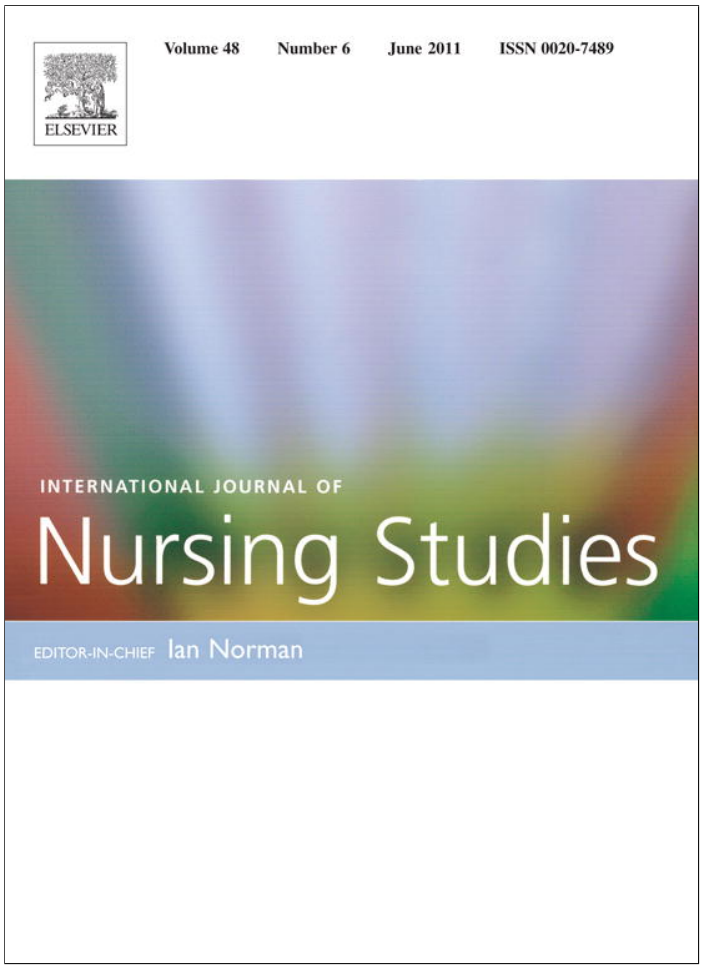

International Journal of Nursing Studies 48 (2011) 672–680

Contents lists available at ScienceDirect

International Journal of Nursing Studies

journal homepage: www.elsevier.com/ijns

The Chinese version of the Outcome Expectations for Exercise scale: Validation study

Ling-Ling Lee [a,*], Yu-Yun Chiu [b], Chin-Chih Ho [c], Shu-Chen Wu [a], Roger Watson [d,e]

[a] Department of Nursing, Tzu-Chi College of Technology, and School of Nursing, Tzu-Chi University, Hualien, Taiwan
[b] Department of Nursing, St. Mary's Medicine, Nursing and Management College, Ilan, Taiwan
[c] Department of Health Care Administration, Central Taiwan University of Science and Technology, Taichung, Taiwan
[d] School of Nursing and Midwifery, University of Sheffield, Sheffield, UK
[e] School of Nursing & Midwifery, University of Western Sydney, Australia

ARTICLE INFO

Article history:
Received 12 May 2010
Received in revised form 28 October 2010
Accepted 4 November 2010

Keywords:
Mokken scaling procedure
OEE-C scale
Older people
Physical activity
Reliability
Self-efficacy
Social cognitive theory
Validity

ABSTRACT

Background: Estimates of the reliability and validity of the English nine-item Outcome Expectations for Exercise (OEE) scale have been tested and found to be valid for use in various settings, particularly among older people, with good internal consistency and validity. Data on the use of the OEE scale among older Chinese people living in the community and how cultural differences might affect the administration of the OEE scale are limited.
Aim: To test the validity and reliability of the Chinese version of the Outcome Expectations for Exercise scale among older people.
Methods: A cross-sectional validation study was designed to test the Chinese version of the OEE scale (OEE-C). Reliability was examined by testing both the internal consistency for the overall scale and the squared multiple correlation coefficient for the single item measure. The validity of the scale was tested on the basis of both a traditional psychometric test and a confirmatory factor analysis using structural equation modelling. The Mokken Scaling Procedure (MSP) was used to investigate if there were any hierarchical, cumulative sets of items in the measure.
Results: The OEE-C scale was tested in a group of older people in Taiwan ($n = 108$, mean age = 77.1). There was acceptable internal consistency (alpha = .85) and model fit in the scale. Evidence of the validity of the measure was demonstrated by the tests for criterion-related validity and construct validity. There was a statistically significant correlation between exercise outcome expectations and exercise self-efficacy ($r = .34$, $p < .01$). An analysis of the Mokken Scaling Procedure found that nine items of the scale were all retained in the analysis and the resulting scale was reliable and statistically significant ($p = .0008$).
Conclusion: The results obtained in the present study provided acceptable levels of reliability and validity evidence for the Chinese Outcome Expectations for Exercise scale when used with older people in Taiwan. Future testing of the OEE-C scale needs to be carried out to see whether these results are generalisable to older Chinese people living in urban areas.

What is already known about the topic?

- The concept of outcome expectations, together with self-efficacy, plays a role in the adoption and maintenance of health behaviours.
- The original English version of the Outcome Expectations for Exercise scale tested among older people has been shown to have good internal consistency and validity.
- Data obtained from the Outcome Expectations for Exercise scale could provide valuable information for the design of physical activity interventions for older people.

* Corresponding author at: Tzu-Chi College of Technology, Department of Nursing, Room 429, No. 880 Chien-Kuo Road, Section 2, Hualien 970, Taiwan.
E-mail addresses: lllee@tccn.edu.tw, lllee666@gmail.com (L.-L. Lee).

0020-7489/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijnurstu.2010.11.001

What this paper adds

- The Chinese version of the Outcome Expectations for Exercise (OEE-C) scale is likely to be internally consistent across its nine items although individual items explain relatively little of the overall variance.
- This study provided evidence for criterion-related validity and construct validity of the OEE-C scale and results of the overall model fit of the scale are acceptable.
- The results of Mokken scaling provide a reliable hierarchy of items thereby providing greater insight into the latent trait, the outcome expectations of exercise, being measured.
- Cultural influences may affect the results of estimation of exercise outcome expectations measured by the OEE-C scale.

1. Background

Self-efficacy is a central concept of Bandura's social learning theory (Bandura, 1977) and is widely applied in predicting health behaviour, including exercise. The underlying assumption of social cognitive theory figures out that behavioural change and the maintenance of that behaviour are a function of the expectations about one's ability to perform a certain behaviour (self-efficacy) and the expectations about the outcome resulting from performing that behaviour (outcome expectations) (Bandura, 1986; Gecas, 1989; Schunk and Carbonari, 1984). Outcome expectations, together with self-efficacy, play a role in the adoption of health behaviours, the modification of unhealthy habits, and the maintenance of change (Bandura, 1991). It is therefore suggests that the stronger the individual's outcome expectations, the more likely that one will initiate and persist with a given activity.

One's outcome expectations for certain behaviour may be inconsistent with self-efficacy of that behaviour occasionally. People may agree that there are health benefits to adopting exercise in general (outcome expectations), whereas they may judge themselves incapable of including regular exercise in their daily life (self-efficacy) because of, for example, limitation of physical functioning or lack of time. Bandura (1997) suggested that people's level of motivation, affective states, and actions are based on beliefs rather than on objective assessments. Belief in behaviour's positive consequence may be at times more important than whether the behaviour has really caused a positive consequence in the past (Bandura, 1997). For this reason, how people behave can often be better predicted through the beliefs they hold than by what the result of an action can actually cause.

Although both self-efficacy and outcome expectations play a role in motivating health behaviours (Bandura, 1989; Maddux and Lewis, 1995), outcome expectancies are particularly important in forming an intention to adopt a new behaviour (Bandura, 1997; Schwarzer, 1992). The construct of outcome expectations was found to be an influential factor in evaluating the effectiveness of smoking cessation (Hertel et al., 2008) and weight loss programmes (Finch et al., 2005). For example, Hertel et al.'s study (2008) explored that individuals hold favourable outcome expectations about cessation were more likely to initiate cessation. Regarding to older adult population, empirical studies have shown that outcome expectations have an important influence on older people's exercise behaviour (Conn, 1998; Jette et al., 1998; Resnick and Spellbring, 2000; Schneider, 1997). A significant relationship between exercise outcome expectations and exercise behaviour has been found in different cross-sectional studies (Conn, 1998; Schneider, 1997). Several studies further reported that outcome expectations were better predictors of exercise behaviour than self-efficacy (Jette et al., 1998; Resnick, 1998).

The original English version of the Outcome Expectations for Exercise (OEE) scale was developed specifically to measure exercise outcome expectations among older people (Resnick et al., 2001; Resnick et al., 2000). The scale was designed for older people in the sense of being written using the older people's own words to describe the benefits they derived from exercise, and only included those items that were identified by older adults themselves as being of benefit. Among the nine items of the OEE scale, five of the items relate to physical benefits and four focus on mental health benefits. Validation of the English nine-item OEE scale have been tested and shown to be valid for use among older American people in the United States, with good internal consistency, validity and clinical feasibility (Resnick, 2005; Resnick et al., 2000, 2001).

Whether or not these estimates are valid for older people from different social and ethnic backgrounds requires further examination. Little data exist on the use of the OEE scale among older Chinese people living in the community and how cultural differences might affect the administration of the OEE scale has not been investigated. Before the OEE scale can be applied more broadly, a cross-cultural evaluation is warranted. The present study aims to assess the reliability and validity of the Chinese version of the Outcome Expectations for Exercise scale when used with older Chinese adults in Taiwan.

2. Methods

The study took place in a rural community of east Taiwan between April and May 2009. The criteria for inclusion in the study were: being resident in a local community, aged 65 years and over. The participants were recruited from a list of older people ($n = 200$) who resided in the community and met the inclusion criteria. The exclusion criteria were: an inability to carry out physical activity independently or the necessity for assistance with physical activity, such as the use of a walker. As suggested by Schwab (1980), a sample of 10 people to each item was required for validating an instrument (Schwab, 1980) and therefore a minimum of 90 participants was needed in the present study.

The original English version of the nine-item Outcome Expectations for Exercise (OEE) scale specifically focusing on the perceived consequences of exercise among older people (Resnick et al., 2001) was translated into Chinese, with the permission of the authors, for use in the present study. A process of translation and adaptation of instru-

ments that was developed by World Health Organization (Unknown) was used to guide the translation and adaptation process of the OEE scale. Forward translation (English to Chinese) was carried out by a bilingual translator and verified via further group discussion following the completion of the translation. The focus of the translation was on the conceptual equivalence of the items (Brislin, 1986). Another translator who had experience of back translation was invited to back translate the scale from Chinese to English and the back-translated instrument was then compared with the original English version to make sure the translation conceptually and linguistically appropriate equivalent to the original English version.

The translated instrument contained nine items measuring the participants' perceived benefit to exercise, such as "exercise makes me feel better physically", "exercise makes my mood better in general". One item of the translated OEE scale were raised as being of concern with regard to the Chinese words employed to describe "enjoy" (item 5 exercise "is an activity I enjoy doing"). A modified Chinese version was developed after a series of meetings with members of the study team. Two professionals in the fields of gerontology and public health in Taiwan were then consulted for the wording and phrasing of each item of the initial Chinese version of the scale. Twenty-two older people living in another community were recruited to pilot test the translated measure for readability and feasibility of completion. Due to the short and simple sentence design of the items in the scale, literate older people could read and answer the items, while for the illiterate older people, the items could also be understood by reading the item for them. However, further explanation was required for both groups of older people for the phrase of "a sense of personal accomplishment" in item 6.

The data of the OEE-C scale were collected by a set of questionnaires and administered by a nurse researcher to a group of older people via face-to-face interviews in their own homes. Following each statement about the outcomes or benefits of exercise, the participants were asked to choose one of five responses ranging from 'strongly agree' to 'strongly disagree'. The scale was scored by summing the numerical ratings for each response and dividing the total by the number of responses. Higher scores indicate lower exercise outcome expectations, with a minimum score of 1 and a maximum score of 5. Data were then reversed as that higher score represents strong exercise outcome expectation for this validation test.

The validity of the OEE-C scale was also tested by using several other measures. These measures included demographic information (i.e. age and gender), exercise self-efficacy, physical activity status, and perceived health. The demographic information included age, gender, years of education, and marital status. Physical activity was measured by the concept of frequency, duration and intensity. Of those, frequency of physical activity was measured by asking how often (times) the participant exercised weekly in a recent month and duration was measured by asking the time (min) involving in exercise each time. The calculation of the mean hours of physical activity was to multiply frequency of adopting exercise per week, which was indicated as times, by the duration, which was indicated as hour(s) each time. Borg's ratings of perceived exertion scale (RPE) (Borg, 1982) was used to monitor participants' intensities of physical activity. The higher the score of RPE, which ranged from one to 15, represents the higher intensity of physical activity and mild to moderate intensity of physical activity was defined between six and eight.

Exercise self-efficacy was measured using the Chinese Self-Efficacy for Exercise (SEE-C) scale (Lee et al., 2009). It was selected for use in the present study because of its prior use among older people (Lee et al., 2007a; Resnick and Jenkins, 2000; Resnick and Nigg, 2003). This scale focuses on self-efficacy expectations relating to the confidence to exercise in the face of encountering barriers (such as bad weather and tiredness). The reliability and validity of this scale and the Chinese version were tested in previous studies (Lee et al., 2009; Resnick, 2000; Resnick and Jenkins, 2000). The participants were asked to choose the response that most reflected their situation. As the original English version, the SEE-C consists of nine in total. The responses for each item were scored from 0 to 10 and represented their degree of confidence about undertaking exercise regularly on various occasions. The scale was scored by summing the numerical ratings for each response and dividing the total by the number of non-missing responses. The mean scores for the self-efficacy of exercise ranged from 0 to 10 with higher scores representing greater exercise self-efficacy. The internal consistency of the SEE-C scale, as measured by Cronbach's alpha was .90 in this study ($n = 108$).

Perceived health was measured by three items used in others' work (Conn, 1998; Hawkes and Holm, 1993) that are summed to create a score of between 3 and 10, with higher scores indicating better perceived health. The participants rated their health as 'very good', 'good', 'poor', or 'very poor'. They then compared their health to others similar in age to themselves and rated it as 'better than average', 'average', or 'worse than average. The participants further rated the extent to which their health interfered with their desired activities as 'not at all', 'somewhat', or 'a great deal'. The criterion validity of this measure of perceived health has been supported by findings that suggest positive correlations with physician visits ($r = .25$) (Kviz and Flaskerud, 1984), number of reported symptoms ($r = .38$) (Kviz and Flaskerud, 1984), number of chronic illnesses ($r = -.46$) (Conn, 1998) and number of prescriptive medications ($r = -.40$) (Conn, 1998). The Cronbach's alpha of the perceived health scale was .63 in the present study ($n = 108$).

3. Statistical analysis

Descriptive analyses of the data were performed with all of the study variables to describe the study participants. Estimates of internal consistency were carried out using Cronbach's alpha and .7 or greater was considered as evidence of the internal consistency of the OEE-C scale (Nunnally and Bernstein, 1994). For a single item measure, an alternative estimate of reliability using a structural equation modelling, a squared multiple correlation coeffi-

cient (R^2), was used (Bollen, 1989), as R^2 estimates the systematic variance in the observed score that can be explained by each item in the measurement model (Bollen, 1989; Jagodzinski and Kuhnel, 1987).

Validity testing for the OEE-C scale was based on both traditional psychometric testing that was analysed with the use of SPSS version 14.0 and a confirmatory factor analysis using structural equation modelling that was analysed with the use of EQS version 6.1. For traditional psychometric testing, criterion-related validity was examined by a multiple regression analysis. Four variables (age, gender, perceived health, and exercise self-efficacy) were included as predictors of physical activity status. Construct validity was tested by the use of know-group technique, which involves differentiating members of one group from another on the basis of their scale scores (Devellis, 2003). This was examined by testing the hypothesis that those who were regular exercisers would have higher exercise outcome expectations than those who exercised less regularly. The construct validity of the OEE-C scale was also examined by testing the hypothesis that there would be a statistically significant relationship between exercise self-efficacy and exercise outcome expectations.

A confirmatory factor analysis was carried out using structural equation modelling. The chi-square statistic, the Normed Fit Index (NFI), the Comparative Fit Index (CFI) and the Steigers' root mean square error of approximation (RMSEA) were used to estimate the model fit. The larger the probability associated with the chi-square, the better the fit of the model to the data (Loehlin, 2004). Because the chi-square statistic is influenced by the sample size, the chi-square divided by degrees of freedom was used to evaluate the model fit.

In addition to the above analysis, the Mokken scaling procedure (MSP) was used to investigate if there were any hierarchical, cumulative sets of items in the questionnaire with the use of MSP version 5 for Windows. Mokken scaling falls within the set of analytical techniques called item response theory (IRT), which examines the relationship between the items in a scale and the latent trait being measured through the ordering of items along the latent trait (van Schuur, 2003). Principally, Loevinger's coefficient (H) is used to judge if a hierarchy of items exists, i.e. when an H of $>.40$ is considered good (Molenaar and Sijtsma, 2000). The reliability of the scale (Rho) can be checked using a test–retest procedure that is analogous to Cronbach's alpha and Rho >0.7 is taken to indicate a reliable scale.

4. Results

The study was approved by the Hualien Education University Ethics Committee and signed informed consent was obtained prior to data collection. Validity and reliability test of the Chinese version of OEE scale was examined in a group of older people living in community ($n = 108$). The mean age of this group of participants was 77.1 ($SD = 5.8$) years. The majority of the participants was male (58.3%), low education level (lower than primary school level, 56.5%) and not living alone (94.1%).

Table 1
Demographic data and mean scores for the study measures ($n = 108$).

	Mean (SD)	Range	n (%)
Age (years)	77.1 (5.77)	66–91	
Male/Female			63 (58.3)/ 45 (41.7)
Education level			
≤6 years			87 (80.6)
>6 years			21 (19.5)
Marital status			
Married			102 (94.4)
Unmarried			6 (5.6)
Living alone			
No			102 (94.4)
Yes			6 (5.6)
The top three chronic diseases			
Hypertension			81 (75.0)
Diabetes mellitus			23 (21.3)
Heart disease			22 (20.4)
Average number of chronic disease per participants	2.3 (1.3)	0–6	
Exercise outcome expectations (measured by OEE-C scale)	4.2 (0.62)	1–5	
Exercise self-efficacy (measured by SEE-C scale)	6.9 (3.10)	0–10	
Perceived health	6.8 (1.55)	3–10	
Self-reported mean hours of mild to moderate physical activity/week	5.1 (5.04)	0–21	
Participating in regular physical activity			82 (75.9)

SD = standard deviation.

The mean scores of the study participants on the OEE-C and SEE-C scale, their perceived health, and levels of physical activity are presented in Table 1. The mean OEE-C score was 4.2 ($SD = 0.62$) on a scale of 1 to 5 and the mean perceived health score was 6.8 ($SD = 1.55$) on a scale of 3–10. More than three quarters (82/108, 75.9%) of the participants took part in regular physical activity. The participants were involved in 5.1 h ($SD = 5.04$) per week of mild to moderate physical activity. The Cronbach's alpha coefficient of the OEE-C scale was .85 ($n = 108$), which was similar to the original English version, which was .89 ($n = 175$) (Resnick et al., 2000). All of the item intercorrelations were positive and statistically significant, which ranged from .27 (item 8 with item 3 and item 9 with item 7) to .78 (item 6 with item 5). The squared multiple correlation coefficients (R^2) ranged from .24 (item 1) to .72 (item 5) (Table 2).

The evidence for the criterion-related validity of the OEE-C scale was supported by the fact that the exercise outcome expectations were significantly related to the mean exercise hours per week ($r = .33$, $p < .01$) (Table 3). A standardised Beta coefficient was .20 ($p = .033$) in a regression model after adjusting for age and gender ($F = 16.987$, $p < .0001$). Exercise self-efficacy was another variable in the model, which was statistically significantly related to exercise behaviour. Overall, the two variables

Table 2
Reliability estimates (R^2) for the OEE-C measure.

Question items	Reliability estimates (R^2)
Exercise...	
1. Makes me feel better in general	.28
2. Makes my mood better in general	.24
3. Helps me feel less tired	.27
4. Makes my muscles stronger	.40
5. Is an activity I enjoy doing	.72
6. Gives me a sense of personal accomplishment	.61
7. Makes me more alert mentally	.33
8. Improves my endurance in performing my daily activities	.35
9. Helps to strengthen my bones	.37

account for 23% of the variance in weekly mean exercise hours.

Evidence of construct validity of the OEE-C scale was supported by a statistically significant difference between the OEE-C scores of those exercising regularly defined as that a total exercise time per week is more than 60 min with at least 10 min per occasion (American College of Sports Medicine, 2007) and those who did not ($t = -2.86$, $p = .005$). Those who exercised regularly had a mean OEE-C score of 4.3 ($SD = 0.59$), whereas those who did not had a mean score of 3.9 ($SD = 0.64$). There was also evidence for the construct validity of OEE-C scale on the basis of a statistically significant correlation between exercise outcome expectations and exercise self-efficacy ($r = .34$, $p < .01$) (Table 3). The moderate association between exercise outcome expectations and exercise self-efficacy further supports the fact that they are distinct constructs.

There was acceptable support for the confirmatory factor analysis of the Chinese version of the Outcome Expectations for Exercise (OEE-C) scale. The Chi-square divided by degrees of freedom (χ^2/df ratio = 2.26) suggested a fair model fit (Arbuckle, 1997) nevertheless, the probability associated with the χ^2 is small ($\chi^2 = 60.92$, $df = 27$, $p < .0001$). The NFI (Normed Fit Index) was .98 and the CFI (Comparative Fit Index) was .99. However, RMSEA (Root Mean Square Error of Approximation) was on a borderline, at .108. Each item was statistically significantly related to the outcome expectations with path coefficients ranging from .49 to .85. The nine items explained in total 40% of the variance in exercise outcome expectations.

The outcome of the Mokken Scaling Procedure analysis is shown in Table 4. Nine items were retained in the analysis and the resulting scale was moderately strong ($H = .43$), reliable (Rho = 0.85) and statistically significant ($p = .0008$). The most readily adopted item, as indicated by the mean score is that exercise improves mood and the least readily adopted is that exercise improves mental alertness. In Mokken scaling terms, therefore, the least difficult item pertains to mood and the most difficult pertains to mental alertness. The arrangement of the items in the scale, in terms of difficulty, runs from those concerned with psychological outcomes: mood, feeling better physically and enjoyable activity. The most difficult

Table 3
Bivariate correlations of the major study variables ($n = 108$).

	Self-reported mean hours of mild to moderate physical activity/week	Exercise outcome expectations (measured by OEE-C)	Exercise self-efficacy (measured by SEE-C)	Perceived health
Self-reported mean hours of mild to moderate physical activity/week	1			
Exercise outcome expectations (measured by OEE-C)	.33**	1		
Exercise self-efficacy (measured by SEE-C)	.46**	.34**	1	
Perceived health		.28**	.22*	1

Note: OEE-C, Chinese version of Outcome Expectations for Exercise scale; SEE-C, Chinese version of Self-Efficacy for Exercise scale.
* $p < .05$ (Pearson's correlation coefficient).
** $p < .01$ (Pearson's correlation coefficient).

Table 4
Mokken scaling analysis of Outcome Expectations of Exercise ($n = 108$).

Item number	Mean score	Item H	Item label
7	3.74	.45	Makes me more alert mentally
9	4.07	.40	Helps to strengthen my bones
4	4.17	.45	Makes my muscles stronger
3	4.22	.37	helps make me feel less tired
8	4.22	.42	Improves my endurance in performing my daily activities
6	4.26	.48	Gives me a sense of personal accomplishment
5	4.30	.53	Is an activity I enjoy doing
1	4.44	.39	Makes me feel better physically
2	4.50	.41	Makes my mood better in general
Total	4.21 ($SD = 5.57$)	.43	Rho = 0.85; $p = .0008$

items are concerned with tangible physical and mental outcomes, such as strength and mental alertness.

5. Discussion

The Chinese Outcome Expectations for Exercise (OEE-C) scale was found to be effective in examining expected outcomes of involving exercise among older people. There was evidence of internal consistency, tested by Cronbach's alpha. However, evidence of the reliability, based on squared multiple correlation coefficients (R^2), was relatively limited. There were three items having an R^2 lower than .3, indicating that less than 30% of the variance in each of the three items was explained by the model. It is possible that the benefits of feeling better physically (item 1), improving one's mood (item 2), and feeling less tired (item 3) were less desirable among this group of participants. This may have resulted from cultural influence. Older Chinese people's intention to exercise is mainly for the sake of their health to prevent them becoming a burden on their children (Lee et al., 2007b). Therefore, additional relevant items, such as "exercise makes me feel more independent", might improve the model fit and explain more of the variance in exercise outcome expectations among older Chinese people. Further exploration of the beneficial effects of exercise as particularly recognised by the older adult population is warranted to ensure that a full construct of exercise outcome expectations is included in this measure. Item 5 (exercise is an activity I enjoy doing) obtained the highest level of R^2 (.72), indicating that the enjoyment of exercise may perform an essential role of being an expected benefit of exercise among older people. The R^2 of item 4 (exercise makes my muscles stronger, R^2 = .40) was low and it might be influenced by the media due to the fact that the fit image of the advertisement came mainly from young adults. Older people may therefore assume that the benefit of exercise in terms of strengthening muscles is aimed at younger people or that limitations in physical functioning prevent them from becoming a member of the exercise society (Lee et al., 2008), which may meanwhile lessen older people's confidence about doing exercise. A previous OEE validation study found that the emphasis in the lay literature on the bone building benefits of exercise makes the R^2 of item 9 (exercise helps to strengthen my bones) increase from .33 to .53 (Resnick et al., 2001). To some extent, these support the use of a media campaign and lay literature to influence and enhance the exercise outcome expectations but it needs particularly to make the target older people to strengthen both their exercise outcome expectations and exercise self-efficacy. Therefore, those items are still important in the measure of exercise outcome expectations among older people in terms of informing the design of physical activity intervention programme.

The evidence of the criterion-related validity of the measure was supported by a positive correlation between exercise outcome expectations and weekly exercise hours. There was also evidence for the construct validity of the measure based on the hypothesis that there was a statistically significant correlation between exercise outcome expectations and exercise self-efficacy and on the basis of the known-group technique as those exercising regularly had a significantly higher OEE-C score than those who did not. Recently, McAuley and colleagues' validation study also found that higher self-efficacy for exercise was significantly associated with social aspects of Outcome Expectations for Exercise among multiple sclerosis patients (McAuley et al., 2010). Nevertheless, these correlations were weak, although they were statistically significant in the present study. This may be related to the fact that many factors influence exercise and other forms of physical activity behaviour, particularly among older people. Confirmatory factor analysis indicated that the fit of the full model to the data was acceptable but relatively lower than that of the original English version. Further testing of this scale with older Chinese people living in urban areas is needed to evaluate the reliability of the OEE-C scale and explore the reasons for the low R^2 values.

The differences in the participants' characteristics between the present study and the previous investigations of the OEE scale (Resnick et al., 2001) include race (Chinese vs. White American), education (fairly well educated vs. well educated), gender (more male vs. more female), the mean age of the participants (77.1 ± 5.8 vs. 85 ± 6.1 years), and marital status (mainly married vs. mainly unmarried). The participants in the original validation study (Resnick et al., 2000), who were living in a continuing care retirement community, were involved in more hours per week of physical activity than the participants in the present study (12.5 ± 8.6 vs. 5.1 ± 5.0), while the scores for the exercise outcome expectations were higher in the present study than in that by Resnick et al.'s other validation study (2001), in which the participants lived independently in their own apartments (4.2 ± 0.62 vs. 3.4 ± 0.82). Whether the differences in the exercise outcome expectations scores between the two groups were from various living environments that might facilitate different ranges of physical activity resources or reside in an enjoyable rural environment requires further investigation.

The results of the Mokken scaling has added value to the present analysis and also to the psychometric analysis of the OEE-C scale as it both complements the classical test theory analysis and provides a reliable hierarchy of items thereby providing greater insight into the latent trait, outcome expectations of exercise, being measured. The items in the Mokken scale are grouped similarly to the outcome of the factor analysis from previous studies with psychological outcomes and the physical outcomes of exercise (Resnick et al., 2000, 2001). However, the results from the Mokken scale procedure in the present study shows how the psychological and the physical outcomes are related in the sense that people more readily perceive psychological outcomes than physical ones. The practical application of this finding is that people who score highly on the physical as well as psychological outcome expectations of exercise may benefit more than people who simply score highly on psychological outcomes, so future study to test the effect of this application is warranted.

The present validation study of the OEE-C scale was carried out among a group of older people who were living

in a rural area; the findings may, therefore, be context specific, which makes it difficult to generalise them to other settings, such as urban areas. Further investigation is needed to test if the findings are consistent over time, between countries, and in different community settings in regard to older people's perception of Outcome Expectations for Exercise. Another limitation of the present study is that the data about the physical activity were collected by the self-report method, which may cause reporting bias, relating to the accuracy of the older people's response to the amount of physical activity. The reliance on self-reported walking and physical activity may limit the validity of the data, previous studies have found that participants over 65 may tend to overestimate their physical activity amount (Sallis and Saelens, 2000; Sims et al., 1999). More objective measures of physical activity, such as including a pedometer, are needed to establish the relationship between exercise outcome expectations and exercise behaviour in a more accurate way.

6. Conclusion

To our knowledge, this is the first study that systematically examines the reliability and validity of the Chinese Outcome Expectations for Exercise (OEE-C) scale among older Chinese people. Acceptable levels of reliability and validity for the scale were found when tested in older people in Taiwan. This indicates that the OEE-C scale can be employed to measure the perceived exercise benefits among older Chinese people living in rural areas. Although the instrument demonstrated evidence of reliability and validity, further improvements are warranted. The applicability of the scale will be explored through its future use in nursing practice (such as in the community), in research studies, and with different age groups (such as younger adults).

Acknowledgments

The authors would like to express their sincere gratitude to all of the participants in this study, for the statistical consultation of Dr. Yi-Hwai Lee, and for the translation consultation of Dr. Hsiang-Ming Hsu and Dr. Shu-Chuan Lau. Appreciation is also expressed to Dr. Barbara Resnick who kindly offered her guidance throughout the course of this research project. We are grateful to the reviewers for their thorough and constructive comments on an earlier draft of this paper.

Conflict of interest

There are no potential conflicts of interest known to any of the contributing authors.

Appendix A. Chinese Outcome Expectations for Exercise scale

Chinese Outcome Expectations for Exercise scale

【中文版運動結果預期量表】

此部份將瞭解您對運動結果的看法，我將唸出以下的句子，請告訴我哪一個答案最能反應您對運動結果的看法。

	非常同意	同意	沒意見(不一定)	不同意	非常不同意
1. 運動讓我覺得身體比較好	1☐	2☐	3☐	4☐	5☐
2. 一般來說，運動會讓我覺得心情比較好	1☐	2☐	3☐	4☐	5☐
3. 運動讓我比較不累	1☐	2☐	3☐	4☐	5☐
4. 運動讓我覺得肌肉比較有力	1☐	2☐	3☐	4☐	5☐
5. 運動是一項我做起來有樂趣的活動	1☐	2☐	3☐	4☐	5☐
6. 運動讓我有成就感	1☐	2☐	3☐	4☐	5☐
7. 運動使我心智上的警覺度比較好	1☐	2☐	3☐	4☐	5☐
8. 運動增加我日常生活活動的耐力	1☐	2☐	3☐	4☐	5☐
9. 運動可以強化我的骨骼	1☐	2☐	3☐	4☐	5☐

References

American College of Sports Medicine, 2007. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. In: Medical Specialty Society American Heart Association, Professional Association, NGC, p. 13.

Arbuckle, J., 1997. Amos Users' Guide Version 3.6. Small Waters Corporation, Chicago.

Bandura, A., 1977. Self-efficacy: toward a unifying theory of behavioral change. Psychological Review 84 (2), 191–215.

Bandura, A., 1986. Cognitive regulators. In: Bandura, A. (Ed.), Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice-Hall, Englewood Cliffs, NJ, pp. 467–480.

Bandura, A., 1989. Human agency in social cognitive theory. American Psychologist 44 (9), 1175–1184.

Bandura, A., 1991. Self-efficacy mechanism in physiological activation and health-promoting behaviour. In: Madden, J. (Ed.), Neurobiology of Learning, Emotion and Affect. Raven Press, New York, pp. 229–270.

Bandura, A., 1997. Theoretical perspectives. In: Bandura, A. (Ed.), Self-efficacy: The Exercise of Control. W.H. Freeman, New York, pp. 1–35.

Bollen, K.A., 1989. Structural Equations with Latent Variables. Wiley, New York.

Borg, G.A., 1982. Psychophysical bases of perceived exertion. Medicine and Science in Sports and Exercise 14 (5), 377–381.

Brislin, R.W., 1986. The wording and translation of research instruments. In: Lonner, W.J., Berry, J.W. (Eds.), Field Methods in Cross-Cultural Research. Sage, Beverly Hills, pp. 137–164.

Conn, V.S., 1998. Older adults and exercise: path analysis of self-efficacy related constructs. Nursing Research 47 (3), 180–189.
Devellis, R.R., 2003. Scale Development: Theory and Applications Second Edition (Applied Social Research Methods). Sage, CA.
Finch, E.A., Linde, J.A., Jeffery, R.W., Rothman, A.J., King, C.M., Levy, R.L., 2005. The effects of outcome expectations and satisfaction on weight loss and maintenance: correlational and experimental analyses – a randomized trial. Health Psychology 24 (6), 608–616.
Gecas, V., 1989. The social psychology of self-efficacy. Annual Review of Sociology 15 (2), 291–316.
Hawkes, J.M., Holm, K., 1993. Gender differences in exercise determinants. Nursing Research 42 (3), 166–172.
Hertel, A.W., Finch, E.A., Kelly, K.M., King, C., Lando, H., Linde, J.A., Jeffery, R.W., Rothman, A.J., 2008. The impact of expectations and satisfaction on the initiation and maintenance of smoking cessation: an experimental test. Health Psychology 27 (3 Suppl.), S197–206.
Jagodzinski, W., Kuhnel, S.M., 1987. Estimation of reliability and stability in single-indicator multiple-wave models. Sociological Methods and Research 15 (3), 219–258.
Jette, A.M., Rooks, D., Lachman, M., Lin, T.H., Levenson, C., Heislein, D., Giorgetti, M.M., Harris, B.A., 1998. Home-based resistance training: predictors of participation and adherence. Gerontologist 38 (4), 412–421.
Kviz, F.J., Flaskerud, J.H., 1984. An evaluation of the index of medical underservice: results from a rural consumer survey. Medical Care 22 (10), 877–889.
Lee, L.L., Arthur, A., Avis, M., 2007a. Evaluating a community-based walking intervention for hypertensive older people in Taiwan: a randomized controlled trial. Preventive Medicine 44 (2), 160–166.
Lee, L.L., Arthur, A., Avis, M., 2008. Using self-efficacy theory to develop interventions that help older people overcome psychological barriers to physical activity: a discussion paper. International Journal of Nursing Studies 45 (3), 1690–1699.
Lee, L.L.s, Avis, M., Arthur, A., 2007b. The role of self-efficacy in older people's decisions to initiate and maintain regular walking as exercise - findings from a qualitative study. Preventive Medicine 45 (1), 62–65.
Lee, L.L., Perng, S.J., Ho, C.C., Hsu, H.M., Liu, S.C., 2009. A preliminary reliability and validity study of the Chinese version of the Self-Efficacy for Exercise scale for older adults in Taiwan. International journal of Nursing Studies 46 (1), 230–238.
Loehlin, J.C., 2004. Latent Variable Models: An Introduction to Factor, Path, and Structural Equation Analysis. Lawrence Erlbaum Associates, Mahwah, NJ.
Maddux, J.E., Lewis, J., 1995. Self-efficacy and Healthy Behavior; Prevention, Promotion and Detection. In: Maddux, J.E. (Ed.), Self-efficacy, Adaptation and Adjustment: Theory, Research, and Application. Plenum Press, New York, p. 395.
McAuley, E., Motl, R.W., White, S.M., Wojcicki, T.R., 2010. Validation of the multidimensional Outcome Expectations for Exercise scale in ambulatory, symptom-free persons with multiple sclerosis. Archives of Physical Medicine and Rehabilitation 91 (1), 100–105.
Molenaar, I.W., Sijtsma, K., 2000. User's Manual MSP 5 for Windows (Software manual). IEC ProGAMMA, Groningen, The Netherlands.
Nunnally, J., Bernstein, I., 1994. Psychometric Theory. McGraw-Hill, New York.
Resnick, B., 1998. Self-efficacy in geriatric rehabilitation. Journal of Gerontological Nursing 24 (7), 34–45.
Resnick, B., 2000. Functional performance and exercise of older adults in long-term care settings. Journal of Gerontological Nursing 26 (3), 7–16.
Resnick, B., 2005. Reliability and validity of the Outcome Expectations for Exercise Scale-2. Journal of Aging and Physical Activity 13 (4), 382–394.
Resnick, B., Jenkins, L.S., 2000. Testing the reliability and validity of the Self-Efficacy for Exercise scale. Nursing Research 49 (3), 154–159.
Resnick, B., Nigg, C., 2003. Testing a theoretical model of exercise behavior for older adults. Nursing Research 52 (2), 80–88.
Resnick, B., Spellbring, A.M., 2000. Understanding what motivates older adults to exercise. Journal of Gerontological Nursing 26 (3), 34–42.
Resnick, B., Zimmerman, S., Orwig, D., Furstenberg, A.L., Magaziner, J., 2001. Model testing for reliability and validity of the Outcome Expectations for Exercise scale. Nursing Research 50 (5), 293–299.
Resnick, B., Zimmerman, S.I., Orwig, D., Furstenberg, A., Magaziner, J., 2000. Outcome Expectations for Exercise scale: utility and psychometrics. Journals of Gerontology Series B – Psychological Sciences and Social Sciences 55B (6), S352–356.
Sallis, J.F., Saelens, B.E., 2000. Assessment of physical activity by self-report: status, limitations, and future directions. Research Quarterly for Exercise and Sport 71 (2 Suppl.), S1–14.
Schneider, J.K., 1997. Self-regulation and exercise behavior in older women. Journal of Gerontology Psychological Sciences 52B (5), 235–241.
Schunk, D.H., Carbonari, J.P., 1984. Self-efficacy models. In: Matarazzo, J.D., weiss, S.M., Herd, J.A., Miller, N.E., Weiss, S.M. (Eds.), Behavioural Health: A Handbook of Health Enhancement and Disease Prevention. Wiley, New York, pp. 230–247.
Schwab, D.P., 1980. Construct validity in organization behavior. In: Staw, B.M., Cummings, L.L. (Eds.), Research in Organization Behavior. JAI Press, Greenwich, CT.
Schwarzer, R., 1992. Self-Efficacy: Thought Control of Action. Hemisphere Publishing Corp., Washington, DC.
Sims, J., Smith, F., Duffy, A., Hilton, S., 1999. The vagaries of self-reports of physical activity: a problem revisited and addressed in a study of exercise promotion in the over 65 s in general practice. Family Practice 16 (2), 152–157.
van Schuur, W.H., 2003. Mokken Scale Analysis: Between the Guttman Scale and Parametric Item Response Theory, pp. 139–163.
World Health Organization, Unknown. Process of translation and adaptation of instruments. Retrieved June 16th, 2009, from http://www.who.int/substance_abuse/.