Psychology of Sport and Exercise 15 (2014) 446–452

Contents lists available at ScienceDirect

Psychology of Sport and Exercise

journal homepage: www.elsevier.com/locate/psychsport

# Comparing sport motivation scales: A response to Pelletier et al.

Chris Lonsdale [a,b,*], Ken Hodge [c], Elaine A. Hargreaves [c], Johan Y.Y. Ng [d]

[a] Australian Catholic University, Australia
[b] University of Western Sydney, Australia
[c] University of Otago, New Zealand
[d] The Chinese University of Hong Kong, Hong Kong

ARTICLE INFO

*Article history:*
Received 6 May 2013
Received in revised form
6 March 2014
Accepted 19 March 2014
Available online 13 April 2014

*Keywords:*
Self-determination theory
Reliability
Validity
Scale development

ABSTRACT

*Objectives:* Pelletier, Rocchi, Vallerand, Deci, and Ryan (2013) proposed a new version of the Sport Motivation Scale (SMS; Pelletier, Fortier, Vallerand, Tuson, & Blais, 1995) as a measure of different types of behavioral regulations in sport, as outlined in self-determination theory (Ryan & Deci, 2000). They examined various aspects of reliability and validity of scale scores, and concluded that the new scale performs better than the original version. They also claimed that the SMS-II is superior to other measures of motivation in sport, including the Behavioral Regulation in Sport Questionnaire (BRSQ) developed by Lonsdale, Hodge, and Rose (2008). By comparing the evidence presented in papers by Pelletier et al. and Lonsdale et al., our objective was to examine the relative merits and shortcomings of the two measures and suggest directions for future research into sport motivation measurement.
*Conclusions:* Both the SMS-II and BRSQ have shown relative strengths and weaknesses. Overall, the construct validity evidence of scores derived from the two measures was similar. There is insufficient information to support the claim that one scale is superior to the other. Researchers are encouraged to make direct comparisons by administering both measures to the same group of participants in future studies.


Pelletier et al. (2013) examined the validity and reliability of scores derived from the Sport Motivation Scale II (SMS-II). The SMS-II, a revision to the Sport Motivation Scale (SMS; Pelletier, Fortier, Vallerand, Tuson, & Blais, 1995), is based on self-determination theory (SDT; Ryan & Deci, 2000), and was designed to measure different types of behavioral regulations in sport. Pelletier et al. conducted a two-staged study and stated they had overcome the problems with the original SMS, concluding that the SMS-II "performs as well as or better than the original scale" (p. 338). Furthermore, Pelletier et al. concluded that compared with the "BRSQ [Behavioral Regulation in Sport Questionnaire] developed by Lonsdale, Hodge, and Rose (2008), we think that the SMS-II represents a scale that better addressed the limitations observed with the original SMS and that showed more consistent results with SDT" (p. 339). In the current paper we examine the evidence related to this claim and suggest future directions for research that may improve sport motivation measurement in the future.[1]

In 2008, we published "The Behavioral Regulation in Sport Questionnaire: Instrument Development and Initial Validity Evidence" (Lonsdale et al., 2008). The measure described in that report was designed to assess sport motivation, as outlined in SDT (see Pelletier et al. (2013) or Lonsdale et al. (2008) for a description of relevant aspects of the theory, including definitions of motivational constructs). One issue with the BRSQ that we acknowledged was that although the majority of evidence showed that there was a distinction between subscales measuring intrinsic motivation, autonomous extrinsic motivation (integrated and identified regulation), controlled motivation (external and introjected regulation), and amotivation, some evidence suggested that scores of integrated and identified regulation were not distinctive. Similarly, external and introjected regulation scores were correlated and showed similar correlations with some constructs (e.g., other types of behavioral regulations). Recently, Pelletier et al. (2013) drew conclusions somewhat similar to our own regarding the nomological validity of BRSQ scores. They wrote,

> "Overall, although [Lonsdale et al.'s (2008)] results showed support for the distinctions between the [BRSQ] self-determined subscales (intrinsic motivation and identified and integrated regulation) and the non-self-determined subscales (external and introjected regulation), the finer discrimination

* Corresponding author. Australian Catholic University, Australia.
*E-mail address:* chris.lonsdale@acu.edu.au (C. Lonsdale).

[1] In the interests of full disclosure, we note that this comparison study involving the BRSQ, the SMS, and the SMS-6 was not part of our original submission to the *Journal of Sport & Exercise Psychology*. At the suggestion of an anonymous reviewer and the editor we collected additional data for this comparison. We thank these individuals for this suggestion, as this direct comparison made our report much stronger.

http://dx.doi.org/10.1016/j.psychsport.2014.03.006
1469-0292/

within each type of category appears to be lacking … there was a lack of discrimination between external and introjected regulation scores in terms of their relationships with amotivation; identified and integrated regulation subscales both had similar high correlations with intrinsic motivation; and there was a lack of discrimination between the self-determined subscales (intrinsic motivation and identified and integrated regulation) and the concepts of flow and burnout." (p. 331).

After acknowledging the limitations of the original SMS and voicing their concerns regarding nomological validity related to some of the BRSQ scores, Pelletier et al. (2013) decided to revise the SMS. They conducted two studies to evaluate the reliability and validity of scores derived from this new measure, entitled the SMS-II. They concluded that compared with the BRSQ, the SMS-II addressed the shortcomings of the SMS better. However, the design of their study allowed them to only directly compare the original SMS and the SMS-II, but not the BRSQ, as it was not included in their data collection. In the absence of direct evidence collected from the same sample of participants, it is our opinion that firm conclusions regarding the relative merits of the SMS-II and the BRSQ are premature.

In this paper we examine the evidence supporting Pelletier et al.'s (2013) claim that the SMS-II is superior to the BRSQ. Currently, comparisons between the SMS-II and the BRSQ and other sport motivation measures (e.g., SMS-6; Mallett, Kawabata, Newcombe, Otero-Forero, & Jackson, 2007) can only be achieved by comparing data collected from different samples. This procedure is clearly not ideal, but at this stage we feel it is necessary to debate the merits of Pelletier et al.'s conclusions. We aim to create an opportunity to enter a respectful, constructive discussion of ways to improve the measurement of SDT constructs in the sport setting.

## Comparison of BRSQ and SMS-II reliability and validity evidence

In the absence of data pertaining to a direct comparison of scores derived using the BRSQ and SMS-II, we compared results presented in the Pelletier et al. (2013; Study 2) and Lonsdale et al. (2008; Study 3) papers. Specifically, we compared the results of reliability and validity testing performed on data gathered from two samples. These two samples were chosen because of their relatively similar mean age (approximately two years older in the New Zealand sample). We decided not to present detailed information from the adult athletes from Pelletier et al.'s Study 1 (mean age = 40.44 years) or the elite athletes from the New Zealand Academy of Sport in our Study 2 (mean age = 25.9 years). While the evidence gathered from these samples was largely similar to that gathered from the adolescent and young adult samples, we felt it was appropriate to limit the sample variation as much as possible. Also, limiting the samples allowed us to present a reasonable amount of data for the reader to digest. Indeed, presenting the results from all analyses in our paper and Pelletier et al.'s would not have added greatly to the discussion and may have made an already nuanced comparison practically inscrutable.

### *Internal consistency*

Both studies examined internal consistency using Cronbach alpha coefficients ($\alpha$). The SMS-II subscale scores had alpha coefficients that ranged from .73 to .86. The BRSQ subscale scores ranged from .76 to .91. Thus, both questionnaires produced scores that would typically be considered internally consistent when measuring these constructs (Nunnally, 1978).

### *Factorial validity*

A comparison of fit indices is limited by the fact that SMS-II and BRSQ have different numbers of items per factor and thus non-equivalent models. Nonetheless, an overview of model fit from confirmatory factor analyses of scale scores from the two studies is presented below with reference to commonly accepted cut-off criteria used to judge model fit. Both models had a significant $\chi^2$ at $p < .001$, suggesting a lack of model fit. However, as the $\chi^2$ is sensitive to sample sizes, model fit based on other fit indices were considered. Hu and Bentler (1999) suggested cut-off criteria for the RMSEA ($\leq$.06), CFI ($\geq$.95) and TLI ($\geq$.95) that have been widely adopted in recent years. Previously, less stringent criteria (RMSEA $\leq$ .08, CFI $\geq$ .90, and TLI $\geq$ .90) were often employed. BRSQ fit statistics were generally strong and met Hu and Bentler's criteria (RMSEA = .07, 90% CI [.06, .08], CFI = .97, TLI = .97). Approximate fit statistics from the SMS-II were also strong, but in some instances fell below Hu and Bentler's suggested cut-off scores (RMSEA = .07, 90% CI [.05, .08], CFI = .94 and TLI = .92). Pelletier et al. (2013) also reported NFI = .90, an index that we did not examine with the BRSQ data (Lonsdale et al., 2008). For the readers' interest, NFI for the analysis using BRSQ scores was .96. It is noteworthy that neither report (Lonsdale et al., 2008; Pelletier et al., 2013) included the SRMR index that Hu and Bentler (1999) showed is particularly important when assessing model fit. For the reader's interest, the SRMR result from our 2008 study was .08, which meets the criterion for good fit ($\leq$.08). Overall, the results generally supported the factorial validity of scores derived from both the BRSQ and the SMS-II, with the BRSQ model fit surpassing Hu and Bentler's cut-off criteria.

### *Nomological validity*

Both studies included examinations of nomological validity that focused on inspection of inter-factor correlations (simplex structure) and correlations between motivation scores and theoretically related constructs. We consider correlations among autonomous motivation scores, then inspect the relations between autonomous motivation scores and controlled motivation, and then scrutinize the correlations among controlled motivation scores. According to SDT, scores representing constructs that are closer together on the self-determination continuum (see Pelletier et al., 2013) should be more strongly and positively correlated (i.e., simplex structure). More distal constructs would be expected to show weaker positive or stronger negative correlations. Finally, we conclude our examination of nomological validity by examining relations between motivation scores and measures of theoretically related constructs. According to SDT, more autonomous motives should be more strongly and positively related with adaptive outcomes, compared with more controlled motives that are expected to be more strongly and positively correlated with maladaptive outcomes. In all instances where we compared correlations across samples, we employed a $z$-test: $(r'_1 - r'_2)/\sqrt{(1/(n_1-1)) + (1/(n_2-1))}$ (Clark-Carter, 2005), where $r'$ represents a Fisher transformation of the Pearson correlation. When comparing correlations within a sample (e.g., comparing the correlation between intrinsic motivation and identified regulation with the correlation between intrinsic motivation and integrated regulation) we followed procedures outlined by Steiger (1980). Details of all analyses can be obtained from the first author.

#### *Relations among autonomous motivation scores*

Pelletier et al. (2013) criticized the simplex structure of the scores derived from the BRSQ subscales intended to represent autonomous forms of extrinsic motivation. Specifically, they

stated that, "identified and integrated regulation subscales both had similar high correlations with intrinsic motivation" (p. 331). As shown in Table 1, BRSQ-Intrinsic Motivation scores showed moderate to strong correlations with integrated regulation ($\Phi = .49$) and identified regulation ($\Phi = .42$). Compared with the BRSQ results, the SMS-II intrinsic motivation scores showed stronger correlations ($p < .05$) with integrated regulation ($\Phi = .78$) and identified regulation ($\Phi = .68$; Pelletier et al., 2013). While the 95% CIs associated with these SMS-II correlations likely do not encompass unity, Tabachnick and Fidell (2007) have suggested that when correlations exceed .70, the distinctiveness of the scores must be questioned. Also, when correlations of this magnitude are present, multi-collinearity is a concern in analyses such as multiple regression (Wilson, Sabiston, Mack, & Blanchard, 2012). Thus, it is doubtful that the three autonomous motivation scores from the SMS-II could be entered into the same regression equation without causing problems with estimation and/or interpretation.

Regarding simplex structure, both the BRSQ and the SMS-II datasets showed a relationship between intrinsic motivation and integrated regulation that was stronger ($p < .05$) than the relationship between intrinsic motivation and identified regulation (a requirement of simplex structure). BRSQ data also indicated that, as hypothesized, the relationship between identified regulation and integrated regulation was stronger than the relationship between identified regulation and intrinsic motivation ($p < .05$). This was not the case in the SMS-II data, as the two correlations were not significantly different ($p = .21$). Thus, support for simplex structure was clearer in the BRSQ data than the SMS-II.

Regarding the strength of correlations, the BRSQ autonomous motivation scores (intrinsic motivation, integrated regulation, and identified regulation) were less strongly correlated than the SMS-II scores ($p < .05$). As a result, Pelletier et al.'s (2013) criticism of the BRSQ intrinsic motivation subscale showing "high" correlations with identified and integrated regulation subscales is unfounded given that the correlations were less than .50 (i.e., the threshold of "strong" suggested by Cohen, 1988). The corresponding SMS-II correlations were, in fact, higher than those found using the BRSQ.

*Relations between scores representing autonomous and controlled forms of motivation*

We examined the relations between autonomous motivation scores and controlled motivation subscale scores. Relations between BRSQ intrinsic motivation and the controlled BRSQ subscale scores followed the simplex pattern (all $p < .05$). However, integrated and identified regulation showed equivalent $\Phi$ correlations with introjected regulation scores ($p = .26$). They also had similar correlations with external regulation scores ($p = 1.00$) and amotivation scores ($p = .41$). This finding prompted us to suggest that "identified and integrated regulation factor scores had similar correlations with the other factors" (Lonsdale et al., 2008, p. 343).

With respect to the SMS-II, the intrinsic motivation subscale produced correlations with the controlled motivation scores that were significantly different from each other (all $p < .05$). However, not all of the SMS-II correlations supported the hypothesized simplex pattern (see Table 2). For example, the correlation between introjected regulation and integrated regulation was significantly stronger ($p < .05$) than the correlation between introjected regulation and identified regulation. This finding is contrary to SDT-based hypotheses. Also, the correlation between external regulation and integrated regulation was similar to the correlation between the external regulation and identified regulation ($p = .30$). Also, the relationship between amotivation and identified regulation was similar to the relationship between amotivation and integrated regulation ($p = .17$). This finding suggested that the SMS-II does not clearly differentiate between the identified and integrated regulation concepts, at least not in a manner that is consistent with theory.

In summary, neither the BRSQ (Lonsdale et al., 2008) nor the SMS-II (Pelletier et al., 2013) evidence was fully supportive of the simplex structure related to correlations between autonomous and controlled motivation scores. The clearest problems with both measures were associated with the identified and integrated regulation scores and their relationships with controlled motivation scores.

*Relations among controlled motivation scores*

Pelletier et al. (2013) criticized the BRSQ, noting that "there was a lack of discrimination between external and introjected regulation scores in terms of their relationships with amotivation" (p. 331). We agreed with this sentiment, writing that "there was no difference between external and introjected regulations scores in terms of their relationships with amotivation" (Lonsdale et al., 2008, p. 343). However, re-analysis of these data using Steiger's (1980) method indicates that the correlation between BRSQ amotivation and external regulation scores ($\Phi = .81$) is, in fact, significantly stronger ($p < .05$) than the relationship between amotivation and introjected regulation scores ($\Phi = .76$), as predicted by theory. That said, this difference was noticeably smaller than that observed in the Pelletier et al.'s (2013) SMS-II data, where the amotivation subscale showed significantly different ($p < .05$) correlations with external ($\Phi = .38$) and introjected ($\Phi = .16$) scores (see Table 3).

*Overall support for simplex pattern of subscale scores*

Overall, the SMS-II and BRSQ both have six subscales, meaning that each correlation matrix has 15 values. When evaluating the simplex structure this means that 20 pairs of correlations need to be examined. Three pairs of BRSQ correlations showed statistically equivalent relationships ($p > .05$); none were mis-ordered. When examining the SMS-II correlations, four showed similar relationships ($p > .05$), and one pair was mis-ordered (i.e., integrated-introjected regulation correlation was stronger than identified–introjected correlation, $p < .05$).

**Table 1**
Correlations ($\Phi$) among autonomous motivation factor scores.

| | Intrinsic motivation | | Integrated regulation | |
|---|---|---|---|---|
| | BRSQ | SMS-II | BRSQ | SMS-II |
| Integrated Regulation | .49 | .78 | – | – |
| Identified Regulation | .42 | .68 | .65 | .70 |

Note: All $\Phi$ values were significantly different from zero ($p < .01$).

**Table 2**
Correlations ($\Phi$) between autonomous and controlled motivation factor scores.

| | Intrinsic motivation | | Integrated regulation | | Identified regulation | |
|---|---|---|---|---|---|---|
| | BRSQ | SMS-II | BRSQ | SMS-II | BRSQ | SMS-II |
| Introjected regulation | −.47 | .33 | .00[a] | .62 | −.03[a] | .45 |
| External regulation | −.54 | .06 | −.16 | .17 | −.16 | .20 |
| Amotivation | −.64 | −.14 | −.26 | −.12 | −.25 | −.08 |

Note.
[a] Indicates $\Phi$ values were not statistically different from zero; all other $\Phi$ values were significant ($p < .01$).

**Table 3**
Correlations ($\Phi$) among controlled motivation factor scores.

| | Introjected regulation | | External regulation | |
|---|---|---|---|---|
| | BRSQ | SMS-II | BRSQ | SMS-II |
| External Regulation | .85 | .50 | – | – |
| Amotivation | .76 | .16 | .81 | .38 |

Note: All $\Phi$ values were significantly different from zero ($p < .01$).

*Relations with theoretically related constructs*

In our scale development studies (Lonsdale et al., 2008), we chose to examine correlations between motivation and two theoretically-related outcomes, flow and athlete burnout. In contrast, Pelletier et al. (2013) examined the relation between SMS-II subscale scores with constructs including perceptions of coach behaviors, task and ego achievement goals, and indicators of well-being (i.e., life satisfaction and vitality). Pelletier et al. (2013, p. 331) suggested that, "there was a lack of discrimination between the [BRSQ] self-determined subscales (intrinsic motivation and identified and integrated regulation) and the concepts of flow and burnout". They also suggested that their correlation analyses involving SMS-II data "provide support for [their] hypotheses regarding associations, on one hand, between athletes' levels of types of motivation and coaches interpersonal behaviors, and, on the other hand, between athletes' types of motivation and various sport related outcomes" (p. 337). We disagree with their interpretations of our results. We also disagree with the conclusions they drew from their own findings with respect to correlations between SMS-II subscales and various outcome variables (i.e., global motivation, achievement goal orientations, life satisfaction, and subjective vitality; Study 1; and interpersonal behavior, achievement goal orientations, life satisfaction, and subjective vitality; Study 2). On balance, we suggest that the nomological evidence regarding the correlations between motivation and theoretically-related constructs actually favors the BSRQ over the SMS-II (see Table 4).

As shown in Table 4, correlations associated with BRSQ subscales followed a pattern that supported the hypothesized self-determination continuum. These results supported the nomological validity of the BRSQ scores. The only exception to this statement was the finding that intrinsic motivation and integrated regulation produced the same correlation with flow (both $r = .36$). None of the other nine pairs of adjacent correlations between BRSQ subscale scores and outcome scores were out of the hypothesized order and all 9 pairs of adjacent correlations were significantly different from the adjacent correlation. In contrast, three of the SMS-II correlations were significantly different, but did not conform to the expected pattern. For example, life satisfaction had a stronger relationship with identified regulation than it did with integrated regulation. A further 15 pairs of adjacent correlations were not significantly different from each other. These findings are contrary to SDT-based hypotheses. Overall, 90% (9 out of 10) pairs of correlations supported the nomological validity of the BRSQ scores. Meanwhile, 51.43% (18 out of 35) pairs of correlations supported the nomological validity of the SMS-II scores.

Pelletier et al. (2013) labeled all the constructs in Table 4 as "outcome variables"; however, we argue that a number of these variables should be considered antecedent variables (i.e., interpersonal behavior, global motivation) or correlates (i.e., achievement goal orientations), but not 'outcome' variables. Only "life satisfaction" and "subjective vitality" could be considered as true outcome variables and both represent global level variables that are likely influenced by motivation from a variety of contexts, not just sport. This issue is important because the nature of the theoretically-related constructs is a crucial consideration when examining nomological validity. Pelletier et al. (2013) did not offer a rationale for their choice of theoretically-related constructs. Also, many of the measures they chose to tap these constructs do not have a great deal of supportive construct validity evidence in the sport context. In our BRSQ development studies (see Lonsdale et al., 2008) we chose a theoretically "adaptive" outcome variable (i.e., flow) and a theoretically "maladaptive" outcome variable (i.e., athlete burnout) to examine the relations between types of motivation (as measured by the BRSQ) and diverse sport-related outcome variables. Also, both of the instruments we chose to measure these constructs had a great deal of supportive evidence in sport-based studies (e.g., Jackson & Eklund, 2002; Raedeke & Smith, 2001).

Finally, the range of correlations associated with the BRSQ was larger across the continuum of motivation types (mean range = .91) than the range associated with SMS-II (mean range = .35). This finding does not speak directly to the nomological validity of subscale scores, but it does suggest that the BRSQ appears to measure motivational constructs that are associated with more divergently adaptive and maladaptive outcomes, when compared with the SMS-II. This finding is in line with SDT tenets that autonomous motives should be more adaptive than controlled motives (Deci & Ryan, 2000).

## Discussion

We congratulate Pelletier et al. (2013) for being proactive in seeking to improve the SMS by developing the SMS-II as an alternative measure. As we stated in our BRSQ development article (Lonsdale et al., 2008, p. 349), "scale development is an on-going process and … we urge researchers to continue the process of psychometric evaluation … and suggest revisions as necessary". Unfortunately, Pelletier et al. did not collect BRSQ data so no psychometric evaluation of the BRSQ scores vis-à-vis the SMS-II scores was possible; nevertheless, we welcome the opportunity to enter into a constructive debate regarding the measurement of SDT constructs in the sport setting.

*Is the SMS-II better than the BRSQ?*

Pelletier et al. (2013) implicitly posed this question in the conclusion of their article and asserted that the SMS-II is a "better" measure of SDT constructs in sport. While we agree that the SMS-II is a substantial improvement over the SMS, we are not convinced by the findings presented by Pelletier et al. that the SMS-II is superior to the BRSQ. There are a number of issues that we believe prompt caution in accepting the SMS-II as a better measure: (i) while both measures produced scores that were internally consistent, factorial validity evidence indicated that the BRSQ met accepted criteria for model fit, but evidence for the SMS-II was mixed; (ii) neither measure produced scores that conformed perfectly with simplex structure; however, the BRSQ appeared to discriminate better among the autonomous motivation subscales scores (see Table 1), while the SMS-II's controlled motivation scores were more clearly differentiated (see Table 3); and (iii) evidence regarding correlations between motivation scores and theoretically-related variables supported the BRSQ over the SMS-II (see Table 4); (iv) all conclusions must be tempered because direct comparisons between the SMS-II and the BRSQ scores are not possible since Pelletier et al. did not collect BRSQ data and in the case of relations with theoretically-related constructs, differing variables were measured in the Pelletier et al. and Lonsdale et al. (2008) studies. Consequently, comparative commentary can only be in narrative form.

**Table 4**
Pearson correlations between BRSQ/SMS-II and theoretically related construct scores.

| | BRSQ | | SMS-II | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Flow | Burnout | Coach Autonomy | Coach Incompetence | Coach Care | Ego Goals | Task Goals | Life Satisfaction | Vitality |
| Intrinsic Motivation | **.36**** | −.50** | .24** | .06 | .21* | .06 | .54** | **.12** | .34** |
| Integrated Regulation | **.36**** | −.23** | **.17*** | −.14 | **.18*** | **.09** | **.45**** | **.11** | **.29**** |
| Identified Regulation. | .21** | −.11* | **.21*** | −.06 | **.15** | **.12** | **.42**** | .20* | **.25**** |
| Introjected Regulation | −.16** | .43** | **.18*** | **.13** | **.08** | **.09** | .29** | −.02 | **.17**** |
| External Regulation | −.25** | .52** | **.13** | **.19*** | **.11** | .26** | −.02 | −.12 | **−.19*** |
| Amotivation | −.31** | .65** | .01 | **.17*** | −.03 | .13* | −.20** | −.30** | −.33** |
| Correlation Range \|r\| | .67 | 1.15 | .23 | .11 | .24 | .07 | .74 | .42 | .67 |

Note: * = $p < .05$, ** = $p < .01$.
Differences in correlations were tested using Steiger's (1980) method. Bolded correlations are not statistically different from correlations in the row above or below ($p < .05$) and, thus, do not support the nomological validity of the BRSQ or SMS-II scores. Underlined correlations are statistically different from the correlation in the row above or below ($p < .05$) and do not conform to the hypothesized pattern. Correlation range = $r$ associated with intrinsic motivation − $r$ associated with amotivation.

*Potential concerns with the SMS-II and the BRSQ*

The SMS-II shares the same stem as the SMS: "Why do you practice your sport?" which could be interpreted as referring only to reasons to be involved in training/practice rather than motives for participation in sport as a whole (including training/practice and competition). As outlined in Lonsdale et al. (2008), we prefer a stem that focuses on sport as a whole in order to avoid any potential ambiguity for the athlete (i.e., "I participate in my sport…."). Both the SMS-II and the BRSQ purport to measure contextual motivation, we believe the BRSQ stem better captures the "context" of sport participation in its totality.

Another potential concern relates to how relevant the items are for the intended constructs (see Table 5 for the list of items in both scales). As presented in our research studies (Lonsdale et al., 2008), the relevance of BRSQ items were rated by a panel of experts in SDT research. Experts rated which constructs the items best represented, and only those receiving unanimous agreement were retained. In contrast, items from the SMS-II were "formulated by experts" (Pelletier et al., 2013, p. 332), but the relevance of the items were not assessed by other researchers in the field. We acknowledge that this lack of assessment is not an indicator of lack of content relevance. Indeed, we agree that most items of the SMS-II were adequate indicators of the constructs they intended to measure. We do, however, feel that there is room for improvement for several items in both the SMS-II and BRSQ. For example, the item "because it teaches me self-discipline" in the BRSQ only measures one of many types of perceived benefits of sports, hence may not be relevant to the much broader construct of identified regulation. As for SMS-II items, some of the intrinsic motivation items may measure intrinsic goal content (i.e., what type of goal an athlete is striving for) rather than intrinsic motivation (Loprinzi et al., 2012). Although intrinsic goal content and intrinsic motivation are similar constructs, they predict psychological well-being independently (Sheldon, Ryan, Deci, & Kasser, 2004).

As for other subscales of the SMS-II, we argue that the item "because I feel better about myself when I do" in the introjected regulation subscale may not necessarily reflect contingent self-worth, but could also represent more autonomous forms of behavioral regulations. Also, the amotivation items "I used to have good reasons for doing sports, but now I am asking myself if I should continue" and "it is not clear to me anymore; I don't really think my place is in sport" measure a *reduction* in motivation over time, rather than *a lack* of motivation. A similar comment was made by Pelletier, Vallerand, and Sarrazin (2007) regarding items from the SMS-6 (Mallett et al., 2007). They argued that a similar item ("I don't seem to be enjoying my sport as much as I previously did") "reflects a decrease in intrinsic motivation and not necessarily the absence of motivation" (p. 618). All in all, we feel that some items from both scales could be improved, and the brief discussion above may serve as a guideline for future development or modification of existing items in the two scales.

We also feel that the representativeness of items in each subscale could be an important aspect of evaluating the merits of measurement scales. Specifically, for both scales, items were discarded from the initial pool during the process of examining the validity of scale scores. For the BRSQ, item elimination was based on statistical considerations (i.e., low factor loadings and potential cross loadings). In contrast, for the SMS-II, items were discarded based on "careful examination of the theory, critical evaluation of the item content, elimination of similar items, and analysis of the initial factor loading" (Pelletier et al., 2013, p. 333). The actual criteria for eliminating or retaining items were unclear. We question the decision to arbitrarily reduce the four item subscales to three item subscales for the SMS-II. Essentially, the elimination of items led to reduced numbers of indicators for each construct (four and three items for the BRSQ and SMS-II, respectively), and may lead to reduced construct coverage regardless of what strategies were chosen to eliminate items (see DeVellis, 2003; Dunn, Bouffard, & Rogers, 1999; Messick, 1980 for detailed discussion). For example, we feel that the items tapping identified regulation in the SMS-II exclusively measured the importance of developing oneself. Athletes can benefit from sport in other ways (e.g., social relations, health), and these may not be tapped using the SMS-II items. In contrast, Pelletier et al. (2013) made an attempt to tap introjected regulation using both approach and avoidance measures of the construct (Roth, Assor, Kanat-Maymon, & Kaplan, 2007), while items in the BRSQ-introjected regulation subscale include mostly avoidance items (75% of items). Unfortunately, neither study conducted an assessment of whether the final sets of items provided sufficient coverage of the constructs being measured (e.g., Ng, Lonsdale, & Hodge, 2011). Therefore, construct representativeness cannot be compared.

*Implications and directions for future research*

Future scale development studies that seek to compare a new measure with an existing measure should collect data with both measures from the same sample. We do not agree that a comparison between the original SMS with the SMS-II was required, given Pelletier et al. (2013) have already stated that the original SMS was less than adequate in terms of content validity. In our previous research studies (Lonsdale et al., 2008), we compared scores derived using the BRSQ, SMS, and the SMS-6. We feel this is the more suitable approach for making direct comparisons. Researchers may also want to revisit important aspects such as content relevance and

**Table 5**
Sport motivation Scale-II and behavioral regulation in sport questionnaire items.

| Sport motivation Scale-II | Behavioral regulation in sport questionnaire |
|---|---|
| Stem: I practice my sport … | Stem: I participate in my sport … |
| Anchors: 1 = *does not correspond at all*, 4 = *corresponds moderately*, 7 = *corresponds exactly* | Anchors: 1 = *Not at all true*, 4 = *Somewhat true*, 7 = *Very True* |
| *Intrinsic Motivation* | |
| Because it gives me pleasure to learn more about my sport. | Because I enjoy it. |
| Because I find it enjoyable to discover new performance strategies. | Because I like it. |
| Because it is very interesting to learn how I can improve. | Because it's fun. |
| | Because I find it pleasurable. |
| *Integrated Regulation* | |
| Because practicing sports reflects the essence of whom I am. | Because it's a part of who I am. |
| Because participating in sport is an integral part of my life. | Because its an opportunity to just be who I am. |
| Because through sport, I am living in line with my deepest principles. | Because what I do in sport is an expression of who I am. |
| | Because it allows me to live in a way that is true to my values. |
| *Identified Regulation* | |
| Because I have chosen this sport as a way to develop myself. | Because the benefits of sport are important to me. |
| Because I found it is a good way to develop aspects of myself that I value. | Because it teaches me self-discipline |
| Because it is one of the best ways I have chosen to develop other aspects of myself. | Because I value the benefits of my sport. |
| | Because it is a good way to learn things which could be useful to me in my life. |
| *Introjected Regulation* | |
| Because I would feel bad about myself if I did not take the time to do it. | Because I would feel ashamed if I quit. |
| Because I feel better about myself when I do. | Because I would feel like a failure if I quit. |
| Because I would not feel worthwhile if I did not. | Because I feel obligated to continue. |
| | Because I would feel guilty if I quit. |
| *External Regulation* | |
| Because people I care about would be upset with me if I didn't. | Because if I don't other people will not be pleased with me. |
| Because I think others would disapprove of me if I did not. | Because I feel pressure from other people to play. |
| Because people around me reward me when I do. | Because people push me to play. |
| | In order to satisfy people who want me to play. |
| *Amotivation* | |
| I used to have good reasons for doing sports, but now I am asking myself if should continue. | But I wonder what's the point. |
| So that others will praise me for what I do. | But I question why I continue. |
| It is not clear to me anymore; I don't really think my place is in sport. | But the reasons why are not clear to me anymore. |
| | But I question why I am putting myself through this. |

construct coverage of items in the future. As discussed in sections above, the relevance and representativeness of items are crucial aspects of construct validity, and systematic assessments are required. Results of these assessments would provide additional evidence for the construct validity of scale scores, or could provide a direction to improve these aspects of the existing scales. For example, one issue that could be considered is the use of both approach and avoidance items when measuring different types of behavioral regulations (i.e., not just introjected regulation).

From a methodological standpoint, collecting data from the same sample using both measures is critical for making direct statistical comparisons of scores derived from the two measures. Apart from direct comparisons of model fit using traditional CFA, models that allow for small cross-loadings, such as ones adopting exploratory structural equation modeling (Asparouhov & Muthén, 2009) or Bayesian structural equation modeling (Muthén & Asparouhov, 2012) frameworks, may be used to address the strong factor correlations found in some studies. Other statistical methods, such as multitrait-multimethod approaches may be employed to further determine the relative merits of the two measures in terms of other aspects of construct validity (e.g., concurrent, discriminant validity).

## Conclusion

It was not our intention in this paper to dismiss the newly created Sport Motivation Scale II (Pelletier et al., 2013). Instead, our purpose was to examine Pelletier et al.'s (2013) claim that the SMS-II is superior to the BRSQ. We believe the evidence highlights the considerable similarities between the two measures in terms of construct validity evidence derived from both scales, including some evidence that was not entirely consistent with theory. Given these similarities, we question whether the SMS-II is superior to the BRSQ or that it addresses the limitations evident in the BRSQ regarding simplex structure. In our opinion, further work is required for both measures. In the meantime, we suggest that authors carefully examine the reliability and validity of scores derived from the measure they choose for their own studies.

## References


Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*(3), 397–438. http://dx.doi.org/10.1080/10705510903008204.

Clark-Carter, D. (2005). *Catalogue of Parametric Tests*. In *Encyclopedia of statistics in behavioral science*.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J: Lawrence Erlbaum.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. http://dx.doi.org/10.1207/S15327965PLI1104_01.

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage.

Dunn, J. G. H., Bouffard, M., & Rogers, W. T. (1999). Assessing item content-relevance in sport psychology scale-construction research: issues and recommendations. *Measurement in Physical Education and Exercise Science, 3*(1), 15–36. http://dx.doi.org/10.1207/s15327841mpee0301_2.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.

Jackson, S. A., & Eklund, R. C. (2002). Assessing flow in physical activity: the Flow State Scale-2 and Dispositional Flow Scale-2. *Journal of Sport & Exercise Psychology, 24*(2), 133–150.

Lonsdale, C., Hodge, K., & Rose, E. A. (2008). The Behavioral Regulation in Sport Questionnaire: instrument development and initial validity evidence. *Journal of Sport & Exercise Psychology, 30*(3), 323–355.

Loprinzi, P. D., Lee, H., Cardinal, B. J., Crespo, C. J., Andersen, R. E., & Smit, E. (2012). The relationship of actigraph accelerometer cut-points for estimating physical activity with selected health outcomes: results from NHANES 2003–06. *Research Quarterly for Exercise and Sport, 83*(3), 422–430. http://dx.doi.org/10.1080/02701367.2012.10599877.

Mallett, C., Kawabata, M., Newcombe, P., Otero-Forero, A., & Jackson, S. (2007). Sport motivation scale-6 (SMS-6): a revised six-factor sport motivation scale. *Psychology of Sport and Exercise, 8*(5), 600–614. http://dx.doi.org/10.1016/j.psychsport.2006.12.005.

Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist, 35*(11), 1012–1027. http://dx.doi.org/10.1037/0003-066X.35.11.1012.

Muthén, B., & Asparouhov, T. (2012). Bayesian SEM: a more flexible representation of substantive theory. *Psychological Methods, 17*(3), 313–335. http://dx.doi.org/10.1037/a0026802.

Ng, J. Y. Y., Lonsdale, C., & Hodge, K. (2011). The Basic Needs Satisfaction in Sport Scale (BNSSS): instrument development and initial validity evidence. *Psychology of Sport and Exercise, 12*(3), 257–264. http://dx.doi.org/10.1016/j.psychsport.2010.10.006.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., Tuson, K. M., Brière, N. M., & Blais, M. R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: the Sport Motivation Scale (SMS). *Journal of Sport & Exercise Psychology, 17*(1), 35–53.

Pelletier, L. G., Rocchi, M. A., Vallerand, R. J., Deci, E. L., & Ryan, R. M. (2013). Validation of the revised sport motivation scale (SMS-II). *Psychology of Sport and Exercise, 14*(3), 329–341. http://dx.doi.org/10.1016/j.psychsport.2012.12.002.

Pelletier, L. G., Vallerand, R. J., & Sarrazin, P. (2007). The revised six-factor Sport Motivation Scale (Mallett, Kawabata, Newcombe, Otero-Forero, & Jackson, 2007): something old, something new, and something borrowed. *Psychology of Sport and Exercise, 8*(5), 615–621. http://dx.doi.org/10.1016/j.psychsport.2007.03.006.

Raedeke, T. D., & Smith, A. L. (2001). Development and preliminary validation of an athlete burnout measure. *Journal of Sport & Exercise Psychology, 23*(4), 281–306.

Roth, G., Assor, A., Kanat-Maymon, Y., & Kaplan, H. (2007). Autonomous motivation for teaching: how self-determined teaching may lead to self-determined learning. *Journal of Educational Psychology, 99*(4), 761–774. http://dx.doi.org/10.1037/0022-0663.99.4.761.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. DOI: 10.1037110003-066X.55.1.68.

Sheldon, K. M., Ryan, R. M., Deci, E. L., & Kasser, T. (2004). The independent effects of goal contents and motives on well-being: it's both what you pursue and why you pursue it. *Personality and Social Psychology Bulletin, 30*(4), 475–486. http://dx.doi.org/10.1177/0146167203261883.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*(2), 245.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education, Inc.

Wilson, P. M., Sabiston, C. M., Mack, D. E., & Blanchard, C. M. (2012). On the nature and function of scoring protocols used in exercise motivation research: an empirical study of the behavioral regulation in exercise questionnaire. *Psychology of Sport and Exercise, 13*(5), 614–622.