BMJ Open Sport & Exercise Medicine

Muscle-Strengthening Exercise Questionnaire (MSEQ): an assessment of concurrent validity and test–retest reliability

Jane Shakespear-Druery [1], Katrien De Cocker [1,2], Stuart J H Biddle [1], Jason Bennie [1]

To cite: Shakespear-Druery J, De Cocker K, Biddle SJH, *et al*. Muscle-Strengthening Exercise Questionnaire (MSEQ): an assessment of concurrent validity and test–retest reliability. *BMJ Open Sport & Exercise Medicine* 2022;**8**:e001225. doi:10.1136/bmjsem-2021-001225

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bmjsem-2021-001225).

Accepted 21 December 2021

[1]Physically Active Lifestyles Research Group (USQ-PALs), Centre for Health Research, University of Southern Queensland, Springfield Central, Queensland, Australia
[2]Department of Movement and Sports Sciences, Ghent University, Gent, Belgium

Correspondence to
Jane Shakespear-Druery; Jane.Shakespear-Druery@usq.edu.au

ABSTRACT

Objectives Muscle-strengthening exercise (MSE) has multiple independent health benefits and is a component of global physical activity guidelines. However, the assessment of MSE in health surveillance is often limited to the constructs of frequency (days/week), with little focus on constructs such as MSE type, muscle groups targeted and intensity. This study describes the test–retest reliability and concurrent validity of the Muscle-Strengthening Exercise Questionnaire (MSEQ), which was developed to assess multiple MSE participation constructs.

Methods The MSEQ was developed to assess the weekly frequency, session duration and intensity, types of MSE (eg, weight machines, bodyweight exercise) and muscle groups targeted. Two convenience samples of adult participants were recruited. Test–retest reliability was completed online by 85 participants. Concurrent validity was assessed for 54 participants using an online 7-day MSE log.

Results The MSEQ shows high test–retest reliability for frequency, duration and level of intensity for each of the four MSE types (using weight machines, bodyweight exercises, resistance exercises and holistic exercises), and for the four types combined (ρ range 0.76–0.91). For muscle groups targeted, the reliability ranged mostly from moderate-to-substantial for each of the four MSE types (κ range 0.44–0.78) and fair-to-moderate for the four types combined (κ range 0.35–0.51). Concurrent validity for frequency, duration and level of intensity for each of the four MSE types, and the four types combined, was moderate-to-high (ρ range 0.30–0.77).

Conclusion The MSEQ shows acceptable reliability and validity for four key MSE constructs. This new MSEQ survey instrument could be used to assess adults' MSE.

Key messages

What is already known

- Muscle-strengthening exercise has multiple independent health benefits.
- Assessment of muscle-strengthening exercise is rare in physical activity surveillance.
- Surveillance instruments assessing muscle-strengthening exercise are mostly limited to frequency (days/week) and duration (minutes/session).

What are the new findings

- We developed a new online muscle-strengthening exercise assessment instrument, the Muscle-Strengthening Exercise Questionnaire (MSEQ).
- The MSEQ has shown acceptable 7-day test–retest reliability.
- The MSEQ has shown adequate validity when using a 7-day muscle-strengthening exercise log as the standard.
- Future population-level health surveillance of muscle-strengthening exercise may include the MSEQ.

INTRODUCTION

Strong scientific evidence demonstrates that regular muscle-strengthening exercise (MSE: including using weight training equipment and machines, resistance bands and doing bodyweight exercises) is linked to optimal health and well-being in adults.[1 2] In brief, meta-analyses and systematic reviews of controlled clinical exercise studies show that MSE leads to enhanced cardiometabolic,[3] musculoskeletal[4] and mental health,[5] and reductions in visceral fat.[6] Recent data from prospective cohort studies suggest that MSE is independently associated with a reduced risk of all-cause and disease-specific mortality,[7 8] cardiovascular disease,[9] type 2 diabetes,[10 11] obesity[12 13] and some cancers.[8]

Despite its multiple independent health benefits, and the fact that MSE was first included as part of the US physical activity guidelines in 2008,[14] and global guidelines since 2010,[15] in comparison to moderate-to-vigorous aerobic physical activity (MVPA: eg, walking, running or cycling) and sedentary behaviour (low energy sitting, reclining or lying posture in waking hours), this exercise modality has received little attention in physical activity epidemiology.[16 17] In particular, research has shown that the assessment of MSE is rare in physical activity surveillance.[18]

Moreover, our recent systematic review of the assessment of MSE within health surveillance highlighted two key limitations in the current assessment of MSE at the population level.[19] First, few surveillance instruments assess MSE participation constructs beyond weekly frequency. Second, unlike aerobic MVPA and sedentary behaviour,[20–22] there is no standardised instrument for assessing MSE in health surveillance.[19]

Developing an understanding of the surveillance of multiple MSE participation constructs is important because clinical exercise studies demonstrate that factors such as type (single vs multijoint; bodyweight vs use of weight machines, etc), duration and intensity of MSE, affect some key outcomes such as skeletal muscle size/endurance/strength.[23 24] The assessment of MSE participation constructs at the population level, that goes beyond simply frequency, is critical for establishing the optimal dose of this exercise modality for health in future studies.[17] Furthermore, a standardised MSE assessment instrument will be essential for the population-level tracking and monitoring of this important and currently understudied health behaviour. In addition, accurate and consistent assessments of physical activity-related behaviours are key for identifying at-risk population subgroups most in need of future large-scale public health interventions.[25]

This study aimed to describe the development of the Muscle-Strengthening Exercise Questionnaire (MSEQ)—a newly designed MSE assessment instrument for adults, with a specific focus on the assessment of its test–retest reliability and concurrent validity.

METHODS

Study population

From January to March 2021, a subsample of participants was recruited from a larger online study on MSE participation, barriers/facilitators and attitudes towards this exercise mode ('Main Study'). In this main study, we recruited a convenience sample of 461 adults (aged ≥18 years) via the use of social media (eg, Facebook, Twitter, Instagram) and professional networks. At the end of the Main Study, respondents were invited to participate in further follow-up research for assessing the reliability and validity of survey items concerning their MSE participation during a usual week (MSEQ, described below). If they agreed, they were allocated, on an alternating (one for one) basis, to either: (1) reliability sample or (2) validity sample. Informed consent was obtained from all participants.

Measures

Muscle-Strengthening Exercise Questionnaire (MSEQ)

The MSEQ was designed to be a brief assessment instrument for delivery in an online format. Specifically, we created a 9-item instrument that assesses key MSE constructs for use in future physical activity surveillance. The final version of the MSEQ is shown in online supplemental digital content 1. The initial development of the MSEQ was broadly guided by several key MSE resources, including the 2009 'American College of Sports Medicine Position Stand on Progression Models in Resistance Training for Healthy Adults',[26] Garber *et al*'s 'Guidance for Prescribing Exercise'[2] and the 2018 Physical Activity Guidelines for Americans.[27] After a review of these sources and consideration of what is practical to include in a surveillance instrument, five MSE participation constructs were assessed: (1) type; (2) frequency; (3) duration; (4) intensity and (5) the muscle groups targeted. The preamble of the MSEQ, the key justifications for choosing each MSE construct and their response items are now described.

Preamble

When developing self-report instruments of physical activity-related behaviours, it is important to provide respondents with some examples of the behaviours of interest to enhance comprehension.[28] Given that at the population level, ~60% of adults do no MSE,[29] an understanding of what constitutes MSE may be limited among the general population.[19] Therefore, to assist respondents, we provided the following statement at the beginning of the survey:

'The next set of questions are about your participation in muscle-strengthening exercise, sometimes called weight or resistance training. When thinking about muscle-strengthening exercise, we are only interested in exercises that you do during your leisure or free time, and not done as part of your work/job, or as a part of household activities (chores).

The types of muscle-strengthening exercise we are interested in include:

- Using weight machines—typically in a gym or fitness centre.
- Bodyweight exercises—including push-ups or sit-ups.
- Resistance exercises—using free weights like dumbbells or resistance bands.
- Holistic exercises—including Yoga, Tai-Chi or Pilates'.

This phrasing was initially tested with a subset of participants (n=10) to assess readability and comprehension. After this consultation, minor changes were made to the final preamble.

Type

At present, MSE surveillance instruments typically include a wide variety of MSE-related activities grouped into one category.[19] For example, the Behavioral Risk Factor Surveillance System (BRFSS), the most commonly used MSE surveillance instrument,[19] combines diverse activities such as 'using weight machines', 'free weights', 'elastic bands', 'yoga' or 'sit-ups or push-ups'.[30] The limitation of combining all MSE types into a single group is that it is not possible to examine patterns and prevalence of different MSE-related behaviours and, most importantly, determine the relationship between separate MSE types and health. To address this limitation, in the MSEQ, we selected four MSE types. These were:

(1) 'Use of weight machines' (eg, leg press, chest press, lat pulldown); (2) 'Bodyweight exercises' (including push-ups, sit-ups); (3) 'Resistance exercises' (using resistance bands or free weights like dumbbells) and (4) 'Holistic exercises' (including Yoga, Tai-Chi and Pilates)'. The terminology for, and examples of, the MSE types are largely consistent with those within the key texts in this field.[2 27 31] For each of these four MSE types, participants were asked to complete the following items (see online supplemental digital content 1).

Frequency (items 2 and 3)

In MSE surveillance, weekly frequency is the most commonly assessed MSE construct.[19] Accordingly, for comparisons to previous instruments, we included a similar question to that used in the BRFSS survey,[30] where respondents were asked for all MSE: 'How many days, in a usual week, do you do muscle-strengthening exercise?' (item 2). Response selections were: (i) 'none'; (ii) '1'; (iii) '2'; (iv) '3'; (v) '4'; (vi) '5', (vii) '6' and (viii) '7 days'. This question was asked separately for each type to understand the frequency of the four specific MSE types given above (item 3). The response options for this question were the same as all MSE (eg, (i) 'none' to (viii) '7 days').

Duration (item 4)

Despite clinical studies showing a positive dose–response relationship between time spent doing MSE and muscle size and strength,[32] duration is rarely assessed in MSE surveillance.[17] Therefore, to gain a better understanding of this construct, respondents were asked: 'In a usual week please indicate how long you spend doing each of the following types of muscle-strengthening exercise? … in a usual session'. Response options in minutes were: (i) '0'; (ii) 'less than 10'; (iii) '10–20', (iv) '21–30'; (v) '31–40'; (vi) '41–50', (vii) '51–60', (viii) '≥60 min spent in a usual session'. This question was asked separately for each of the four MSE types.

Muscle groups targeted (items 5–8)

The 2008 Physical Activity Guidelines for Americans first introduced the recommendation that when doing MSE, an adult should engage all major muscle groups.[14] This recommendation is based on the clinical evidence that suggests that using several large muscle groups is more effective in maintaining and increasing muscle strength and bone mineral density,[33] compared with using the smaller muscle groups.[34] However, few existing MSE instruments assess muscle groups targeted.[35] To guide which muscle groups to include in the MSEQ, we used the American College of Sports Medicine (ACSM) definitions, which define all major muscle groups as seven separate groups: legs, hips, back, abdomen, chest, shoulders, and arms.[36] Accordingly, in the MSEQ, respondents were asked, 'In a usual week please indicate which muscle groups you use when you do each of the following types of muscle-strengthening exercise?'

Response options of (i) 'yes' or (ii) 'no' were provided for the following seven different muscle groups, and to assist with respondent comprehension, we provided examples of MSE activities that target each group: (i) 'legs (eg, squats, lunges, bridges)'; (ii) 'hips (eg, side leg raises, bridges)'; (iii) 'back (eg, lat pulldown, bent-over row)'; (iv) 'abdomen (eg, crunches, sit-ups)'; (v) 'chest (eg, bench press, push-ups)'; (vi) 'shoulders (eg, lat raise, overhead press)' and (vii) 'arms (eg, bicep curl, tricep dips)'. This question was asked separately for each of the four MSE types.

Intensity (item 9)

Current global physical activity guidelines state: 'Adults should also do muscle-strengthening activities at moderate or greater intensity'.[15] However, current MSE surveillance instruments do not generally assess intensity.[19] To assess this key MSE participation construct, the MSEQ used the previously validated visual analogue scale developed by Robertson *et al*,[37] (See online supplemental digital content 1). Specifically, participants were asked: 'For each of the following types of muscle-strengthening exercises, please say how hard (level of intensity) you usually exercise'. The response options were provided on a 12-point scale: (i) 'not applicable' (ii) '0 extremely easy'; (iii) '1'; (iv) '2 easy'; (v) '3'; (vi) '4 somewhat easy'; (vii) '5'; (viii) '6 somewhat hard'; (ix) '7'; (x) '8 hard'; (xi) '9'; and (xii) '10 extremely hard'. This question was asked for each of the four MSE types.

Test–test reliability assessments

To examine test–test reliability, participants allocated to the reliability sample were sent an individualised survey link approximately 7 days after completing the first online survey. Each participant responded to the same set of questions described above.

Concurrent assessments (7-day MSE log)

To examine concurrent validity, participants allocated to the validity sample were asked to complete a 7-day MSE log. Approximately 7 days after completing the first online survey, participants were sent an individualised link to complete an MSE log for 7 consecutive days. During this week and on each day, participants were asked, 'Did you do any muscle-strengthening exercise today?' (response options: (i) 'yes'; or (ii) 'no'). Those reporting no MSE were automatically taken to the end of the survey, advising them they would receive the invitation to complete the MSE log the next day, or if day 7, they were directed to the end of the survey.

Participants who responded 'yes' to doing MSE on a given day were then asked to respond to the same set of questions described above. All response options were the same as the original survey. Based on the 7-day MSE log responses, we created weekly averages for all MSE and each MSE type for comparison to the original survey. To calculate the average data for the frequency, duration, muscle groups used, and intensity, we collated each

response to these items from the 7-day MSE log, then divided the respective values by the total number of days of reported MSE in the 7-day log. For example, for frequency, the original 'times per usual week' value (from the baseline responses to the MSEQ) was compared with the total times per week value calculated from the 7 day MSE log.

Statistical analysis

Statistical Package for the Social Sciences V.26 (SPSS, IBM) was used to conduct all data analyses. Descriptive statistics were used to describe the characteristics of the participants included in the two individual samples (see table 1), and significance was set at p<0.05 throughout.

To examine the test–retest reliability of the MSEQ items, four statistical tests were used to assess the relative agreement between the participant responses to the initial MSEQ survey (test) against the responses to the follow-up survey (retest). For the continuous variables (frequency, duration and intensity), we used intraclass correlation coefficients (ICC) and Spearman's rank correlation coefficients (Spearman's rho; ρ). For the dichotomous variable (muscle groups used), we used Kappa coefficients (κ) and percentage agreement. We applied a two-way random effects model for ICC analysis, including tests for absolute agreement. We presented ICC and Spearman's r value and its 95% CI for each frequency, duration, and level of intensity question. We present the κ value and 95% CI and the percentage agreement for the muscle groups used. Using previously established ICC correlation coefficient thresholds,[38] a value of 0.00–0.10 suggests virtually no agreement, 0.11–0.40 slight agreement, 0.41–0.60 fair agreement, 0.61–0.80 moderate agreement and ≥0.80 substantial agreement. Applying standard classifications[39] we interpreted Spearman's r as <0.30 low, 0.30–0.50 moderate, or >0.50 high agreement. For κ, we classified the strength of agreement as either <0.00 poor, 0.00–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, or 0.81–1.00 almost perfect.[40] For percentage agreement, we adopted the following classifications:<60% poor, 60%–74% moderate or ≥75% excellent.[41]

To examine the concurrent validity of the MSEQ items on frequency, duration and intensity, ICCs and Spearman's r are presented to show the relative agreement between the responses to the initial MSEQ survey against the 7-day MSE log as the standard. For validity assessments, we used the same ICC and Spearman's r thresholds for the level of agreement used for the reliability.

To investigate the viability of a short version of the MSEQ (see Discussion), we combined the response of all four MSE types (using weight machines, bodyweight exercises, resistance exercises and holistic exercises) to create a fifth type ('all types of muscle-strengthening exercise'). We examined the validity and reliability using the same for each of the individual types described above.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

RESULTS

A summary of the participant characteristics and weekly frequency of MSE for each sub-sample is shown in table 1. Full data were available from 85 participants included in the reliability sample and 54 in the validity sample. Overall, the sociodemographic characteristics were similar for each sample (see online supplemental digital content 3 for a copy of the sociodemographic questions). In brief, over half were female, ~70% aged between 18 and 54 years, ~54% living in Australia and ~60% living in urban areas. Most were university qualified and self-rated their health as good-excellent, just under half were married, and over half were employed and working ≥40 hours in a usual week. While the validity sample met the MSE guideline ≥2 days/week, ~30% of the reliability sample did not.

Test–retest reliability

The results of the test–retest reliability of the MSEQ are shown in tables 2 and 3. For the frequency of MSE days during a usual week, there was a substantial agreement for all four types of MSE (ICC range: 0.85–0.95) and fair agreement when all types were combined (ICC 0.58; 95% CI 0.40 to 0.73). For MSE duration, items for each type showed substantial agreement (ICC range: 0.88–0.96) and moderate agreement for all types combined (ICC 0.69; 95% CI 0.55 to 0.80). For the level of intensity, there was substantial agreement across all four types (ICC range: 0.89–0.93) and moderate agreement for the combined analysis (ICC 0.51; 95% CI 0.31 to 0.68). Spearman's rank correlations were high for all four MSE types, and all types combined for the frequency, duration and level of intensity of MSE (ρ range 0.76–0.91).

For the muscle groups used (see table 3), there was substantial to an almost perfect agreement for using weight machines (κ range 0.61–0.85) for all groups, except for 'abdomen' (κ=0.33; 95% CI −0.01 to 0.65). Moderate to almost perfect agreement was shown for muscle groups used when doing body weight exercises (κ range 0.51–0.83) except for 'back' (κ=0.40; 95% CI 0.14 to 0.64). The agreement for muscle groups using resistance bands or free weights (κ range 0.44–0.84) was similar to body weight exercises. However, the 'chest' agreement was only fair (κ=0.33; 95% CI 0.02 to 0.60). Agreements for holistic exercises were less diverse for all seven muscle groups (moderate to substantial), with κ ranges from 0.57 to 0.76. When all MSE types were combined, the agreement for each of the muscle groups was moderate (κ range 0.41–0.51) except for 'abdomen' (κ=0.35; 95% CI 0.17 to 0.52). Percentage agreement for four types of MSE and the types combined was excellent (range 75.9%–96.3%), except for 'back' when doing

Table 1 Sample characteristics of the test–retest reliability and concurrent validity samples and Weekly frequency of muscle-strengthening exercise

Characteristic	Test–retest reliability sample (n=85) % (n)	Concurrent validity sample (n=54) % (n)
Sex		
Male	40.0 (34)	38.9 (21)
Female	60.0 (51)	61.1 (33)
Age (years)		
18–34	32.9 (28)	37.0 (20)
35–54	38.8 (33)	35.2 (19)
≥55	28.3 (24)	27.8 (15)
Country		
Australia	63.5 (54)	38.9 (21)
Other	36.5 (31)	61.1 (33)
Region		
Urban	62.4 (53)	59.3 (32)
Regional/remote	37.6 (32)	40.7 (22)
Education		
Primary school/some high school	2.4 (2)	7.4 (4)
Year 12 or equivalent	10.6 (9)	3.7 (2)
Higher education (below degree)	11.8 (10)	9.3 (5)
University qualification	75.3 (64)	79.6 (43)
Marital status		
Married	47.1 (40)	53.7 (29)
Not married (defacto/separated/divorced/widowed)	24.7 (21)	18.5 (10)
Never married	28.2 (24)	27.8 (15)
Work situation		
School/university (full time)	20.0 (17)	18.5 (10)
Paid employment/self-employed/unpaid work	58.8 (50)	64.8 (35)
Not working/other	21.2 (18)	16.7 (9)
Self-rated health		
Excellent	20.0 (17)	33.3 (18)
Very good	41.2 (35)	50.0 (27)
Good	27.1 (23)	13.0 (7)
Fair/poor	11.8 (10)	3.7 (2)
Body mass index classification		
Underweight or normal weight	51.8 (44)	48.1 (26)
Overweight	30.6 (26)	29.6 (16)
Obese	16.5 (14)	20.4 (11)
Frequency of muscle-strengthening exercise (days/week)		
0	27.1 (23)	0 (0)
1	3.5 (3)	0 (0)
2	16.5 (14)	13.0 (7)
3	18.8 (16)	25.9 (14)
4	10.6 (9)	16.7 (9)
≥5	23.5 (20)	44.4 (24)

Table 2 Test–retest reliability* of the Muscle-Strengthening Exercise Questionnaire (MSEQ) assessing days per week, duration and level of intensity for each type of muscle-strengthening exercise and all types combined

	Test–retest reliability*	
	Interclass correlations coefficient† (95% CI)	Spearman's r‡ (95% CI)
How many days, in a usual week, do you do muscle-strengthening exercise?	0.92 (0.86 to 0.95)	0.85 (0.72 to 0.94)
Days per week by type		
Use weight machines	0.85 (0.74 to 0.91)	0.79 (0.60 to 0.94)
Body weight exercises	0.93 (0.88 to 0.96)	0.86 (0.73 to 0.93)
Use resistance bands or free weights	0.95 (0.92 to 0.97)	0.87 (0.73 to 0.95)
Holistic exercises	0.95 (0.91 to 0.97)	0.83 (0.65 to 0.95)
All types of muscle-strengthening exercise	0.58 (0.40 to 0.73)	0.91 (0.85 to 0.95)
Duration (minutes/session)		
Use weight machines	0.90 (0.82 to 0.94)	0.82 (0.67 to 0.93)
Body weight exercises	0.88 (0.79 to 0.93)	0.77 (0.60 to 0.87)
Use resistance bands or free weights	0.89 (0.82 to 0.94)	0.79 (0.62 to 0.90)
Holistic exercises	0.96 (0.94 to 0.98)	0.88 (0.74 to 0.96)
All types of muscle-strengthening exercise	0.69 (0.55 to 0.80)	0.89 (0.78 to 0.95)
Level of intensity		
Use weight machines	0.90 (0.83 to 0.94)	0.82 (0.66 to 0.95)
Body weight exercises	0.93 (0.89 to 0.96)	0.78 (0.61 to 0.90)
Use resistance bands or free weights	0.93 (0.88 to 0.96)	0.76 (0.59 to 0.88)
Holistic exercises	0.89 (0.82 to 0.94)	0.81 (0.63 to 0.94)
All types of muscle-strengthening exercise	0.51 (0.31 to 0.68)	0.82 (0.66 to 0.92)

*Test and retest of MSEQ were conducted a maximum of 14 days apart.
†Intraclass correlations coefficient between test and retest and its 95% CI.
‡Spearman's rank correlation between test and retest and its 95% CI.

body weight exercises and 'abdomen' when using resistance bands or free weights (range 72.2%–72.7%).

Concurrent validity

The concurrent validity of the MSEQ are shown in table 4. Compared with the 7-day MSE log, days in a usual week displayed fair-to-moderate concurrent validity for each MSE type and all types combined (ICC range: 0.56–0.78). The results for Spearman's rank correlation analyses were 'high' for the frequency of all four MSE types and all types combined (ρ range 0.58–0.77). For the duration, the MSEQ items for the use of machines, resistance bands, holistic exercise and all types combined showed a fair-to-moderate agreement (ICC range: 0.46–0.78) and slight agreement for bodyweight exercise (ICC=0.39; 95% CI −0.04 to 0.64). The correlation was high for the duration of all types combined (ρ=0.73; 95% CI 0.59 to 0.83), with similar results observed for the individual MSE types (ρ range 0.73–0.74), except for body weight exercises that were moderate (ρ=0.34; 95% CI 0.08 to 0.56). For the level of intensity, a moderate agreement was shown for the use of weight machines, resistance bands/free weights, and holistic exercises (ICC range: 0.69–0.78), and slight agreement was observed for bodyweight exercises and all types combined (ICC range: 0.18–0.35). The results of the Spearman's rank correlation analyses for the level of intensity were mixed, ranging from moderate (body weight exercises: ρ=0.30; 95% CI 0.03 to 0.55) to high (ρ range 0.63–0.71) for the individual MSE types, and a moderate result for all MSE types combined (ρ=0.45; 95% CI 0.18 to 0.67).

DISCUSSION

This study describes the test–retest reliability and concurrent validity of a newly developed online survey instrument assessing muscle-strengthening exercise. The MSEQ was specifically designed to assess multiple MSE participation constructs (eg, frequency, duration, intensity, muscle groups) across different types of MSE (eg, use weight machines, body weight exercises, use resistance bands or free weights, and holistic exercises) in adults. Overall, among our active, young, and well-educated sample, the MSEQ showed substantial test–retest reliability and adequate validity when using a 7-day MSE log as the standard. While these findings need to be replicated in studies with a more representative sample, this

Table 3 Test–retest reliability* of the Muscle-Strengthening Exercise Questionnaire (MSEQ) items assessing muscle groups targeted for each type of muscle-strengthening exercise and all types combined

Type of muscle-strengthening exercise	Kappa statistic† (95% CI)	% Agreement
Use weight machines		
Legs	0.70 (0.49 to 0.88)	85.2
Hips	0.61 (0.34 to 0.85)	87.0
Back	0.74 (0.54 to 0.89)	87.0
Abdomen	0.33 (-0.01 to 0.65)	81.5
Chest	0.78 (0.59 to 0.93)	88.9
Shoulders	0.85 (0.67 to 0.96)	92.6
Arms	0.81 (0.63 to 0.96)	90.7
Body weight exercises		
Legs	0.51 (0.19 to 0.79)	85.2
Hips	0.83 (0.65 to 0.96)	92.6
Back	0.40 (0.14 to 0.64)	72.7
Abdomen	0.71 (0.41 to 0.93)	90.7
Chest	0.54 (0.23 to 0.80)	85.2
Shoulders	0.53 (0.26 to 0.74)	77.8
Arms	0.57 (0.33 to 0.79)	79.6
Use resistance bands or free weights		
Legs	0.55 (0.23 to 0.81)	87.0
Hips	0.66 (0.45 to 0.85)	83.3
Back	0.49 (0.16 to 0.74)	81.5
Abdomen	0.44 (0.19 to 0.67)	72.2
Chest	0.33 (0.02 to 0.60)	75.9
Shoulders	0.84 (0.55 to 1.00)	96.3
Arms	0.63 (0.22 to 0.92)	92.6
Holistic exercises		
Legs	0.58 (0.34 to 0.79)	79.6
Hips	0.61 (0.38 to 0.81)	81.5
Back	0.74 (0.50 to 0.91)	88.9
Abdomen	0.66 (0.45 to 0.85)	83.3
Chest	0.69 (0.44 to 0.88)	87.0
Shoulders	0.57 (0.34 to 0.78)	81.5
Arms	0.76 (0.53 to 0.94)	90.7
All types of muscle-strengthening exercise		
Legs	0.41 (0.22 to 0.58)	79.6
Hips	0.51 (0.34 to 0.66)	79.6
Back	0.45 (0.26 to 0.63)	81.5
Abdomen	0.35 (0.17 to 0.52)	77.8
Chest	0.41 (0.22 to 0.59)	77.8
Shoulders	0.44 (0.24 to 0.60)	83.3
Arms	0.47 (0.29 to 0.64)	77.8

*Test and retest of MSEQ were conducted a maximum of 14 days apart.
†Kappa coefficient of agreement between test and retest and its 95% CI.

study suggests that the MSEQ has potential for use in future physical activity surveillance.

Few studies have reported on the reliability and validity of existing MSE survey instruments, but[19] comparing our findings to similar studies is limited. For weekly MSE frequency, the MSEQ shows similar reliability and stronger validity compared with the MSE item from the BRFSS using a physical activity log.[30] However, the MSEQ expands on the BRFSS by assessing the frequency of four different types of MSE and muscle groups targeted, duration and intensity. A recent study examined the reliability and validity of MSE items (using a 7-day diary) from the Cancer Prevention Study-3.[42] That study assessed the MSE frequency and duration of similar MSE types using a single MSE question. In comparison, the MSEQ shows stronger reliability and similar validity to that study. While the study[42] included similar MSE types, expanding on the BRFSS, the MSEQ is more extensive as it allows for the additional assessment of intensity and muscle groups targeted.

Compared with commonly used MVPA surveillance instruments, the MSEQ showed stronger reliability and validity. For example, compared with the frequency and duration items in the Global Physical Activity Questionnaire (GPAQ), the MSEQ was superior for both 7-day test–retest reliability and concurrent validity using activity logs.[43] Validity of the MSEQ is stronger when compared with the leisure-time frequency and duration items contained in the International Physical Activity Questionnaire (IPAQ)-Long.[44] The potential reason for stronger reliability and validity observed in MSE, compared with MVPA, is likely because MSE is easier to recall and a more memorable physical activity.[17] Moreover, that in this study a high percentage of participants also meet the MSE guidelines.

When designing the MSEQ, we decided to solely target MSE-related behaviours within the context of leisure time, and consequently not to include any activities accrued during occupational (eg, labouring/lifting) and domestic tasks (eg, carry shopping bags, gardening). This decision was to avoid any potential misclassification of other MSE-related behaviours. Furthermore, occupation-related physical activity is often undertaken at low/moderate intensity for long durations with limited time for recovery.[45] Moreover, it has been argued that the repetitive nature of undertaking MSE outside the context of leisure time may negatively influence health. For example, MSE within the occupational and domestic context may result in an increased risk of musculoskeletal disorders (eg, back, shoulder, neck injuries/pain) and arthritis/rheumatic diseases (eg, osteoarthritis, rheumatoid arthritis).[46]

MSEQ: short and long format

For potential use in future health surveillance, we adapted the MSEQ to be consistent with existing self-reported physical activity surveillance instruments, such as the widely used IPAQ[20] and GPAQ.[47] Specifically, we

Table 4 Concurrent validity* of the Muscle-Strengthening Exercise Questionnaire (MSEQ) assessing days per week, duration and level of intensity for each type of muscle-strengthening exercise and all types combined

	Concurrent validity*	
	Interclass correlations coefficient† (95% CI)	Spearman's r‡ (95% CI)
How many days, in a usual week, do you do muscle-strengthening exercise?	0.64 (0.37 to 0.80)	0.49 (0.25 to 0.70)
Days per week by type		
Use weight machines	0.69 (0.16 to 0.86)	0.76 (0.59 to 0.89)
Body weight exercises	0.74 (0.56 to 0.85)	0.59 (0.36 to 0.76)
Use resistance bands or free weights	0.72 (0.35 to 0.86)	0.64 (0.38 to 0.84)
Holistic exercises	0.78 (0.45 to 0.89)	0.77 (0.60 to 0.89)
All types of muscle-strengthening exercise	0.56 (0.37 to 0.71)	0.58 (0.35 to 0.76)
Duration (minutes/session)		
Use weight machines	0.60 (0.18 to 0.79)	0.73 (0.56 to 0.87)
Body weight exercises	0.39 (–0.04 to 0.64)	0.34 (0.08 to 0.56)
Use resistance bands or free weights	0.78 (0.37 to 0.90)	0.74 (0.55 to 0.86)
Holistic exercises	0.70 (0.38 to 0.84)	0.73 (0.54 to 0.87)
All types of muscle-strengthening exercise	0.46 (0.24 to 0.64)	0.73 (0.59 to 0.83)
Level of intensity		
Use weight machines	0.73 (0.38 to 0.87)	0.71 (0.52 to 0.85)
Body weight exercises	0.35 (–0.07 to 0.61)	0.30 (0.03 to 0.55)
Use resistance bands or free weights	0.78 (0.21 to 0.91)	0.68 (0.52 to 0.82)
Holistic exercises	0.69 (0.43 to 0.83)	0.63 (0.41 to 0.78)
All types of muscle-strengthening exercise	0.18 (–0.06 to 0.42)	0.45 (0.18 to 0.67)

*To assess the concurrent validity, MSEQ baseline responses were compared with a 7-day MSE log.
†Intraclass correlations coefficient between test and retest and its 95% CI.
‡Spearman's rank correlation between test and retest and its 95% CI.

developed two versions of the MSEQ, the MSEQ-Short and MSEQ-Long, each designed to be used in either a self-administered or interview-administered format (full versions shown in online supplemental digital content 2). The MSEQ-Short is a brief 6-item instrument that assesses any engagement in MSE ('yes' or 'no'), the usual weekly frequency (number of days), duration (minutes spent), intensity (range from 0 to 10), type of muscle-strengthening exercise ('yes' or 'no' response to the four types of MSE), and muscle groups targeted ('yes' or 'no' response to seven muscle groups). The MSEQ-Long is a 20-item instrument that assesses the usual weekly frequency, duration, intensity, and the muscle groups targeted (similar responses as in MSEQ-Short), separately for all four types of MSE (weight machines, bodyweight exercises, resistance exercises and holistic exercises). As shown in online supplemental digital content 3, table, the reliability and validity of the MSEQ-Short items displayed mostly fair-to-moderate agreement and was moderate-to-high for most of the MSEQ-Long items (see online supplemental digital content 4, table). These preliminary data suggest that the MSEQ-Short and MSEQ-Long have promise as a standardised MSE surveillance instrument. However, we urge caution, as both instruments have not yet been tested for reliability and validity in the format provided in online supplemental digital content 2. We now call for future studies to assess the psychometric properties of the MSEQ-Short and MSEQ-Long, with diverse population subgroups (eg, older adults, those from differing income/education levels) and translated into different languages.

Limitations

A key limitation of this study was our recruitment of a non-representative sample, which is likely to affect the generalisability of our findings. A further limitation was the self-reported nature of the online responses to the survey. There is a risk of responder recall bias (eg, social desirability or over-reporting/under-reporting of actual behaviour). However, there is no device-based measurement available for the assessment of MSE. This behaviour is routinely assessed by self-report in physical activity surveillance. A further limitation is that we were unable to establish the validity of the muscle groups targeted items, as this was not possible when comparing a single 'yes' or 'no' response in the 'main survey' to the daily 'yes' or 'no' response in the 7-day diary. A key strength of this study is that it is one of the first to assess the test–retest

reliability and concurrent validity of questions specifically developed to explore the constructs of MSE beyond frequency and duration. Moreover, the inclusion of the assessment of MSE intensity is a unique and important component that is not currently well understood.[48]

CONCLUSION

The newly developed MSEQ displayed adequate test–retest reliability and concurrent validity in assessing multiple MSE participating constructs. Given that the current study included a sample of young, well educated, and active adults, further research is needed to examine whether these findings are generalisable to more representative samples.

Twitter Jane Shakespear-Druery @JaneDruery, Katrien De Cocker @KatrienDeCocker, Stuart J H Biddle @stuart_biddle and Jason Bennie @JasonBennie

Acknowledgements This research has been supported by an Australian Government Research Training Programme Scholarship and a University of Southern Queensland Research Training Program Stipend Scholarship.

Contributors JS-D, KDC, SJHB and JB conceptualised the study and developed the research plan. JS-D created the online survey, collected all data and conducted the data analysis. JSD and JB drafted the initial manuscript. JB, KDC and SJHB provided guidance on the study and critically reviewed the manuscript. JB is acting as guarantor. All authors read and approved the final version of the manuscript, and agreed with the order of presentation of the authors.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Ethical approval was obtained from the University of Southern Queensland Human Ethics Committee in May 2020 (H20REA233). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request from the corresponding author.

ORCID iDs

Jane Shakespear-Druery http://orcid.org/0000-0002-4111-6494
Katrien De Cocker http://orcid.org/0000-0001-7510-4419
Stuart J H Biddle http://orcid.org/0000-0002-7663-6895
Jason Bennie http://orcid.org/0000-0002-8668-8998

REFERENCES

1 Piercy KL, Troiano RP, Ballard RM, *et al*. The physical activity guidelines for Americans. *JAMA* 2018;320:2020–8.
2 Garber CE, Blissmer B, Deschenes MR, *et al*. American College of sports medicine position stand. quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. *Med Sci Sports Exerc* 2011;43:1334–59.
3 Ashton RE, Tew GA, Aning JJ, *et al*. Effects of short-term, medium-term and long-term resistance exercise training on cardiometabolic health outcomes in adults: systematic review with meta-analysis. *Br J Sports Med* 2020;54:1–9.
4 Grgic J, Schoenfeld BJ, Davies TB, *et al*. Effect of resistance training frequency on gains in muscular strength: a systematic review and meta-analysis. *Sports Med* 2018;48:1207–20.
5 Gordon BR, McDowell CP, Hallgren M, *et al*. Association of efficacy of resistance exercise training with depressive symptoms: meta-analysis and meta-regression analysis of randomized clinical trials. *JAMA Psychiatry* 2018;75:566–76.
6 Khalafi M, Malandish A, Rosenkranz SK, *et al*. Effect of resistance training with and without caloric restriction on visceral fat: a systemic review and meta-analysis. *Obes Rev* 2021;22:e13275.
7 Saeidifard F, Medina-Inojosa JR, West CP, *et al*. The association of resistance training with mortality: a systematic review and meta-analysis. *Eur J Prev Cardiol* 2019;26:1647–65.
8 Nascimento W, Ferrari G, Martins CB, *et al*. Muscle-strengthening activities and cancer incidence and mortality: a systematic review and meta-analysis of observational studies. *Int J Behav Nutr Phys Act* 2021;18:69.
9 Shiroma EJ, Cook NR, Manson JE, *et al*. Strength training and the risk of type 2 diabetes and cardiovascular disease. *Medicine Science in Sports Exercise* 2017;49:40–6.
10 Grøntved A, Rimm EB, Willett WC, *et al*. A prospective study of weight training and risk of type 2 diabetes mellitus in men. *Arch Intern Med* 2012;172:1306–12.
11 Grøntved A, Pan A, Mekary RA, *et al*. Muscle-strengthening and conditioning activities and risk of type 2 diabetes: a prospective study in two cohorts of US women. *PLoS Med* 2014;11:15.
12 Brellenthin AG, Lee D-C, Bennie JA, *et al*. Resistance exercise, alone and in combination with aerobic exercise, and obesity in Dallas, Texas, us: a prospective cohort study. *PLoS Med* 2021;18:e1003687.
13 Mekary RA, Grøntved A, Despres J-P, *et al*. Weight training, aerobic physical activities, and long-term waist circumference change in men. *Obesity* 2015;23:461–7.
14 U.S. Department of Health and Human Services. *2008 physical activity guidelines for Americans*. Washington, DC: U.S. Department of Health and Human Services, 2008.
15 Bull FC, Al-Ansari SS, Biddle S, *et al*. World Health organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med* 2020;54:1451.
16 Strain T, Fitzsimons C, Kelly P, *et al*. The forgotten guidelines: cross-sectional analysis of participation in muscle strengthening and balance & co-ordination activities by adults and older adults in Scotland. *BMC Public Health* 2016;16:1108.
17 Bennie JA, Shakespear-Druery J, De Cocker K. Muscle-strengthening exercise epidemiology: a new frontier in chronic disease prevention. *Sports Med Open* 2020;6:1–8.
18 Milton K, Varela AR, Strain T, *et al*. A review of global surveillance on the muscle strengthening and balance elements of physical activity recommendations. *J Frailty Sarcopenia Falls* 2018;3:114–24.
19 Shakespear-Druery J, De Cocker K, Biddle SJH, *et al*. Assessment of muscle-strengthening exercise in public health surveillance for adults: a systematic review. *Prev Med* 2021;148:106566.
20 Craig CL, Marshall AL, Sjöström M, *et al*. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;35:1381–95.
21 Armstrong T, Bull F. Development of the world Health organization global physical activity questionnaire (GPAQ). *J Public Health* 2006;14:66–70.
22 Prince SA, LeBlanc AG, Colley RC, *et al*. Measurement of sedentary behaviour in population health surveys: a review and recommendations. *PeerJ* 2017;5:e4130.
23 Nunes JP, Grgic J, Cunha PM, *et al*. What influence does resistance exercise order have on muscular strength gains and muscle hypertrophy? A systematic review and meta-analysis. *Eur J Sport Sci* 2021;21:149–57.
24 Paoli A, Gentil P, Moro T, *et al*. Resistance training with single vs. multi-joint exercises at equal total load volume: effects on body composition, cardiorespiratory fitness, and muscle strength. *Front Physiol* 2017;8:1105.
25 Troiano RP, Stamatakis E, Bull FC. How can global physical activity surveillance adapt to evolving physical activity guidelines? needs, challenges and future directions. *Br J Sports Med* 2020;54:1468.
26 Ratamess NA, Alvar BA, Evetoch TE. Progression models in resistance training for healthy adults. *Med Sci Sports Exerc* 2009;41:687–708.
27 U.S. Department of Health and Human Services. *Physical activity guidelines for Americans*. 2nd edn. Washington, DC: U.S. Department of Health and Human Services, 2018.
28 Baranowski T. Validity and reliability of self report measures of physical activity: an information-processing perspective. *Res Q Exerc Sport* 1988;59:314–27.
29 Bennie JA, Kolbe-Alexander T, Seghers J, *et al*. Trends in muscle-strengthening exercise among nationally representative samples of United States adults between 2011 and 2017. *J Phys Act Health* 2020;17:512.

30 Yore MM, Ham SA, Ainsworth BE, *et al*. Reliability and validity of the instrument used in BRFSS to assess physical activity. *Med Sci Sports Exerc* 2007;39:1267–74.
31 American College of Sports Medicine. American College of sports medicine position stand. progression models in resistance training for healthy adults. *Med Sci Sports Exerc* 2009;41:687–708.
32 Dankel SJ, Mattocks KT, Jessee MB, *et al*. Do metabolites that are produced during resistance exercise enhance muscle hypertrophy? *Eur J Appl Physiol* 2017;117:2125–35.
33 Zhao R, Zhao M, Xu Z. The effects of differing resistance training modes on the preservation of bone mineral density in postmenopausal women: a meta-analysis. *Osteoporos Int* 2015;26:1605–18.
34 Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee Report. In: *Physical activity guidelines Advisory Committee report*. Washington, DC: U.S. Department of health and Human Services, 2008.
35 Loustalot F, Carlson SA, Kruger J, *et al*. Muscle-strengthening activities and participation among adults in the United States. *Res Q Exerc Sport* 2013;84:30–8.
36 Pollock ML, Gaesser GA, Butcher JD. American College of sports medicine position stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Med Sci Sports Exerc* 1998;30:975–91.
37 Robertson RJ, Goss FL, Rutkowski J, *et al*. Concurrent validation of the Omni perceived exertion scale for resistance exercise. *Med Sci Sports Exerc* 2003;35:333–41.
38 Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res* 1998;7:301–17.
39 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Lawrence Erlbaum Associates, 1988.
40 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.
41 Saelens BE, Frank LD, Auffrey C, *et al*. Measuring physical environments of parks and playgrounds: EAPRS instrument development and inter-rater reliability. *J Phys Act Health* 2006;3:s190–207.
42 Subbiah K, Rees-Punia E, Patel AV. Reliability and validity of self-reported Muscle-strengthening exercise in the cancer prevention Study-3. *Med Sci Sports Exerc* 2021;53:888–93.
43 Keating XD, Zhou K, Liu X, *et al*. Reliability and concurrent validity of global physical activity questionnaire (GPAQ): a systematic review. *Int J Environ Res Public Health* 2019;16. doi:10.3390/ijerph16214128. [Epub ahead of print: 26 10 2019].
44 Hagströmer M, Oja P, Sjöström M. The International physical activity questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr* 2006;9:755–62.
45 Holtermann A, Krause N, van der Beek AJ, *et al*. The physical activity paradox: six reasons why occupational physical activity (opa) does not confer the cardiovascular health benefits that leisure time physical activity does. *Br J Sports Med* 2018;52:149.
46 Centers for Disease Control and Prevention. Work-related musculoskeletal disorders & ergonomics Atlanta, Georgia: Centers for Disease Control and Prevention. Available: https://www.cdc.gov/workplacehealthpromotion/health-strategies/musculoskeletal-disorders/index.html [Accessed 27 Jul 2021].
47 World Health Organization. Global Physical Activity Questionnaire (GPAQ). In: *Prevention of noncommunicable diseases department*. Geneva, Switzerland: World Health Organization, 2020.
48 Giovannucci EL, Rezende LFM, Lee DH. Muscle-strengthening activities and risk of cardiovascular disease, type 2 diabetes, cancer and mortality: a review of prospective cohort studies. *J Intern Med* 2021;290:789-805.