

Development and validation of MyLifeTracker: a routine outcome measure for youth mental health

Benjamin Kwan¹
Debra J Rickwood^{1,2}
Nic R Telford²

¹Faculty of Health, University of Canberra, Bruce, ACT, ²headspace National Youth Mental Health Foundation, Melbourne, VIC, Australia

Purpose: Routine outcome measures are now being designed for session-by-session use, with emphasis on clinically meaningful items and sensitivity to change. Despite an increasing mental health service focus for young people aged 12–25 years, there is a lack of outcome measures that are designed to be used across this age group. Consequently, MyLifeTracker (MLT) was developed as a brief mental health outcome measure designed for young people for routine use. It consists of the following five items targeting areas of importance to young people: general well-being, day-to-day activities, relationships with friends, relationships with family, and general coping.

Participants and methods: The measure was tested with 75,893 young people aged 12–25 years attending headspace centers across Australia for mental health-related issues.

Results: MLT showed a robust unidimensional factor structure and appropriate reliability. It exhibited good concurrent validity against well-validated measures of psychological distress, well-being, functioning, and life satisfaction. The measure was further demonstrated to be sensitive to change.

Conclusion: MLT provides a psychometrically sound mental health outcome measure for young people. The measure taps into items that are meaningful to young people and provides an additional clinical support tool for clinicians and clients during therapy. The measure is brief and easy to use and has been incorporated into an electronic system that routinely tracks session-by-session change and produces time-series charts for the ease of use and interpretation.

Keywords: MyLifeTracker, youth mental health, routine outcome measure, routine outcome monitoring, adolescent and young adult

Introduction

Routine outcome measurement (ROM) is now widely used in mental health services across Australia, New Zealand, Netherlands, Canada, Germany, the USA, Norway, and the UK.¹ This is in response to calls for greater quality assurance and demonstration of service effectiveness. Increasingly, decisions regarding government funding are based on evidence of cost-effective mental health outcomes.² Furthermore, there is expanding implementation of session-by-session ROM,³ whereby clients complete outcome measures at each session, targeting areas of cognition and emotion, functioning, and quality of life, with the aim of quantitatively informing clinicians about their clients' progress.⁴

Some concerns are consistently raised by stakeholders about the use and implementation of ROM. Studies have shown typically low completion rates of ROM, which

Correspondence: Benjamin Kwan
Faculty of Health, University of Canberra,
Building 12D22, Bruce, ACT 2601,
Australia
Email ben.kwan@canberra.edu.au

suggests a number of barriers in their clinical application.^{5–7} Clinicians have identified practical barriers, such as constraints around time and lack of training provided for the use of ROM.^{8–10} There are concerns around the lack of timely feedback of outcomes to clinicians, measures being meaningless to clients, and measures lacking sensitivity to change.^{11,12} Additionally, lack of administrative support and appropriate information technology support reduces the feasibility and acceptability of ROM.¹³ Consequently, services find it challenging to identify appropriate outcome measures for both clinical work and service evaluation.¹⁴

Young people, clinicians, and parents have identified through qualitative research on session-by-session ROM that measures need to be brief, sensitive to change, and meaningful to both young people and clinicians.^{10,15} Notably, measures administered electronically using computer devices were found to be engaging to young people, and timely feedback using computer-enabled visual representations to display outcomes was useful to both young people and clinicians; electronically administered outcome measures are now being designed for routine use, with increased focus on measure brevity and sensitivity to change.¹⁰ Training and support tools for clinicians are also being developed for the use of ROM.¹⁶ A number of measures have been designed for session-by-session use including the Outcome Rating Scale,¹⁷ Outcome Questionnaire-45 (OQ-45),¹⁸ Youth Outcome Questionnaire (Y-OQ),¹⁹ Clinical Outcomes in Routine Evaluation-Outcome Measure,²⁰ Session-by-Session Strengths and Difficulties Questionnaire (S × S-SDQ),²¹ and Symptoms and Functioning Severity Scale (SFSS).²²

Session-by-session ROM provides clinicians with regular feedback on their clients' progress, which enables appropriate decision making during treatment and can be used as a therapeutic tool to discuss progress in-session with a client.¹¹ A growing evidence base shows that ROM can improve communication between clinician and client, reduce client dropout, increase the accuracy of diagnosis, maintain positive outcomes for longer periods, and improve outcomes for those who are not improving or who are deteriorating.^{4,23–25} This suggests that ROM feedback could be particularly beneficial for population groups who are difficult to engage in mental health services.²⁶

A developmentally distinct group who have the highest rates of mental illness but the lowest levels of help-seeking are adolescents and young adults.^{27–29} Collectively termed “young people”, there is an increasing shift away from the historical categorization of adulthood starting at the age of 18 years, as a number of complex changes occur socially, physically, neurologically, and psychologically up to the age

of 25 years.³⁰ Traditional mental health services are organized around a demarcation between “Child and Adolescent Mental Health Services” (CAMHS) and a major transition to “Adult Mental Health Services” at the age of 18 years. It is now well understood that this service dichotomy is particularly inappropriate for mental health services, as it undermines continuity of care at a time of greatest need and heightened vulnerability.³¹

Mental health reform in Australia has reoriented mental health services to focus on early intervention for young people aged 12–25 years, and this approach has been gaining momentum in other countries.^{31,32} headspace is Australia's Youth Mental Health Foundation, which was initiated in 2006 to provide evidence-based early intervention mental health services to young people, implementing a model that breaks down the barriers to help-seeking.³³ Similarly, *Jigsaw* was a setup to strengthen the community's capacity to support young people aged 12–25 years in Ireland and there are related initiatives burgeoning worldwide.³⁴

With the current strong focus on service innovation in youth mental health, researchers are exploring how to measure and identify therapeutic change for young people.^{35–37} A systematic review of mental health outcome measures for young people aged 12–25 years did not identify any measures designed specifically for this age range.³⁷ The review focused on general mental health outcome measures, rather than single mental disorder-specific measures, and identified the target age group for each measure. Results identified 29 outcome measures that met the eligibility criteria including eight cognition and emotion measures, nine functioning measures, six quality of life measures, and six multidimensional measures.³⁷ It revealed that the design of outcome measures has been in line with the traditional split of mental health services between those younger than 18 years and those 18 years and older.³⁸

There were five measures within this review that were shown to have been used across the age range of 12–25 years, although these measures were originally developed and validated for an adult population. The review also found three ROMs used in this age group, the Y-OQ, S × S-SDQ, and SFSS, but all were designed to be used with children and adolescents and had not been validated with young adults older than 18 years.³⁷ The reorientation of mental health services means outcome measures specific to the youth transition period spanning adolescence and young adulthood is urgently needed to track outcomes relevant to young people.^{37,38} Additionally, ROM feedback provides another therapeutic tool for supporting this highly vulnerable age group.³⁷

Development of MyLifeTracker (MLT)

headspace National Youth Mental Health Foundation is an Australian innovation in youth mental health services offering mental health, alcohol and other drugs, health care, and vocational services to young people aged 12–25 years through a primary care platform. Established with funding from the Australian government in 2006, by 2018, there will be >100 headspace centers across Australia.³³ The centers prioritize an early intervention approach to youth mental health and are designed to redress the barriers to help-seeking that young people experience, by being no or low cost, youth focused, inclusive, and nonstigmatizing.³⁶ The centers are supported by an online service, eheadspace, which facilitates even earlier intervention via online chat, email, and a phone service.

In 2013, headspace centers implemented an innovative data collection system that collects information at each occasion of service for a young person.³⁶ The data set has multiple purposes, including service monitoring and evaluation and providing therapeutically useful information for clinicians. A particular challenge, however, was determining appropriate mental health outcome measures for use with this age range, as there were none available that targeted those aged 12–25 years or that were appropriate across a diverse range of mental health presentations.^{37,38} In response, headspace developed a measure called “MyLifeTracker” to supplement the other measures being used in the data collection system.

MLT is a brief mental health outcome measure designed for young people aged 12–25 years. Its items measure current self-reported quality of life in the following five different areas of importance to young people: general well-being, day-to-day activities, relationships with friends, relationships with family, and general coping. The measure is intended to be used at every session. It is completed electronically on a tablet or desktop device, and the information is immediately available to the clinician in the visual representation of graphs over time. The measure was developed specifically for use in headspace through a literature review; consultation with clinical staff, clients, and the headspace youth reference group; and pilot testing. The aim was to develop an outcome measure that reflected the important areas of life for young people aged 12–25 years and that could be used for routine monitoring of client progress by being responsive to change.

The current study examines the psychometric properties of MLT to determine its reliability, validity, sensitivity to change, and whether it is appropriate for the entire age range from 12 to 25 years and for both males and females. It was hypothesized that the five items constituting MLT would represent a single factor and that it would show good internal

consistency. MLT was expected to display adequate test–retest reliability and show appropriate concurrent validity with standardized mental health outcome measures. Finally, it was hypothesized that MLT would display appropriate sensitivity to change over time.

Participants and methods

Participants

Participants were 75,893 adolescents and young adults between the ages of 12 and 25 years who commenced their first episode of care at a headspace center for a mental health-related issue between July 1, 2015 and March 31, 2017. A total of 25.3% of the participants were in early adolescence (12–14 years), 31.9% of the participants were in mid-adolescence (15–17 years), 28.3% of the participants were in late adolescence (18–21 years), and 14.4% of the participants were young adults (22–25 years). There were more females (58.3%) than males (40.4%), and 1.3% of them reported a nonbinary gender option. Participants were predominantly from major cities (66.0%), followed by inner regional areas (24.5%), outer regional areas (8.4%), and remote areas (1.1%). A total of 8.5% of the participants were Aboriginal and Torres Strait Islander young people. A total of 8.9% of the participants were from culturally and linguistically diverse backgrounds, and 17.6% of the participants identified their sexuality as lesbian, gay, bisexual, other sexualities, or questioning.

At initial presentation, the majority of participants had not been assessed or did not meet threshold diagnostic criteria for a mental disorder; however, primary presenting issue was reported by the client and clinician. As the service is aimed at early intervention, young people who are at-risk or have subclinical presentations are encouraged to present. The initial primary presenting issues include depressive symptoms (29.6%), anxiety symptoms (28.6%), anger issues (7.4%), stress related (5.0%), conflict in home environment (4.6%), difficulty with personal relationships (3.9%), alcohol and other drug use (2.8%), behavioral problems (2.5%), suicidal thoughts or behaviors (2.3%), trauma (2.2%), grief (1.6%), bullying (1.3%), and a range of other mental health symptoms and situational stressors (≤1%).

Procedure

On presentation for each occasion of service at a headspace center, young people were given access to an electronic tablet device or a desktop computer where they were provided with information about the data collection system and consent to answering a series of questions, which varied depending

on the occasion of service. The time taken to complete the questions varied according to visit as different question sets were presented at different times. The initial questions at first presentation took up to 15 minutes to complete; questions at subsequent visits generally took <5 minutes. Clinicians also completed a series of questions related to each occasion of service after the young person's session. The data were encrypted and uploaded to a national data warehouse, which is used for monitoring and evaluation. The data set includes demographic characteristics, clinical presentation, and outcome measures.

Ethics approval was obtained through quality assurance processes, comprising initial consideration and approval through the headspace Board research subcommittee. The consent processes were reviewed and endorsed by an independent body, the Australasian Human Research Ethics Consultancy Services. Follow-up data collection processes were approved by Melbourne Health Quality Assurance ethics. The data were collected primarily for service provision and quality improvement purposes, and parental consent was not routinely sought for those younger than 18 years, who were assessed as mature minors unless otherwise indicated. Young people were able to opt out of data collection if they chose, or at the discretion of their parents.

Measures

Demographic characteristics comprised age, gender, sexual orientation, Aboriginal and/or Torres Strait Islander background, cultural and linguistic backgrounds, living situation, and current work/study status. Client clinical presentation was determined by self-reported reasons for presentation and clinician-reported presenting issues.

Outcome measures

Routine outcome monitoring measure

MLT is a five-item self-report measure to assess current quality of life in areas of importance to young people. It asks young people to self-report how they have been feeling over the last week, or, since their last headspace session, in relation to their "general well-being (emotional, physical, spiritual)", "day-to-day activities (study, work, leisure, self-care)", "relationships with friends", "relationships with family", and "coping (dealing with life, using your strengths)". Responses are given on a sliding scale anchored at 0 and 100 with the chosen score visible; the anchors are accompanied by a visual analog of a sad and happy face. Total MLT scores were calculated by averaging across the five items, ranging from 0 to 100, with a higher score indicating a higher quality of life.

Psychological distress

The Kessler Psychological Distress Scale 10 (K10)³⁹ is a 10-item self-report measure of nonspecific psychological distress in the anxiety-depression spectrum. An example item is "In the last 4 weeks, how often did you feel hopeless?" The response format ranged from 1 (none of the time) to 5 (all of the time) for each item. Scores are summed, and the total score ranged from 10 to 50, with higher scores indicating more distress. The K10 has been validated for use in the general Australian population.⁴⁰ Cronbach's α was 0.91 in the present study.

Life satisfaction

Life satisfaction is a seven-item self-report measure that was adapted from the Brief Multidimensional Students' Life Satisfaction Scale from the Peabody Treatment Progress Battery.²² The questions asked young people to indicate how satisfied they are with different areas of their life, in relation to "family life", "friendships", "romantic relationships", "school/work experience", "yourself", "where you live", and "life overall". The response format was a 0–10 scale for each item, 0 (worst possible) to 10 (best possible). Total life satisfaction scores were calculated by averaging across the seven items, ranging from 0 to 10, with a higher score indicating more positive life satisfaction. The original measure was validated with young people aged 11–18 years in the USA and had a Cronbach's α of 0.77.²² Cronbach's α was also 0.77 in the present study.

Well-being

Three items taken from the 14-item Warwick–Edinburgh Mental Well-Being Scale (WEMWBS)⁴¹ were used to measure self-reported positive mental health and well-being. The three items were selected through consultation with headspace young people and staff. The questions asked young people to rate the frequency of their feelings and thoughts over the last 2 weeks, in relation to "I've been interested in new things", "I've been feeling useful", and "I've been feeling good about myself". Responses were on a 5-point scale from 1 (none of the time) to 5 (all of the time). Total well-being scores were calculated by averaging across the three items, ranging from 1 to 5, with a higher score indicating more positive well-being. WEMWBS has been validated with teenagers in England and Scotland revealing a Cronbach's α of 0.87.⁴² Cronbach's α was 0.72 in the present study.

Social and occupational functioning

The Social and Occupational Functioning Assessment Scale (SOFAS)⁴³ is a single-item, clinician-report assessment of a

client's current functioning at the time the evaluation is made, determined irrespective of the severity of the client's psychological symptoms. Scores range from 1 to 100, with higher scores representing increased functioning across a variety of activities. For example, a score of 1–10 is indicative of a persistent inability to maintain minimal personal hygiene; a score of 41–50 is indicative of serious impairment in social, occupational, or school functioning; and a score of 91–100 is indicative of superior functioning in a wide range of activities. The SOFAS is derived from the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) Global Assessment of Functioning Scale (Axis V), which has shown good internal consistency with the Cronbach's α of 0.80.⁴³

Data analyses

First, we conducted an exploratory factor analysis (EFA) of MLT items using maximum likelihood factoring. A confirmatory factor analysis (CFA) was then conducted to replicate the factor structure of MLT. The baseline data set was randomly split into two, with the EFA performed on one half and the CFA performed on the other. As the Chi-square test is highly sensitive to sample size, the following fit indices were used for the CFA to assess the model fit: root mean square error of approximation (RMSEA < 0.06, with the lower bound of its 90% CI < 0.05 to indicate "close" fit), the comparative fit index (CFI > 0.95), the Tucker–Lewis Index (TLI > 0.95), and the Standardized Root Mean Square Residual (SRMR < 0.08).⁴⁴ Additionally, a multi-group CFA was performed to test measurement invariance across age and gender groups.⁴⁵ The following indices have been recommended: CFI > 0.95 for configural invariance and $\Delta\text{CFI} \leq 0.01$ for metric invariance.⁴⁶

Internal consistency of MLT using the first visit data was evaluated using the Cronbach's α coefficient. Pearson's correlation analysis was conducted to examine the test–retest reliability of MLT scores across the first and second headspace visits that were within a 2-week time period. Internal consistency and test–retest reliability were further examined as a function of age group and gender (male/female).

Concurrent validity was examined using a Pearson's correlation analysis, which explored the associations of MLT scores at participants' initial session with the standardized outcome measures of psychological distress, life satisfaction, well-being, and social and occupational functioning. Concurrent validity was further examined as a function of age group and gender.

Sensitivity to change was examined using a number of methods including effect sizes,⁴⁷ Reliable Change Index (RCI),⁴⁸ and growth curve modeling.⁴⁹ Effect sizes were

estimated for participants' average MLT change scores between the first and last sessions. RCIs were calculated between participants' first and last MLT scores, and percentages of participants who improved, did not change, or deteriorated were reported. RCIs were also calculated for K10 scores to allow comparisons with MLT. RCI is an index that was developed to detect the minimum reliable amount of change in scores while accounting for measurement error.⁴⁸ Effect sizes and RCIs were further examined as a function of age group and gender.

Growth curve modeling estimated average rates of change in MLT scores across participants' episode of care. This approach was utilized as it can estimate a mean intercept and mean slope of the pooled sample trajectory (within-person patterns).⁴⁹ Maximum likelihood estimation procedures were used. Weeks in treatment were used over session number as time between sessions varied among participants, and this approach has been recommended in the literature exploring youth psychotherapy change.⁵⁰ Only cases below the 90th percentile (26 weeks) of treatment length were used to avoid effects of outliers. Inspection of MLT data suggested a cubic pattern rather than a linear pattern over time (Bayesian Information Criterion [BIC]_{linear} = 1892911, BIC_{cubic} = 1890673, and BIC_Δ = 2238, where smaller BICs indicate a better model fit). Growth curve modeling was further examined as a function of age group and gender. All analyses were conducted using SPSS Version 23 and Amos Version 23.

Results

Descriptives

Table 1 provides descriptive statistics for MLT scores and the standard outcome measures at first headspace visit. It reveals that females generally reported greater psychological distress, lower life satisfaction, and lower well-being than males ($P < 0.001$, $\eta_p^2 = 0.04$, 0.01, and 0.02, respectively). In addition, young adults generally reported greater psychological distress, lower life satisfaction, and lower well-being than adolescents ($P < 0.001$, $\eta_p^2 = 0.07$, 0.06, and 0.03, respectively). These trends across age and gender groups were also seen with the MLT scores ($P < 0.001$, $\eta_p^2 = 0.06$ and 0.03, respectively). Clinician reports of participants' social and occupational functioning appeared constant across gender and age group ($P < 0.001$, $\eta_p^2 = 0.004$ and 0.001, respectively).

Exploratory factor analysis

The eigenvalues generated by the EFA were 3.17, 0.67, 0.64, 0.32, and 0.21. Standardized factor loadings for the five MLT items were uniformly strong and positive (range 0.53–0.90).

The one factor solution explained 56% of the variance. The eigenvalues and factor loadings supported a unidimensional solution for MLT.

Confirmatory factor analysis

Fit indices obtained from the CFA indicated an excellent level-of-fit of the data to the single-factor model: RMSEA=0.049, 90% CI (0.045, 0.053), CFI=0.995, TLI=0.990, and SRMR=0.015. Measurement invariance tests across age and gender groups indicated that MLT displayed configural invariance (CFI=0.990) and metric invariance (Δ CFI=0.006).

Internal consistency

MLT showed good internal consistency with Cronbach's α =0.84. Internal consistency was further examined as a function of age group and gender, and Cronbach's α ranged from 0.79 to 0.86.

Test-retest reliability

Due to variability of time between participants' initial and second sessions, only a test-retest period within a 2-week

time frame was applied (n=18890). MLT test-retest reliability showed a coefficient of 0.78. Test-retest reliability was further examined as a function of participant age group and gender. Coefficients were highly similar, ranging from 0.74 to 0.78 across the age group and gender combinations.

Concurrent validity

Table 2 provides the correlation coefficients for the relationships of MLT with the standardized outcome measures. MLT was strongly negatively correlated with psychological distress and strongly positively correlated with life satisfaction and well-being. It was also positively correlated with social and occupational functioning, but these associations were weak in strength. The correlation coefficients between MLT and the standardized outcome measures were consistent in strength across age and gender groups.

Sensitivity to change

The number of sessions attended by participants ranged from 1 to 30 sessions, with an average attendance of 4.16 sessions (SD=4.37). Change was determined by comparing the first and last session scores. Overall, participants demonstrated

Table 1 Descriptive statistics for MyLifeTracker and standardized outcome measures at first headspace visit, by age group and gender

Outcome measures	12–14 years, mean (SD)	15–17 years, mean (SD)	18–21 years, mean (SD)	22–25 years, mean (SD)
Males	n=6243 mean (age)=13.07	n=7772 mean (age)=16.00	n=7451 mean (age)=19.37	n=3903 mean (age)=23.09
MyLifeTracker	61.88 (21.14)	52.70 (20.17)	45.78 (19.51)	44.35 (19.33)
Psychological distress	21.63 (8.19)	25.79 (8.75)	29.05 (8.63)	29.51 (8.60)
Life satisfaction	6.54 (1.83)	5.80 (1.79)	5.09 (1.84)	4.85 (1.86)
Well-being	3.10 (0.83)	2.77 (0.78)	2.61 (0.77)	2.60 (0.77)
Social and occupational functioning	63.50 (17.09)	62.87 (16.37)	60.77 (16.82)	61.06 (17.06)
Females	n=9818 mean (age)=13.21	n=12,474 mean (age)=15.93	n=10,000 mean (age)=19.36	n=4786 mean (age)=23.03
MyLifeTracker	49.21 (20.51)	43.45 (18.09)	41.05 (17.80)	41.73 (18.48)
Psychological distress	27.08 (9.11)	30.36 (8.49)	31.54 (8.26)	30.96 (8.24)
Life satisfaction	5.69 (1.82)	5.29 (1.69)	5.04 (1.72)	4.98 (1.75)
Well-being	2.64 (0.80)	2.48 (0.73)	2.46 (0.73)	2.50 (0.74)
Social and occupational functioning	63.70 (16.25)	62.96 (16.06)	62.01 (16.59)	62.37 (17.11)

Table 2 Pearson correlation coefficients between MyLifeTracker and standardized outcome measures, by age group and gender

	12–14 years				15–17 years				18–21 years				22–25 years				All ages			
	PD	LS	WB	SOF	PD	LS	WB	SOF	PD	LS	WB	SOF	PD	LS	WB	SOF	PD	LS	WB	SOF
Males																				
MyLifeTracker	–0.61	0.72	0.61	0.17	–0.65	0.73	0.63	0.19	–0.66	0.74	0.60	0.21	–0.67	0.71	0.64	0.26	–0.68	0.75	0.64	0.21
Females																				
MyLifeTracker	–0.66	0.74	0.63	0.20	–0.64	0.72	0.60	0.19	–0.64	0.71	0.60	0.24	–0.65	0.70	0.60	0.26	–0.66	0.72	0.61	0.22
All participants																				
MyLifeTracker	–0.67	0.74	0.65	0.18	–0.66	0.73	0.63	0.18	–0.66	0.72	0.61	0.22	–0.66	0.70	0.62	0.25	–0.68	0.74	0.64	0.21

Note: All $P < 0.05$.

Abbreviations: PD, psychological distress; LS, life satisfaction; WB, well-being; SOF, social and occupational functioning.

change with a moderate effect size ($d=0.52$). Effect sizes were similar across the age group and gender combinations, with Cohen's d ranging from 0.47 to 0.58.

RCI was estimated as an 18.27-point change for MLT, and Table 3 provides the percentages of young people showing reliable change across age group and gender. The proportions of reliable change for the K10 are also provided in Table 4 for comparison with MLT. RCI was estimated as a 7-point change for the K10 based on previous research.³⁶ MLT showed more participants reliably improving compared to the K10 (−1.5%) and fewer participants reliably deteriorating than the K10 (+2.5%). Therefore, MLT detected more

Table 3 Proportion of young people showing reliable change for MyLifeTracker between first and last session ratings, by age group and gender

	Change category (%)		
	Improvement	No change	Deterioration
12–14 years			
Male (n=4066)	29.0	65.4	5.6
Female (n=6864)	29.2	65.3	5.5
15–17 years			
Male (n=4853)	28.4	66.7	4.9
Female (n=8314)	25.8	68.9	5.3
18–21 years			
Male (n=4641)	28.0	67.3	4.7
Female (n=6481)	28.4	67.0	4.6
22–25 years			
Male (n=2433)	30.2	65.8	4.0
Female (n=3092)	30.6	65.1	4.3
All participants (n=41,444)	28.2	66.8	5.0

Table 4 Proportion of young people showing reliable change for K10 between first and last session ratings, by age group and gender

	Change category (%)		
	Improvement	No change	Deterioration
12–14 years			
Male (n=4064)	23.7	68.2	8.1
Female (n=6871)	26.1	64.0	9.9
15–17 years			
Male (n=4838)	25.5	66.5	8.0
Female (n=8319)	25.8	66.5	7.7
18–21 years			
Male (n=4643)	27.9	65.8	6.3
Female (n=6474)	28.3	65.7	6.0
22–25 years			
Male (n=2432)	29.3	64.8	5.8
Female (n=3092)	30.0	64.1	5.9
All participants (n=41,439)	26.7	65.8	7.5

Abbreviation: K10, Kessler Psychological Distress Scale 10.

improvement and less deterioration among the participants when compared with the K10.

Change in MLT followed a cubic pattern, which had a significant contribution in the model ($P<0.001$). The initial status of MLT was -46.0 ($\beta=46.53$, standard error [SE]=0.09, $P<0.001$). The positive effect of linear growth ($\beta=1.48$, SE=0.02, $P<0.001$) suggested that the MLT increased at the beginning; the negative effect of the quadratic growth ($\beta=-0.08$, SE=0.003, $P<0.001$) indicated a deceleration in the rate of change; and the positive effect of cubic growth ($\beta=0.002$, SE=0.001, $P<0.001$) revealed that such deceleration gradually diminished over time. As shown in Figure 1, the average rate of change during early therapy was relatively steep but began to flatten out between weeks 10 and 20; then, between weeks 21 and 26, there was another slight upward trend.

Age and gender groups were examined as a time-invariant covariate to explore any group differences in change over time. Although the initial scores for MLT were significantly different across age and gender combinations ($\beta=-2.36$, SE=0.05, $P<0.001$), as shown in Table 1, the shape of the change did not differ significantly across age and gender groups ($P>0.05$).

Discussion

The current study aimed to examine the psychometric properties of MLT to determine its reliability, validity, sensitivity to change, and whether it is appropriate for the entire age range from 12 to 25 years and for both males and females.

A unidimensional factor structure was determined using EFA and, then, confirmed through a CFA. This provides support for clinicians and services to use MLT as a single total score to track mental health outcomes. Importantly, the measure was shown to display measurement invariance across age and gender combinations. The reliability of MLT was demonstrated through good internal consistency and robust test–retest reliability.

Evidence for the concurrent validity of MLT was found, as correlations between MLT and standardized outcome measures were as hypothesized. Specifically, MLT was strongly related to psychological distress, life satisfaction, and well-being. Weak correlation coefficients between the client-reported MLT and the clinician-rated SOFAS argue in favor of the necessity of implementing routine client-reported outcome measures rather than relying solely on clinician subjective judgment of treatment progress.^{8,11,12} Past research has shown significant differences between self-report versus clinician-reported ratings of similar measures.^{51,52} Similar correlation coefficients were evident

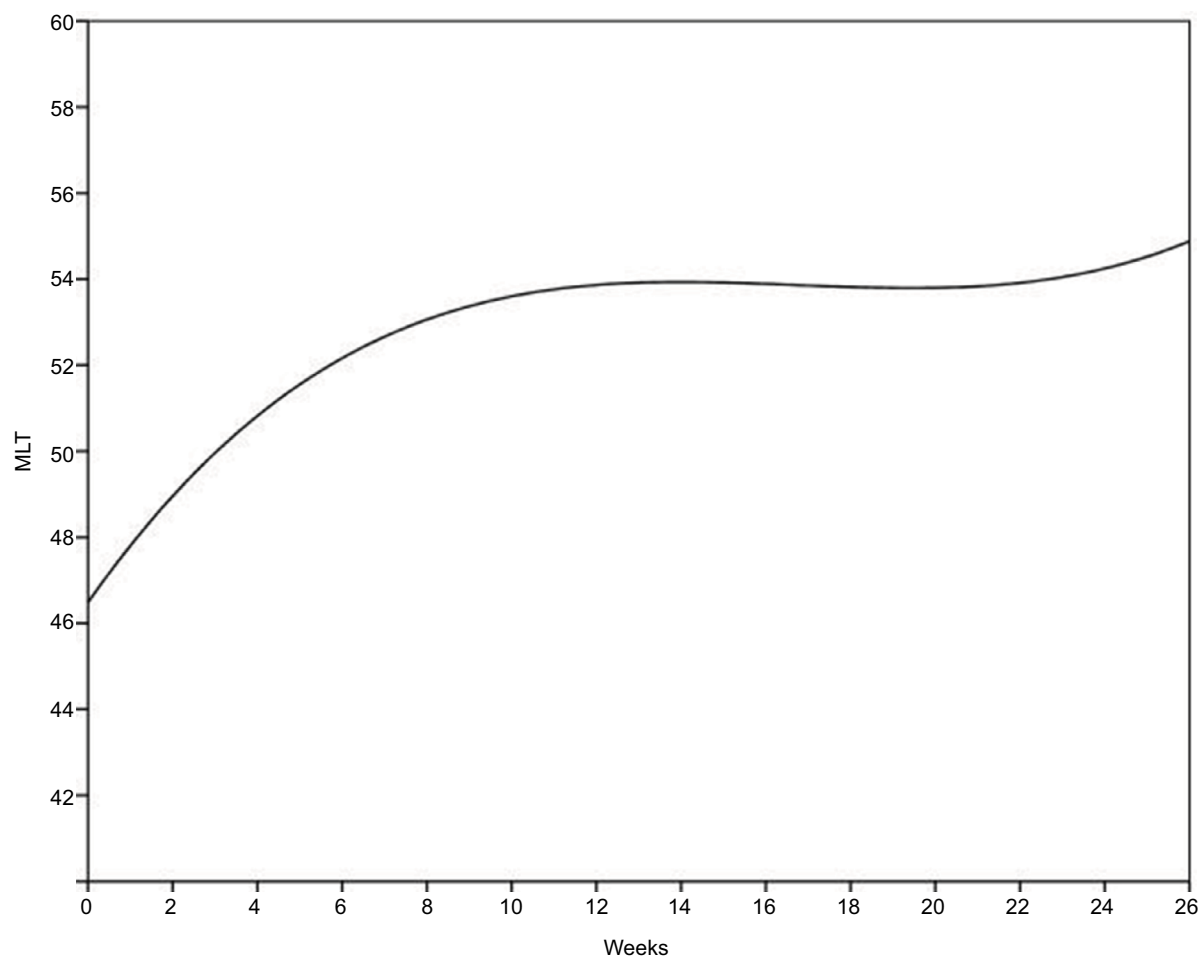


Figure 1 Average rate of change in MLT across weeks in treatment.
Abbreviation: MLT, MyLifeTracker.

for various age groups and for males and females, further demonstrating concurrent validity across age and gender for MLT.

The sensitivity to change of MLT was shown via significant improvements that were observed for males and females in all age groups. Importantly, these changes were moderate in effect size. RCI proportions across MLT and the K10 were generally comparable, but MLT showed higher proportions of participants improving and lower proportions of participants deteriorating when compared with the K10. It should be noted, however, that some participants would not be able to reliably change as their initial score may be too close to the upper or lower limits of the measure (eg, if a participant's initial score on MLT was 82, they could not reliably improve by the required 18.27 points as the upper limit of the measure is 100).

MLT RCI percentages and effect sizes are comparable to those reported from other naturalistic youth mental health settings. A study in the Netherlands explored naturalistic

change in a youth mental health clinic targeting young adults aged 16–23 years attending an average of 7.5 sessions. The results showed an overall effect size of 0.55 and 19% of clients reliably improving when rated on the Global Assessment of Functioning scale.⁵³ A study exploring change in parent-reported SDQ in England CAMHS showed an overall effect size of $d=0.40$, and reliable improvement was seen among 16.5% of clients after a 4- to 8-month period.⁵⁴ Australian data from acute inpatient CAMHS from 2006 to 2007, using the Health of the Nation Outcome Scales for Children and Adolescents, showed 23.8% reliable improvement and 1.8% reliable deterioration from admission to discharge.⁵⁵

The growth curve pattern demonstrated by MLT showed a steep increase during the early stages of therapy, subsequent flattening out, and then a slight increase again later in therapy. There were similar rates of change across age and gender groups. This cubic pattern of change has been observed in other research with children and adolescents using the parent-reported Y-OQ.⁵⁰ Similarly, this pattern of

change was seen with clients from a university counseling center using the OQ-45, which is a self-report measure for adults.⁵⁶ A theory for this pattern of change is the “good enough level” model, whereby clients are said to change at different rates and those that have reached sufficient change terminate treatment, and therefore, a plateau is seen within the growth curve.⁵⁶ Interestingly, there is some evidence that gains, particularly sudden gains, occur earlier in the treatment for younger clients. For example, Gaynor et al⁵⁷ found that in children and adolescents, 85% of sudden gains occurred before session five and all sudden gains by the 10th session; whereas for an adult sample, only half of sudden gains had occurred by session five.⁵⁸ As such, rates of change shown by MLT is comparable to other well-validated measures in similar youth-based settings; however, due to the naturalistic clinical setting in the current sample, MLT could not be compared with a control group.

This study offers evidence for the validity of MLT as a brief and practical measure that can be used in routine outcome monitoring with adolescents and young adults. Information is quickly and easily obtained from young people and able to be displayed graphically for clinicians to track treatment progress. MLT fills a gap in mental health outcome measures for young people aged 12–25 years.³⁷ The measure has been specifically designed for and now validated with this age group, which is preferable to using measures designed for adult populations. MLT targets valued parts of a young person’s life including relationships, coping, and well-being, rather than focusing on traditional measures of diagnostic symptoms and functioning. The shift of mental health service provision to the transition period of adolescence and early adulthood means outcome measures for this age range are vital, and MLT has been shown to be psychometrically sound across this entire age range. The measure enables clinicians working with young people to gain a quick regular snapshot of overall client progress and provides a valid measure to use to assess service effectiveness.

Furthermore, each item on MLT targets current quality of life in areas of importance to young people themselves, including general well-being, day-to-day activities, relationships with friends, relationships with family, and coping. These areas were developed through extensive consultation with headspace clients and their clinicians and shown to be meaningful to young people and clinically useful to clinicians. While the items are all tapping into a single dimension, the level of change within each item can identify areas where therapy is being effective or not – for example, relationships and coping skills. If outcomes are routinely

monitored by clinicians and used as a therapeutic tool with clients, this can improve outcomes, specifically for identifying clients who are deteriorating or not improving.^{4,12,23,59} Such monitoring provides clinicians with another source of information on a session-by-session basis to help with treatment planning.

However, the current results should be considered in light of several limitations. First, the study was conducted in Australia within an early intervention youth mental health setting and the psychometric properties of MLT may not be generalizable to other settings, such as public mental health services, or other age groups, such as children younger than 12 years or older adults. Further testing is required to support it being used more broadly. Another limitation is that the current study has not provided norms or clinically significant change values. This would provide a general benchmark for clinicians to compare against when providing treatment to young people and enhance the practical utility of the measure to aid in its interpretation.

Further research should explore clinicians’ and young people’s views of MLT to confirm that it measures meaningful changes. A qualitative approach has been suggested to better understand this developmentally sensitive construct of therapeutic change.³⁵ Furthermore, future research should consider the clinical utility of MLT as a feedback system in routine outcome monitoring to determine how clinicians and clients can use the measure to discuss outcomes and areas of need for targeted intervention.

The push for ROM should not only be motivated by a need for increased quality assurance and effectiveness but also provide clinically useful and meaningful tools for clinicians and clients to enhance the therapeutic process. MLT provides a psychometrically sound mental health outcome measure for young people aged 12–25 years. The measure is brief and easy to use with a young person. It can be incorporated into an electronic system that routinely tracks session-by-session change and produce time-series charts for the ease of use and interpretation. MLT taps into areas that are meaningful to young people and provides an additional support to clinicians and clients during therapy.

Author contributions

BK completed the literature review, analyzed the data, and drafted the article. DJR supervised the design, reviewed the results, and revised the article. NRT supervised the design and reviewed the results. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work.

Disclosure

All authors are either employed by headspace National or a headspace center. The authors report no other conflicts of interest in this work.

References

- Roe D, Drake RE, Slade M. Routine outcome monitoring: an international endeavour. *Int Rev Psychiatry*. 2015;27(4):257–260.
- Coombs T, Stapley K, Pirkis J. The multiple uses of routine mental health outcome measures in Australia and New Zealand: experiences from the field. *Australas Psychiatry*. 2011;19(3):247–253.
- Wolpert M, Fugard AJ, Deighton J, Gözrig A. Routine outcomes monitoring as part of children and young people's Improving Access to Psychological Therapies (CYP IAPT): improving care or unhelpful burden? *Child Adolesc Mental Health*. 2012;17(3):129–130.
- Bickman L, Kelley SD, Breda C, de Andrade AR, Riemer M. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. *Psychiatr Serv*. 2011;62(12):1423–1429.
- Batty MJ, Moldavsky M, Foroushani PS, et al. Implementing routine outcome measures in child and adolescent mental health services: from present to future practice. *Child Adolesc Mental Health*. 2013;18(2):82–87.
- Hatfield DR, Ogles BM. The use of outcome measures by psychologists in clinical practice. *Prof Psychol Res Pract*. 2004;35(5):485–491.
- Johnston C, Gowers S. Routine outcome measurement: a survey of UK child and adolescent mental health services. *Child Adolesc Mental Health*. 2005;10(3):133–139.
- Boswell JF, Kraus DR, Miller SD, Lambert MJ. Implementing routine outcome monitoring in clinical practice: benefits, challenges, and solutions. *Psychother Res*. 2015;25(1):6–19.
- Hall CL, Moldavsky M, Taylor J, et al. Implementation of routine outcome measurement in child and adolescent mental health services in the United Kingdom: a critical perspective. *Eur Child Adolesc Psychiatry*. 2014;23(4):239.
- Hall CL, Taylor J, Moldavsky M, et al. A qualitative process evaluation of electronic session-by-session outcome measurement in child and adolescent mental health services. *BMC Psychiatry*. 2014;14(1):113.
- Bickman L. A measurement feedback system (MFS) is necessary to improve mental health outcomes. *J Am Acad Child Adolesc Psychiatry*. 2008;47(10):1114.
- Kelley SD, Bickman L. Beyond outcomes monitoring: measurement feedback systems (MFS) in child and adolescent clinical practice. *Curr Opin Psychiatry*. 2009;22(4):363.
- Lucock M, Halstead J, Leach C, et al. A mixed-method investigation of patient monitoring and enhanced feedback in routine practice: barriers and facilitators. *Psychother Res*. 2015;25(6):633–646.
- Wolpert M, Cheng H, Deighton J. Measurement issues: review of four patient reported outcome measures: SDQ, RCADS, C/ORS and GBO—their strengths and limitations for clinical use and service evaluation. *Child Adolesc Mental Health*. 2015;20(1):63–70.
- Merry S, Stasiak K, Parkin A, et al. *Child and Youth Outcome Measures: Examining Current Use and Acceptability of Measures in Mental Health Services and Recommending Future Directions*. Auckland, New Zealand: Health Research Council; 2004.
- Law D, Wolpert M. *Guide to Using Outcomes and Feedback Tools with Children, Young People and Families*. North London: CAMHS Press; 2014.
- Miller SD, Duncan B, Brown J, Sparks J, Claud D. The outcome rating scale: a preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *J Brief Ther*. 2003;2(2):91–100.
- Lambert MJ, Hansen NB, Harmon SC. Outcome Questionnaire System (The OQ System): development and practical applications in healthcare settings. *Developing and Delivering Practice-Based Evidence: A Guide for the Psychological Therapies*. 2010:141–154.
- Ridge NW, Warren JS, Burlingame GM, Wells MG, Tumbliin KM. Reliability and validity of the youth outcome questionnaire self-report. *J Clin Psychol*. 2009;65(10):1115–1126.
- Evans JM-C, Margison F, Barkham M, Audin K, Connell J, Graeme McGrath C. CORE: clinical outcomes in routine evaluation. *J Mental Health*. 2000;9(3):247–255.
- Hall CL, Moldavsky M, Taylor J, et al. Innovations in practice: piloting electronic session-by-session monitoring in child and adolescent mental health services: a preliminary study. *Child Adolesc Mental Health*. 2015;20(3):171–174.
- Bickman L, Athay M, Riemer M, et al. *Manual of the Peabody Treatment Progress Battery*. Nashville: Vanderbilt University; 2010.
- Carlier IV, Meuldijk D, Van Vliet IM, Van Fenema E, Van der Wee NJ, Zitman FG. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. *J Eval Clin Pract*. 2012;18(1):104–110.
- Lambert MJ, Shimokawa K. Collecting client feedback. *Psychotherapy*. 2011;48(1):72–79.
- Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *J Consult Clin Psychol*. 2010;78(3):298–311.
- Lutz W, De Jong K, Rubel J. Patient-focused and feedback research in psychotherapy: where are we and where do we want to go? *Psychother Res*. 2015;25(6):625–632.
- Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. *Lancet*. 2013;382(9904):1575–1586.
- Costello EJ, He JP, Sampson NA, Kessler RC, Merikangas KR. Services for Adolescents with psychiatric disorders: 12-month data from the National Comorbidity Survey–Adolescent. *Psychiatr Serv*. 2014;65(3):359–366.
- Copeland WE, Shanahan L, Davis M, Burns BJ, Angold A, Costello EJ. Increase in untreated cases of psychiatric disorders during the transition to adulthood. *Psychiatr Serv*. 2015;66(4):397–403.
- Patton GC, Viner R. Pubertal transitions in health. *Lancet*. 2007;369(9567):1130–1139.
- McGorry PD, Goldstone SD, Parker AG, Rickwood DJ, Hickie IB. Cultures for mental health care of young people: an Australian blueprint for reform. *Lancet Psychiatry*. 2014;1(7):559–568.
- Hetrick S, Bailey A, Smith K, et al. Integrated (one-stop shop) youth health care: best available evidence and future directions. *Med J Austr*. 2017;207(10 Suppl):S5–S18.
- Rickwood DJ, Telford NR, Parker AG, Tanti CJ, McGorry PD. Headspace—Australia's innovation in youth mental health: who are the clients and why are they presenting? *Med J Aust*. 2014;200(2):108–111.
- McGorry PD, Bates T, Birchwood M. Designing youth mental health services for the 21st century: examples from Australia, Ireland and the UK. *Br J Psychiatry*. 2013;202(s54):s30–s35.
- Donald IN, Rickwood DJ, Carey TA. Understanding therapeutic change in young people—a pressing research agenda. *J Psychother Integr*. 2014;24(4):313.
- Rickwood DJ, Mazzer KR, Telford NR, Parker AG, Tanti CJ, McGorry PD. Changes in psychological distress and psychosocial functioning in young people visiting headspace centres for mental health problems. *Med J Aust*. 2015;202(10):537–542.
- Kwan B, Rickwood DJ. A systematic review of mental health outcome measures for young people aged 12 to 25 years. *BMC Psychiatry*. 2015;15(1):279.
- Brann P, Alexander M, Coombs T. Routine outcome measurement in youth mental health: a comparison of two clinician rated measures, HoNOSCA and HoNOS. *Psychiatry Res*. 2012;200(2):884–889.
- Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychol Med*. 2002;32(6):959–976.

40. Slade T, Grove R, Burgess P, Kessler psychological distress scale: normative data from the 2007 Australian National Survey of Mental Health and Wellbeing. *Aust N Z J Psychiatry*. 2011;45(4):308–316.
41. Tennant R, Hiller L, Fishwick R, et al. The Warwick–Edinburgh mental well-being scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes*. 2007;5(1):63.
42. Clarke A, Friede T, Putz R, et al. Warwick–Edinburgh Mental Well-being Scale (WEMWBS): validated for teenage school students in England and Scotland. A mixed methods assessment. *BMC Public Health*. 2011;11(1):487.
43. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. *Am J Psychiatry*. 1992;149(9):1148–1156.
44. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equation Model*. 1999;6(1):1–55.
45. van de Schoot R, Lugtig P, Hox J. A checklist for testing measurement invariance. *Eur J Dev Psychol*. 2012;9(4):486–492.
46. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct Equation Model*. 2002;9(2):233–255.
47. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale: Lawrence Earlbaum Associates; 1988:2.
48. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991;59(1):12.
49. Singer JD, Willett JB. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford: Oxford University Press; 2003.
50. Warren JS, Nelson PL, Mondragon SA, Baldwin SA, Burlingame GM. Youth psychotherapy change trajectories and outcomes in usual care: community mental health versus managed care settings. *J Consult Clin Psychol*. 2010;78(2):144.
51. Cuijpers P, Li J, Hofmann SG, Andersson G. Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clin Psychol Rev*. 2010;30(6):768–778.
52. Forkmann T, Scherer A, Boecker M, Pawelzik M, Jostes R, Gauggel S. The clinical global impression scale and the influence of patient or staff perspective on outcome. *BMC Psychiatry*. 2011;11(1):83.
53. Falkenström F. Does psychotherapy for young adults in routine practice show similar results as therapy in randomized clinical trials? *Psychother Res*. 2010;20(2):181–192.
54. Wolpert M, Görzig A, Deighton J, Fugard AJB, Newman R, Ford T. Comparison of indices of clinically meaningful change in child and adolescent mental health services: difference scores, reliable change, crossing clinical thresholds and ‘added value’ — an exploration using parent rated scores on the SDQ. *Child Adolesc Mental Health*. 2015;20(2):94–101.
55. Burgess P, Pirkis J, Coombs T. Modelling candidate effectiveness indicators for mental health services. *Aust N Z J Psychiatry*. 2009;43(6):531–538.
56. Baldwin SA, Berkeljon A, Atkins DC, Olsen JA, Nielsen SL. Rates of change in naturalistic psychotherapy: contrasting dose–effect and good-enough level models of change. *J Consult Clin Psychol*. 2009;77(2):203.
57. Gaynor ST, Weersing VR, Kolko DJ, Birmaher B, Heo J, Brent DA. The prevalence and impact of large sudden improvements during adolescent therapy for depression: a comparison across cognitive-behavioral, family, and supportive therapy. *J Consult Clin Psychol*. 2003;71(2):386.
58. Tang TZ, Luborsky L, Andrusyna T. Sudden gains in recovering from depression: are they also found in psychotherapies other than cognitive-behavioral therapy? *J Consult Clin Psychol*. 2002;70(2):444.
59. Whipple JL, Lambert MJ, Vermeersch DA, Smart DW, Nielsen SL, Hawkins EJ. Improving the effects of psychotherapy: the use of early identification of treatment and problem-solving strategies in routine practice. *J Couns Psychol*. 2003;50(1):59.

Psychology Research and Behavior Management

Publish your work in this journal

Psychology Research and Behavior Management is an international, peer-reviewed, open access journal focusing on the science of psychology and its application in behavior management to develop improved outcomes in the clinical, educational, sports and business arenas. Specific topics covered in the journal include: Neuroscience, memory and decision making; Behavior

modification and management; Clinical applications; Business and sports performance management; Social and developmental studies; Animal studies. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/psychology-research-and-behavior-management-journal>

Dovepress