick <u>one</u> box in each column)

Full-time paid work, for an employer	☐ 1
Part-time paid work, for an employer	☐ 2
Full time self-employed paid employment	☐ 3
Part time self-employed paid employment	☐ 4
Flexible work schedule negotiated with employer	☐ 5
Project or contract work (short term and full time)	☐ 6
Project or contract work (short term and part time)	☐ 7
Fully retired, no paid work	☐ 8
Full time homemaker	☐ 9
Full time student	☐ 10
Unable to work due to health or disability issue	☐ 11
Unemployed and seeking work	☐ 12
Other Please specify:	☐ 13

YOUR FINANCIAL WELLBEING

Next we ask about your financial circumstances, please be assured that your answers to these questions are completely confidential.

Please indicate both your personal and household income below. If needed, see notes at the back of the questionnaire to help work out your income.

Q28 From all sources of income, what do you expect your annual personal income before tax to be this financial year?

(Please tick one box)

1	loss
2	zero income
3	$1 - $5,000
4	$5,001 - $10,000
5	$10,001 - $15,000
6	$15,001 - $20,000
7	$20,001 - $25,000
8	$25,001 - $30,000
9	$30,001 - $35,000
10	$35,001 - $40,000
11	$40,001 - $50,000
12	$50,001 - $60,000
13	$60,001 - $70,000
14	$70,001 - $100,000
15	$100,001 - $150,000
16	$150,001 - $200,000
17	$200,001 or more

Q29 From all sources of income, what do you expect your annual <u>household</u> income <u>before tax</u> to be this financial year?

(Please tick one box)

1	loss
2	zero income
3	$1 - $5,000
4	$5,001 - $10,000
5	$10,001 - $15,000
6	$15,001 - $20,000
7	$20,001 - $25,000
8	$25,001 - $30,000
9	$30,001 - $35,000
10	$35,001 - $40,000
11	$40,001 - $50,000
12	$50,001 - $60,000
13	$60,001 - $70,000
14	$70,001 - $100,000
15	$100,001 - $150,000
16	$150,001 - $200,000
17	$200,001 or more

Q30 For the following questions, please indicate whether or not you have (or have access to) the item:

(Please tick one box on each line)	Yes, I have it	No, because I don't want it	No, because of the cost	No, for some other reason
Telephone	☐1	☐2	☐3	☐4
Washing machine	☐1	☐2	☐3	☐4
At least two pair of good shoes	☐1	☐2	☐3	☐4
Suitable clothes for important or special occasions	☐1	☐2	☐3	☐4
Personal computer	☐1	☐2	☐3	☐4
Home contents insurance	☐1	☐2	☐3	☐4
Enough room for family/whānau to stay the night	☐1	☐2	☐3	☐4

Q31 For the following questions, please indicate whether or not you do the activity:

(Please tick one box on each line)	Yes, I do it	No, because I don't want to	No, because of the cost	No, for some other reason
Keep the main rooms of your home adequately heated	☐1	☐2	☐3	☐4
Give presents to family/whānau or friends on birthdays, Christmas or other special occasions	☐1	☐2	☐3	☐4
Visit the hairdresser at least once every three months	☐1	☐2	☐3	☐4
Have holidays away from home for at least a week every year	☐1	☐2	☐3	☐4
Have a holiday overseas at least every three years	☐1	☐2	☐3	☐4
Have a night out for entertainment or socialising at least once a fortnight	☐1	☐2	☐3	☐4
Have family/whānau or friends over for a meal at least once every few months	☐1	☐2	☐3	☐4

Q32   The following are a list of things some people do to help keep costs down. In the last 12 months, have you done any of these things?

(Please tick <u>one</u> box on each line)	Not at all	A little	A lot
Gone without or cut back on fresh fruit and vegetables to help keep down costs	☐₁	☐₂	☐₃
Continued wearing clothing that was worn out because you couldn't afford a replacement	☐₁	☐₂	☐₃
Put off buying clothes for as long as possible to help keep down costs	☐₁	☐₂	☐₃
Stayed in bed longer to save on heating costs	☐₁	☐₂	☐₃
Postponed or put off visits to the doctor to help keep down costs	☐₁	☐₂	☐₃
NOT picked up a prescription to help keep down costs	☐₁	☐₂	☐₃
Spent less time on hobbies than you would like to help keep down costs	☐₁	☐₂	☐₃
Gone without or cut back on trips to the shops or other local places to help keep down costs	☐₁	☐₂	☐₃

The following questions are about your material standard of living – the things that money can buy. Your material standard of living does NOT include your capacity to enjoy life. You should NOT take your health into account.

Q33   Generally, how would you rate your material standard of living? (Please tick <u>one</u> box)

High	Fairly high	Medium	Fairly low	Low
☐₁	☐₂	☐₃	☐₄	☐₅

Q34   Generally, how satisfied are you with your current material standard of living? (Please tick <u>one</u> box)

Very satisfied	Satisfied	Neither satisfied nor dissatisfied	Dissatisfied	Very dissatisfied
☐₁	☐₂	☐₃	☐₄	☐₅

Q35   How well does your total income meet your everyday needs for such things as accommodation, food, clothing and other necessities? (Please tick <u>one</u> box)

Not enough	Just enough	Enough	More than enough
☐₁	☐₂	☐₃	☐₄

Q36   Below are statements that people have made about their standard of living. Please indicate how true these statements are for you.

(Please tick <u>one</u> box on each line)	Not true for me at all				Definitely true for me
I can afford to go to a medical specialist if I need to	□₁	□₂	□₃	□₄	□₅
I am able to visit people whenever I wish	□₁	□₂	□₃	□₄	□₅
I am able to give to others as much as I want	□₁	□₂	□₃	□₄	□₅
I am able to do all the things I love	□₁	□₂	□₃	□₄	□₅
I expect a future without money problems	□₁	□₂	□₃	□₄	□₅
My choices are limited by money	□₁	□₂	□₃	□₄	□₅

YOUR PERSONAL SITUATION

Q37   What gender do you identify as? (Please tick <u>one</u> box)

□₁   **Male / Tāne**

□₂   **Female / Wāhine**

□₃   **Gender diverse**

Q38   When were you born?

D	D	/	M	M	/	1	9	Y	Y	**DD/MM/YYYY**

Q39   Which one of these statements is true about you? (Please answer for your **current**, marriage, partnership or situation). (Please tick <u>one</u> box)

□₁ I am married   □₂ I am a widow or widower

□₃ I am in a civil union/de facto/partnered relationship   □₄ I am single

□₅ I am divorced or permanently separated from my legal husband or wife

Q40   What is your highest educational qualification? (Please tick <u>one</u> box)

☐₁   No qualifications

☐₂   Secondary school qualifications (e.g., School Certificate, University entrance, NCEA)

☐₃   Post-secondary certificate, diploma, or trade diploma

☐₄   University degree

Q41   Please indicate below which ethnic group or groups you belong to: (Please tick <u>all that apply</u>)

☐₁   New Zealand European   ☐₁   Niuean

☐₁   Māori   ☐₁   Chinese

☐₁   Samoan   ☐₁   Indian

☐₁   Cook Island Māori   ☐₁   Tongan

☐₁   Other (please specify e.g., Dutch, Japanese, Tokelauan)

Q42   Please indicate below which ethnic group you feel you identify with the most: (Please tick <u>one</u> box)

☐₁   New Zealand European   ☐₅   Niuean

☐₂   Māori   ☐₆   Chinese

☐₃   Samoan   ☐₇   Indian

☐₄   Cook Island Māori   ☐₈   Tongan

☐₉   Other (please specify e.g., Dutch, Japanese, Tokelauan)

Guide notes

Why do you want to know my income?

Information such as income are used to help determine how well respondents to the New Zealand Health, Work and Retirement survey represent the general New Zealand population and whether income is a feature in ageing well. All of the answers you give are kept confidential.

How do I work out my annual personal/household income?

Remember:

- If you and your spouse / partner earn income jointly, only include your part of that income when reporting your personal income.
- Count any payments that are taken out of your income **before** you get it, such as repayments of student loans, union fees, fines or child support.
- DON'T count loans (including student loans), inheritances, sale of household or business assets, lottery wins, matrimonial / civil union / de facto property settlements or one-off lump sum payments.
- DON'T count money given by members of the same household to each other. For example, pocket money given to children, or money given for housekeeping expenses by a flatmate.

Calculating annual income before tax: If you know your weekly or fortnightly income **after tax**, use this table to work out your annual income **before tax**.

After tax weekly income$	After tax fortnightly income $	Before tax annual income $
up to 86	up to 172	21 – 5,000
87 – 172	173 – 343	5,001 – 10,000
173 – 256	344 – 512	10,001 – 15,000
257 – 335	513 – 671	15,001 – 20,000
336 – 414	672 – 829	20,001 – 25,000
415 – 493	830 – 987	25,001 – 30,000
494 – 573	988 – 1,145	30,001 – 35,000
574 – 652	1,146 – 1,303	35,001 – 40,000
653 – 805	1,304 – 1,610	40,001 – 50,000
806 – 939	1,611 – 1,879	50,001 – 60,000
940 – 1,074	1,880 – 2,147	60,001 – 70,000
1,075 – 1,459	2,148 – 2,918	70,001 – 100,000
1,460 – 2,102	2,919 – 4,203	100,001 – 150,000
2,103+	4,204+	150,001+

Standard NZ Super: these are the approximate standard **before tax** rates for NZ Super.

	Fortnightly before tax	Annual before tax
Single, living alone	$900.20	$23,405.20
Single, sharing accommodation	$827.20	$21,507.20
Married person or partner in a civil union or de facto relationship	$681.60	$17,721.60
Married or in a civil union or de facto relationship, both qualify	$645.56	$16,784.56

Thank you for completing the 2018 Health, Work & Retirement pilot survey!

We greatly appreciate your help with this questionnaire. If you have questions about the study, please contact the Health and Ageing Research Team on our free-phone number 0800 100 134 or email: hart@massey.ac.nz.

If you have any additional thoughts about any of the topics in the survey, or experienced any difficulties in completing the survey online, please let us know in the box below.