Helping people access and use quality evidence

YOUTH EMPLOYMENT PATHWAY QUESTIONNAIRE 2.0

Psychometric properties of a preliminary scale measuring the ability to obtain and retain employment

30 April 2021

STANDARD OF PROOF

A BRIEF OVERVIEW

This short report documents the psychometric properties of the Youth Employment Pathway questionnaire (hereafter the questionnaire) that was designed first, to identify needs of young people aged 15-24 in New Zealand that are 'most at risk' of long-term unemployment; and second, to monitor these individual's progress towards employment.

The questionnaire was developed as part of a New Zealand government initiative – *He Poutama Rangatahi* – to help community providers identify needs, monitor and report on progress for their engaged 'at-risk' young people. The questionnaire design built upon extensive engagements with community groups, iwi, employers and local agencies across four regions in New Zealand. Communities (mainly rural) identified critical factors relevant to obtaining and sustaining employment for youth. These critical factors were considered alongside a review of published literature, identifying factors that limit or contribute to employment for young people.

A test instrument was developed and then administered by community providers working with youth in these regions, and resulted in a dataset of 547 young people. We applied the Rasch unidimensional measurement model to these data, checking the psychometric properties and determining the extent to which there is sufficient justification for intended use and interpretation of the data derived from the questionnaire. More importantly, the Rasch model produced diagnostic information, identifying and then testing modifications to the questionnaire so that it can better measure the trait – here, the ability to obtain employment.

The initial results:

1. **Confirmed the modified questionnaire provides a reliable scale measuring a single construct with stable measurement properties of internal consistency.** Specifically, the Rasch Person Separation index of reliability was very high (0.8737), and Chi-Square statistics showed a significant result overall (x^2 = 230.9998, df = 80, $p < 0.001$). We tested the persons-fit of the model ($x = -0.24$; standard deviation =1.22), and although items showed some misfit ($x = 0.0826$, standard deviation of 2.1350), the misfit was not severe as indicated by the Item Characteristic Curves, and further testing was done of the impact on reliability by removing items.
2. **Showed that this scale (as a total score) demonstrated a very strong relationship with employment** ($t(680) = -18.7378$, $p < 0.0001$). These preliminary results are very promising. They provide evidence to promote further use by community providers who work to help young people move towards sustained employment, and subsequent testing to confirm reliability and predictive validity.

The modified questionnaire is provided overleaf.

YOUTH EMPLOYMENT PATHWAY QUESTIONNAIRE

Name: _____ Date of birth: _____ Date: _____

	1	2	3	4	5
1. My relative priorities	I don't want a job at the moment.	I may want a job, but it's not a priority for me.	It is a priority for me to get and keep a job.	It is a top priority for me to get and keep a job.	
2. When there are new skills to learn	If I think it is hard, I am not going to try.	If I think it is hard, I may avoid trying.	I will try to learn new skills, although I know I will make mistakes.	I willingly take on learning new skills, and I learn from my mistakes.	
3. What I believe	I have little confidence in my ability to do paid work.	I have some confidence in my ability to do paid work	I am mostly confident in my ability to do paid work.	I am fully confident in my ability to do paid work.	
4. When I have tasks	I am consistently late or absent, and sometimes do not complete tasks.	I am sometimes late or absent, and may not complete tasks.	I am rarely late or absent, and often complete tasks.	I am never late or absent, and always complete tasks.	
5. When there are challenges	I often give up when things are not going well.	I may give up but I try to work things out when things are not going well.	I often find solutions to challenges.	I welcome challenges, and create solutions.	
6. Working with others	I won't work with other people.	I will work only with my friends or family.	I will work with others if necessary.	I will work with others, and can get on with a variety of people.	I prefer working with others. I get on easily with a variety of people.
7. Work experience	I have no work experience at all.	I have a few months of work experience.	I have 1 to 2 years of work experience.	I have more than 2 years of regular work experience.	
8. Relevant training	I have no relevant training for the job I want.	I have completed some of the relevant training for the job I want.	I have completed most of the relevant training for the job I want.	I have completed all the necessary training for the job I want.	
9. Reading	I struggle to read simple documents and signs. I sometimes guess what they mean.	I can read simple documents (e.g. short forms) and signs (e.g. on the street) easily.	I can understand ideas in long text, and can find information easily.	I can understand ideas in long and complex text easily, and I can put information from several complex texts together to create a general understanding.	
10. Numbers	I struggle to do any mathematics, including adding and subtracting.	I can add and subtract easily, or count forward and backward in 1s and 10s.	I can easily add and subtract, counting forward and backward in 1s and 10s; as well as multiply numbers easily (7 x 8 = 56).	I can add, subtract, multiply and divide numbers, as well as find and convert fractions to percentages easily (3/12 = 1/4 = 25%).	
11. The influential people in my life	They would try to stop me from getting a job.	They would not try to stop me but would not help me get a job.	They may ask me about getting a job and may sometimes help me.	They would help me to get or keep a job.	They would provide me with regular, ongoing support to get and keep a job.

	1	2	3	4	5
12. How I feel about my community	I don't connect to my community. I feel like I don't belong here.	I have a limited sense of belonging. I feel like it is not really my community.	I have some sense of belonging to my community.	I have a strong sense of belonging to my community. I feel a pride of place.	
13. My connection to culture	I have a limited or no sense of connection to my culture.	I connect but don't feel the need to participate with my cultural language, history and practices.	I connect and participate with my cultural language, history and practices when I can.	I have a strong sense of connection to my culture. 'I live and breathe my culture'.	
14. My housing	I don't have a safe home, or I have an uncertain housing situation.	I have a safe home.			
15. My other responsibilities	I don't have any support or services that can help me with my caregiving responsibilities if I get a job.	I have support or services that can help me with my caregiving responsibilities, but they may not be reliable. If I get a job I may need time off for caregiving.	I have no caregiving responsibilities, or I have reliable support available when I get a job.		
16. My physical ability for work	I am not fit for some physical activities required for work (poor fitness or uncontrolled chronic illness e.g. diabetes, asthma.)	I may not be fit but I can do most physical activities required for work (moderate fitness or somewhat controlled chronic illness e.g. diabetes, asthma).	I am fit for all physical activities required for work (good fitness and controlled or no chronic illness e.g. diabetes, asthma).		
17. Broader health	I have regular uncontrolled moods or behaviours, or my mental health is not well managed.	I mostly manage my moods or behaviours, or any mental health issues. I access any necessary treatment or medication.	I have controlled, stable moods or behaviours. I have no mental health issues.		
18. Alcohol or drug use	I drink alcohol or take drugs and feel out of it daily.	I drink alcohol or take drugs socially multiple times a week.	I drink alcohol or take drugs experimentally, mostly on weekends or at social events.	I have not drunk alcohol or taken drugs in the last six months, or I don't use them at all.	
19. Access to jobs	I can't access any jobs (because of transport, location, my age).	There are very few jobs accessible to me (because of transport, location, my age).	There are some jobs accessible to me.	There are relevant jobs accessible and available to me.	
20. Any financial barriers	I can't get a job because I will lose my (WINZ) benefit, or the costs of working are too expensive (e.g. materials, transport).	I would have to find specific types of jobs so I don't lose my benefit, or the costs of working may be too expensive (e.g. materials, transport).	There are financial barriers to me getting a job and I don't know if these will affect me over time.	There may be financial barriers to me getting a job, but these won't affect me.	There are no obvious financial barriers to me getting a job (e.g. loss of benefit, costs of materials, transport).

WHY ASSESS A QUESTIONNAIRE INSTRUMENT?

Organisations regularly collect and use data for various purposes. Some organisations require evidence for accountability purposes to measure if the organisation or programme is meeting established standards. Other times, evidence serves more of a learning purpose and supports continuous improvement. Data can provide a greater understanding of their strengths and weaknesses, or measuring and tracking progress over time to determine whether any activity or remediation was successful (and therefore should continue).

The questionnaire was developed as part of the monitoring activities of He Poutama Rangatahi. In the first year of this initiative, 19 providers were funded. As part of their service delivery they were asked to administer a questionnaire within their engagements with young people. The questionnaire was designed to assist the providers understand the strengths of young people and identify areas requiring further support. It was also intended to measure and track progress over time according to 20 indicators. Indicators would then comprise a composite score representing growth towards the ability to sustain employment.

The intended use of the questionnaire highlights the importance of providing accurate information. If the scores misrepresent a young person's ability to get into and sustain work then inappropriate decisions might be made that could have negative effects on them. For example, they may be placed into work too early to simply get them into a job, which may have a lasting negative impact on their future employment opportunities. Having inaccurate information may be worse than having no information.

HOW DID WE ASSESS THE QUESTIONNAIRE?

The critical factors believed to limit or enable young people's ability to get and maintain a job were derived initially over a six-month period, whereby New Zealand government agency officials engaged with community groups, iwi, employers and local agencies across four regions in New Zealand. These regions were Northland/Te Taitokerau, Eastern Bay of Plenty (Opotiki), Gisborne/Te Tairāwhiti and Hawke's Bay.

The identified critical factors included aspects about the individual such as: physical and mental health, drug and alcohol use, personal attitudes, literacy, numeracy, work experience and training, and having adequate housing. Other critical factors focused on the system, such the financial restrictions and the system-driven incentives/disincentives of employment, as well as accessibility to employment in relation to transport and networks. Further factors included community relationships and employer attitudes.

These factors were reviewed alongside research exploring the factors that influence employability. We searched for evidence of 'what works' in terms of young people accessing employment in New Zealand, as well as initiatives that were shown to enable employment. [1] [2] [3] [4] [5] [6]

The original questionnaire was developed from the community discussions and the research evidence. It aimed to measure the ability of young people to obtain and sustain employment – referred to here as 'ability' (our measure) – across the 20 items. Each item focused on one indicator, and included a five-point ordinal response scale describing levels of the indicator. The test instrument was workshopped with four Māori rangatahi, aged 15-24 years old, who provided feedback and ideas on each theme, and the language in relation to each point on the response scale.

We then discussed the instrument with community workers ('providers') engaged in the He Poutama Rangatahi initiative, who were working with young people (aged 15-24

[1] Smits, R. (2017). Kaikohe GROW: End of Project Report. Te Puni Kōkiri, Ministry for Primary Industries, Ministry of Social
[2] Auckland Co-Design Lab. (2016). The Attitude Gap Challenge: A South Auckland Employment and Skills Challenge. Auckland Council, Ministry of Business, Innovation and Employment, and Ministry of Social Development.
[3] Savage, C. (2016). Ako Whakaruruhau: Supporting Māori Apprenticeship Success Through Mentoring and Building Employer Capability. Summary Report. Ako Aotearoa: The National Centre for Tertiary Teaching Excellence.
[4] Ball, C., Crichton, S., Templeton, R., Tumen, S., Ota, R., and MacCormick, C. (2016). Characteristics of Children at Greater Risk of Poor Outcomes as Adults. The Treasury Analytical Paper 16/01.
[5] Crichton, S., Templeton, R. and Tumen, S. (2015). Using Integrated Administrative Data to Understand Children at Risk of Poor Outcomes as Young Adults. New Zealand Treasury Analytical Paper 15/01.
[6] Research New Zealand. (2016). Mid to Far North Employer Engagement. Prepared for Ministry for Primary Industries. MPI Technical Paper.

years) and helping them develop a pathway from education to employment. These providers were providing tailored support to each young person relevant to their specific contexts and within their community. Although the questionnaire was originally conceptualised as a self-assessment tool, the providers believed that some young people would require assistance. Administration procedures differed for the providers – some providers administered the questionnaire with individual young people, while others facilitated the questionnaire with groups of young people.

The data were collected over the initial few months of the initiative, and collated by providers into a provided spreadsheet. The subsequent data represent 547 young people, and were collected by eight providers within Northland/Te Taitokerau, Eastern Bay of Plenty (Opotiki) and Hawke's Bay. Of these responses, 18 of the 20 categories showed responses by at least 95% of the participants. Two items showed greater non-completion rates, cultural connection, which was added to the tool after data had begun being collected, and caregiving.[7]

The data were analysed as an assessment of the instrument. At this stage, it was important to use a model that could inform modifications to the questionnaire, and make completion easier for young people. Standard of Proof, with the support of the University of Western Australia, applied the Rasch unidimensional measurement model to check the psychometric properties and determine the extent to which there is sufficient justification for intended use and interpretation of the data derived from the questionnaire. Of even greater importance at this stage, the Rasch model produced diagnostic information about the questionnaire. The benefits of the Rasch model are that it identified where items were not performing as expected, and then allows the team to test modification to the questionnaire so that it can better measure the trait.

The analytical process was iterative, reflecting on the data and modifications alongside daily panel discussions with staff from both Standard of Proof and the University of Western Australia. The discussions focused on the questionnaire and initiative, alongside the graphical and statistical results (Rumm 2030 was used to apply the Rasch measurement model). When the model identified anomalies in the results in relation to the content of the items and definitions of the categories, these were discussed and different modifications were made iteratively alongside reviewing the results. The process refined the questionnaire so that it is both logical and supported by the model.

The Rasch model has the added advantage of providing interval level data for each person. Person locations identify the person's ability on a logit scale alongside the difficulty of our measure. Using the model also means we can measure growth better than other approaches when distributions show responses close to the maximum score on the scale. This level of data provides greater analytical power and opportunity to then test relationships between this score and potentially predictive factors, such as employment. We applied a T-Test to explore the relationship between these variables,

7 Although caregiving provided a response for those without such responsibilities, the options may have not been considered after the theme was deemed irrelevant to the individual.

providing initial evidence about the validity of the questionnaire as a measure relevant for employment, but this would need to be assessed further, specifically for predictive validity.

The analysis and subsequent modifications made to the questionnaire improved the reliability and made it easier for respondents.

The original scale included 20 items intending to measure the ability of individuals to achieve sustained employment. None of the 20 items were removed, as the analytical process showed that each item was measuring something unique, and therefore there were no redundant items (no item residual correlations exceeding 0.3). Further, each item contributed to the overall measure. Item Characteristic Curves showed that each item individually contributed to and differentiated the 'ability' of the young person. We tested removing items that showed limited ability to discriminate ability (e.g. 'housing') but removing items negatively impacted the reliability, which showed that they do work with other items to provide information. We therefore retained all 20 items.

The original scale included five responses across all 20 items, measuring degrees of ability for sustained employment. The application of the model showed collapsing specific response categories for 13 items improved the meaningful distinction between these different responses (e.g. people will more easily distinguish between response category 1 versus response category 2). Item Characteristic Curves showed that each item individually contributed to and differentiated the 'ability' of the young person. We tested removing items that showed limited ability to discriminate ability (e.g. 'housing') but removing items negatively impacted the reliability, which showed that they do work with other items to provide information.

Collapsing the response categories also improved the order of the response categories (i.e. response 1 is easier to achieve than a higher-level response 2). The probability of achieving or exceeding a higher order threshold increases in the order expected, as demonstrated in the Threshold Probability Curve. The task also improved the ability of the item to differentiate between those with higher and lower abilities. Category Probability Curves, Threshold Probability Curves, and Item Characteristic Curves similarly demonstrate increasing probabilities of achieving higher order response thresholds as person ability increases.

These 13 identified item responses were rescored, and results were compared statistically and discussed as any changes were made. The recommended modifications improved each item with near-null effects on the overall measure. The Rasch Person Separation Index of reliability decreased marginally from 0.881 (non-modified tool) to 0.874 (modified version of the tool). These new categories were simplified to reduce text and make language simpler, and in light of the experiences and feedback provided by regions.

All items need to measure the same thing, irrespective of the respondent. For instance, you want a scale to measure the weight of girls or boys to be the same, as it intends to

represent a construct – physical weight – separate from respondent gender. If the scale measures them differently they are not comparable. DIF was diagnosed statistically and graphically. For the variables age (14-18, 19-21, 22-25, older), gender (male and female) and region (Northland/Te Taitokerau, Eastern Bay of Plenty (Opotiki), Gisborne/Te Tairāwhiti and Hawke's Bay). One item was significant for age group (for persons at the same 'ability' level older persons scored higher in work experience than younger persons). This was a logical difference and expected to be a true measured effect and not a function of the instrument. The physical health item showed DIF for gender (for persons at the same 'ability' level males scored higher). The item was resolved to equate for these differences, but the statistics showed a reduction in reliability after the split so the original item was retained. It is believed naturally occurring regional differences can explain the DIF for access to employment, physical health and influential people.

Deleting or resolving items in this case would have improved fit but would have destroyed the item parameter invariance and in comparing means of groups or change of a group or individual, it would have changed the variable and ignored relevant features of the construct. Therefore, no items were deleted or resolved to account for DIF.

Some items, once modified, were clearly easier for young people to endorse than other items. Modifying the order may support young people in completing the questionnaire, but it can also guide providers by showing them those areas that are hardest for our young people to achieve. The relative easiness (to complete the questionnaire) was considered and the resultant order of the items changed to make completion easier for young people.

The results from the model showed a significant Chi-Square result overall ($x2 = 230.9998$, $df = 80$, $p<0.001$). Person fit residuals were good, with a mean near 0 (-0.24) and standard deviation near 1 (1.22), showing that the persons fit the model. The item fit residuals had a mean of 0.0826 and standard deviation of 2.1350; showing some item misfit. However, the misfit was not severe, as indicated by the Item Characteristic Curves and further testing of the impact on reliability of removing items. The Rasch Person Separation index of reliability is 0.8737 (very high). The traditional reliability index, Cronbach's coefficient, was also calculated after removing items with incomplete data and it was also very high (0.8757). The person 'ability' estimates were normally distributed with a marginal positive skew (mean on the logit scale was 0.938, standard deviation 0.966) when compared with the difficulty of the items, which were evenly spread across the difficulty scale (mean on the logit scale constrained to be 0, and standard deviation of 0.717).

Further analyses were also undertaken to examine the validity of the questionnaire in relation to employment. The aim of the questionnaire is to measure and monitor progress in the ability of young people to get and sustain a job. We cannot yet say that the questionnaire is 'predictive' of employment using our current data. However we were able to test the relationship between the questionnaire score (using the logit

scale from our modified questionnaire) and the young people's employment status (employed more than 10 hours per week versus not employed).

The evidence showed that employed young people had significantly greater overall scores than those who were not employed. Independent t-tests showed a significant difference ($t(680) = -18.7378$, $p < 0.0001$) between young people that were employed ($n = 61$, mean = 1.845, SE = 0.145) than those that were unemployed ($n=467$, mean=0.7977, SE = 0.1452), and the effect size was large ($r = 0.583$).This result provides initial evidence about the validity of the questionnaire as a relevant measure for employment.

WHAT DOES THIS MEAN?

The questionnaire can be reported as reliable.

Application of the model produced a reliable scale measuring a single construct with stable measurement properties of internal consistency and very high reliability.

The questionnaire measures a range of abilities relevant to young people aged 15-24 in New Zealand.

It is important to identify that the questionnaire can adequately measure the range of young people whom providers are enrolling, ranging from those that are 'most able' to those that are 'least able'. Otherwise, the scale could not adequately describe all our young people, nor would it likely be able to measure growth. The ability scale was found to adequately measure the range of ability for employment within our cohort, albeit the scale was a little easy for them.

The total score can be used to demonstrate and measure progress in overall 'ability'.

We assessed if the questionnaire measures a single concept. The ability to use a single measure (an overall score) of a young person's ability is a very useful and succinct way to assess ability at any point in time and measure change over time. The ability scale was found to measure a single concept, which we are theoretically interpreting here as *ability to sustain employment.* Multiple fit statistics were considered, including the overall summary statistics (overall item/trait chi-square statistic was significant) along with the fit residuals for persons and items. Further graphical evidence, such as through the Item Characteristic curves, were used to determine if the items added information to the overall discriminability of the tool. A review of the Principal Component Summary also showed no obvious 'elbow' in the eigen values.

Initial results show a strong relationship between the questionnaire total score and employment.

Although we cannot yet say that the questionnaire is 'predictive' of employment with our current data, we found that our measure first, has a very strong relationship with employment; and second, has some validity as a measure relevant for employment. It would be further recommended to test the predictive ability of the questionnaire over time.

Authorship and acknowledgements

Standard of Proof was created out of passion. Passion for making a difference, for making the lives of individuals, communities and organisations better through evidence. We pride ourselves on providing the highest standard of evidence that is appropriate and useful for the context.

Standard of Proof provides specialist services in:

- Evaluation: We encourage high standards of evidence, we promote relevant and inclusive processes, and we focus on informing decisions.
- Monitoring: We inspire progress through evidence, and we make quality data accessible.
- Measurement: We design, test and validate measures, and we enable efficient and accurate measurement practice.
- Research: We review, analyse and synthesise existing evidence, and we investigate new ideas and concepts.
- Data science and analytics: We apply statistical modelling, data analysis, and machine learning to unlock data.

The author of this short report is Dr Patricia Vermillion Peirce.

We worked with our partners at the University of Western Australia, namely Professor David Andrich and Ida Marais, to ensure the process was credible and quality assured.

We would like to extend our sincere gratitude to the communities that provided insights into the enablers and barriers to employment for young people, as well as He Poutama Rangatahi providers and rangatahi who tested and provided feedback to the instrument.

www.standardofproof.nz

Disclaimer

The instrument and data collected from it, are the perceptions of individuals about themselves. The information should be treated ethically, and within the legal bounds of privacy and information management. Such responsibilities fall upon the collector to ensure adequate consent has been gained for its intended use, and that the information is used accordingly.

Copyright and use

Standard of Proof (2021). *Youth Employment Pathway questionnaire 2.0: psychometric properties of a preliminary scale measuring the ability to obtain and retain employment.*


STANDARD OF PROOF
Evaluation Research Strategy