

Development and validation of a structured observation scale to measure physical activity instructors' need-supportive and -thwarting strategies

YONGJU HWANG, PH.D¹, TAEWOO KIM, PH.D², YANGYANG DENG, PH.D³, JUNHYUK PARK, MSC⁴, SANGA YUN, MSC⁵, SAMI YLI-PIIPARI, PHD⁶,

¹ Department of Kinesiology University of Alabama Birmingham UNITED STATES

² Brevard College College of Exercise Science UNITED STATES

^{3, 4, 5, 6} Department of Kinesiology University of Georgia UNITED STATES

Published online: August 31, 2025

Accepted for publication: August 15, 2025

DOI:10.7752/jpes.2025.08177

Abstract

Problem Statement: A key challenge for every physical education teacher is motivating students to participate. Instruction is an essential motivational tool within a teacher's pedagogical toolbox. From this instructional standpoint, a central question is which strategies most effectively support participant motivation. Given the practical need to accurately evaluate the impact of these strategies and the growing number of experimental studies on psychological need-supportive instruction, reliable and valid scales to measure need-support are essential. **Purpose:** This study aimed to develop a need-supportive and -thwarting instruction observation scale (NSTI) and perform psychometric testing of the scale. **Approach:** A sample of 20 physical education classes and 540 observations consisting of 110 college-aged students were recruited. **Results:** First, the content validity was established by using the Delphi method followed by the steps to establish factorial and construct validity. Between-rater correlation (.80 to .95) suggested strong internal consistency. Factorial validity was excellent, factor loading ranging between .88 and .99, with 92% (CI 95% [.89, .95]) of the total variance explained. Several factor solutions were tested. Confirmatory factor analysis supported the one-factor solution ($\chi^2(27) = 35.45$; CFI = .99; TLI = .99; RMSEA = .07). **Conclusions:** In conclusion, NSTI had a desired validity and reliability to measure instructor's need-supportive and -thwarting behaviors in physical activity. This study fills an important gap in Self-Determination Theory-centered literature and represents an advancement of the literature focusing on how to measure need-supportive instruction in physical activity validly and reliably.

Keywords: Self-determination, motivation, validity, reliability

Introduction

Motivation is one of the main factors driving humans' volitional behaviors (Ryan & Deci, 2017). This is especially true in the physical activity contexts, e.g. physical education, exercise, or sport, as instructors' and learners' motivation do not always meet (Yli-Piipari, 2017). From the motivational perspective, one of the key questions pertaining to effective instruction is which instructional strategies are best in supporting participant motivation. More effectively instructor can support participant motivation the better learning outcomes will be. Self-Determination Theory (SDT) is a prominent theoretical framework to understand the quality of participant motivation (Deci & Ryan, 1985, 2000; Ryan & Deci, 2017). Based to the SDT, need-supportive instruction is instruction that supports learners' three basic psychological needs of competence, autonomy, and relatedness, whereas need-thwarting instruction violates those needs (Manninen et al., 2022; Manninen & Yli-Piipari, 2021; Tessier et al., 2010). Need-supportive instruction improves people's optimal development and well-being by nurturing their self-determined motivation (Ryan & Deci, 2017). Although the relationships between need-supportive instruction and positive outcomes have been evidenced in the experimental studies (Manninen et al., 2022), objective measures to assess need-supportive or -thwarting instruction are not fully developed. Thus, the aim of this study was to develop and validate a structured observation scale to measure participants' need-supportive and -thwarting strategies in the physical activity context.

According to the SDT, three basic psychological needs are essential for human development and well-being, and those needs impact the quality of motivation (Ryan & Deci, 2002). SDT identifies six different types of motivational regulations based on one's perception of the locus of control, and these motivational regulations are hypothesized to range along with a continuum of self-determination (Deci & Ryan, 1985, 2000; Ryan & Deci, 2017). Intrinsic motivation (i.e., behaviors that are performed within an individual's pleasure and inherent satisfaction) is located at the end of the continuum. Following the continuum of self-determination from fully internalized motive to external motive, extrinsic forms of regulations, integrated regulation (i.e., behaviors that are performed within an individual's goals and values), identified regulation (i.e., behaviors that are performed as valuable and important to an individual), introjected regulation (i.e., behaviors that are performed to achieve others' approval or avoid internal pressures and feeling guilty), and external regulation (i.e., behaviors that are

performed to avoid negative consequences or achieve reward) are located at the middle of the continuum and are regarded extrinsic in nature. Finally, amotivation (i.e., a total absence of motivation) is located at the other far end of the continuum. SDT hypothesizes that intrinsic motivation, integrated motivation, and identified regulation are related to positive affective, cognitive, and behavioral outcomes, while introjected, extrinsic, and amotivation are related to negative outcomes (Deci & Ryan, 1985; Vallerand, 1997).

There have been several studies that have validated and used SDT-centered observational instruments to measure instructional behaviors (Chen, Reeve, & Moon, 2012; Reeve, 2004; Sarrazin et al., 2006; Smith et al., 2015; Tessier, Sarrazin, & Ntoumanis, 2008; 2010). Cheon et al.'s (2012) questionnaire focused on the autonomy-supportive teaching practices, including four autonomy-supportive and controlling behaviors on the 1–7 bipolar response scale. This scale evolved from the Reeve and colleagues' (2004) study, in which the most suitable instructional behaviors were selected from the existing SDT centered literature (Deci, 1995; Deci et al., 1991, 1994; Reeve, 1996, 1998; Ryan & La Guardia, 1999). It is noteworthy, that this observation scale did not include two other dimensions of the need-supportive instruction, namely competence- or relatedness-support. Haerens et al.' (2013) validation work, on the other hand, included all three dimensions of psychological need-support. Their instructional practices were based on the literature and 21 need-supportive observed in physical education teaching behaviors (e.g., Reeve et al., 2004; Reeve & Jang, 2006; Tessier et al., 2008; Vansteenkiste, Zhou, Lens, & Soenens, 2005). During their validation process, raters recorded the 21 need-supportive behaviors using a 4-point frequency scale at every five-minute interval of the class. The scale ranged from 0 (never observed) to 1 (sometimes observed), to 2 (often observed), and to 3 (observed all the time) (Haerens et al., 2013). This scale measures solely need-supportive not need-thwarting instructional practices. Van den Berghe and colleagues' (2013) observation instrument, on the other hand, focused on need-thwarting behaviors. Specifically, the measurement comprised a set of 16 items of need-thwarting behaviors with the specific dimensions of control, coldness, and chaotic instruction (e.g., Bartholomew, Ntoumanis, Ryan, & Bosch, et al., 2011; Belmont, Skinner, Wellborn, & Connell, 1988; Ryan & Deci, 2006).

Similar to the Haerens et al. (2013) observation instrument, the 16 need-thwarting behaviors were coded on a 4-point scale ranging from 0 to 3 for every five minutes of the lesson (Van den Berghe et al., 2013). Both scales Haerens et al. (2013) and Van den Berghe et al. (2013) have shown to display good intra- and inter-rater reliability. However, we argue that both scales are limited in measuring need-supportive and -thwarting instructional behaviors in a real-life instructional environment. First, due to learning environments being dynamic and interactive and instructors displaying different instructional behaviors (sometimes repeating the same instructional behaviors) multiple times in one lesson, sliding scale recording these changes would be a more authentic measurement tool compared to the 5-minute lesson interval strategy of Haerens et al.'s (2013) scale. Second, the scale by Van den Berghe et al. (2013) requires raters to use two different coding systems to simultaneously assess teachers' need-supportive and -thwarting behaviors, which does not represent the authentic nature of the instructional process. Instead, the need-supportive and -thwarting behaviors are occurring in the continuum, as represented in the Cheong et al. (2012) scale. E.g. a need-supportive instructional decision will move the instructional support towards need-support and away from need-thwarting. To our knowledge, there are no observational instrument measuring this authentic nature of teaching and learning. Thus, the aim of this study was to develop and validate need-supportive instruction observation scale to assess instructors' need-supportive and -thwarting behaviors simultaneously in physical activity context and carry out a psychometric testing of the validity and reliability of the scale.

Material & methods

Scale Development

In the Phase, a bank of items was formed completing an extensive review of the central databases (e.g., PUBMED and PsycINFO) to identify both qualitative and quantitative studies that have employed questionnaires focusing on need-supportive and -thwarting instructional strategies (including autonomy-support) (Bartholomew et al., 2011a, 2011b; Manninen et al., 2022; Manninen & Yli-Piipari, 2021; Tessier et al., 2010). In addition, we reviewed the existing questionnaires in the area of SDT (Cheon et al., 2012; Haerens et al., 2013; Van den berg et al., 2013). After reviewing the existing evidence, 21 need-supportive teaching strategy items were selected. To accompany the need-supportive items, we drafted the bipolar need-thwarting items. In Phase 2, a Delphi technique was utilized to seek convergence on a topic from a panel of experts (Helmer, 1967; Hsu & Sandford, 2007). First, three members of the research group conducted a content analysis to determine the degree to which items are representing the need-supportive and -thwarting instructional strategies in the PA context.

After the content analysis, 14 items were selected and submitted for the international experts' review. The inclusion criteria for the panelists was that they needed to have published peer-reviewed journal articles on either SDT and need-supportive instruction or SDT and autonomy-supportive instruction. The authors reviewed the curriculum vitae for each potential panelist to judge their level of expertise based on the inclusion criteria. A convenience sample of eight potential panelists was approached via email and asked to perform content analysis. Three experts agreed to participate, and they performed accuracy analysis rating each item statement (1 = low accuracy ... 5 = high accuracy) (i.e., separately for need-supportive and -thwarting strategies). In addition, they

were asked to give additional commentary on the items, if their accuracy rating was below 4. Based on the experts' feedback, nine maximum scoring items were selected for the final phase. Finally, each item was transferred back to the bipolar format with the need-supportive behaviors on the left side (scored as 3) and the need-thwarting behavior on the left side (scored as -3). Figure 1 presents a description of the specific teacher behaviors.

Need-Supportive Instruction Observation Scale		
Items	Rating	Items
1. Instructor provides tangible goals of the lesson and lays out reasonable expectations relating to participants' individual development and/or growth.	3 2 1 0 -1 -2 -3	1. Instructor provides inadequate goals and unrealistic or extrinsic (e.g., being better in than others) expectations.
2. Instructor provides chances to choose drills at an optimal skill level.	3 2 1 0 -1 -2 -3	2. Instructor provides same tasks for everyone at the same time and rejects any differentiation or pressures participants to participate her/his way.
3. Instructor offers participants a rationale and expected learning outcomes for tasks and exercises.	3 2 1 0 -1 -2 -3	3. Instructor pushes participants to perform tasks without explaining why they should do what the instructor asks them to do. Instructor ignores participants' efforts to understand the tasks.
4. Instructor provides positive and informational feedback that signifies competence or improves competence.	3 2 1 0 -1 -2 -3	4. Instructor makes all decisions, provides simple tasks requiring lower level of cognitive engagement, and no time to work independently. Instructor provides no opportunities for participant ownership.
5. Instructor provides meaningful choice within the skill drills and flexibility with appropriate techniques.	3 2 1 0 -1 -2 -3	5. Instructor forces participants to follow instructor's decision and offers no choice.
6. Instructor provides new guidelines, tips and advice for participants leading to more engagement and higher-level learning.	3 2 1 0 -1 -2 -3	6. Instructor ignores chances to provide guidance, tips, or advice.
7. Instructor encourages participants to persist (e.g., "come on, you can do it").	3 2 1 0 -1 -2 -3	7. Instructor does not put any effort and energy into the lesson.
8. Instructor is close to the participants, moves around, and interacts with participants.	3 2 1 0 -1 -2 -3	8. Instructor does not move around the space and physically distances (her)himself from the participants.
9. Instructor asks participants questions about their interests, problems, values, or wishes addressing them by their first name.	3 2 1 0 -1 -2 -3	9. Instructor detaches from the participants and ignores their opinions or questions (acting cold).

Figure 1. A description of need-supportive and -thwarting instructional teacher behaviors

Participants

A sample consisted of 20 physical education classes, 540 observations, and 110 college-aged students ($M_{age} = 20.56 \pm 2.25$; 62.7% females) segregated to four separate classes ($M_{size} = 27.5$). These cases were part of the local university's basic physical education courses. Students selected individually one out of four classes based on the classes' suitability to the overall study schedule. One teacher with experience in teaching physical education taught 20 (5 lessons for each group) 50 min lessons in two weeks. Nine raters (3/class) were randomly selected to observe and rate the instructional practices. The study protocol was approved by the Institutional Review Board of the local university, and all participants provided a written consent before the study.

Procedures

Before the observation began, the observers participated in a 90-minute training session, which took place four days prior to the start of the study. During this training, four researchers who were involved in the study explained each item on the observation scale to the observers in detail. Following this, a researcher demonstrated a session of juggling instruction, after which each observer independently graded the session using the observation scale. The group then discussed their individual grades together to ensure consistency in their observations. During the validation process, each observer scored all five lessons independently. A male instructor who had ten years of experience teaching public school physical education taught all the lessons. Lessons were conducted in September 2021 during the fall semester (met 3-times/week following the Mon, Wed, Fri schedule). Neither the teacher nor students were using facemasks. The content of the lesson was juggling, and each lesson lasted 45 min. The five 45-min lessons were held during 11 days, and all four groups followed the same content and structure. The lessons commenced with a standardized 10-min warm-up segment followed by a 25-min skill-learning section, and a 10-min closure.

Raters were asked to use the NSTI scale to evaluate the classes. If a teacher observed conducted one need-supportive instructional behavior, raters were advised to move sliding from 0 (default) to 1. Subsequently, if a teacher performed a need-thwarting instructional behavior, raters were advised to slide the rating one step towards the need-thwarting side.

Instrument

To measure need-supportive and -thwarting strategies, the NSTI scale consisting of nine need-supportive and need-thwarting teaching strategies was used. The NSTI scale comprised a 7-point frequency scale, ranging from 3 (need-supportive behavior observed three times), to 2 (need-supportive behavior observed two times), to 1 (need-supportive behavior observed once), to 0 (need-supportive behavior never observed), to -1 (need-thwarting behavior observed once), to -2 (need-thwarting behavior observed two times), to -3 (need-thwarting behavior observed three times) (see Figure 1). The NSTI observational scale consist of three need-supportive and – thwarting teaching dimensions of autonomy support (e.g. ‘Instructor provides chances to choose drills at an optimal skill level’) and autonomy thwarting (e.g. ‘Instructor provides same tasks for everyone at the same time and rejects any differentiation or pressures participants to participate her/his way’), relatedness support (e.g. ‘Instructor is close to the participants, moves around, and interacts with participants’) and relatedness thwarting (e.g. ‘Instructor detaches from the participants and ignores their opinions or questions (acting cold)’), and competence support (e.g. ‘Instructor provides positive and informational feedback that signifies competence or improves competence’) and competence thwarting (e.g. ‘Instructor makes all decisions, provides simple tasks requiring lower level of cognitive engagement, and no time to work independently. Instructor provides no opportunities for participant ownership’).

Statistical analyses

After cleaning the data, intra-class correlation coefficients (ICC) and their 95% confidence intervals based on a mean-rating ($k = 3$), absolute-agreement, 2-way random effects model were examined using the SPSS statistical package version 23 (SPSS Inc, Chicago, IL). When interpreting the findings, values less than .5 are indicative of poor reliability, values between .5 and .75 indicate moderate reliability, values between .76 and .90 indicate good reliability, and values greater than .90 indicate excellent reliability (Koo & Lee, 2016).

Next, an exploratory factor analysis (EFA) was conducted. Bartlett’s test of sphericity (Bartlett, 1950) was used to ensure that the correlation matrix was not random, and the Kaiser-Meyer-Olkin (KMO) statistic (Kaiser, 1974) was required to be above a minimum of .80. After confirming that the correlation matrix was factorable, EFA was conducted. A common factor analysis was selected over a principal component analysis because the intent was to identify a latent factor structure (Fabrigar et al., 1999). An iterated extraction method with initial communalities estimated by squared multiple correlations was employed because of its relative tolerance of non-normality and demonstrated ability to recover weak factors (Briggs & MacCallum, 2003; Guttman, 1956). Following the advice of Velicer et al. (2000), parallel analysis (Horn, 1965), MAP (Velicer, 1976), and the visual scree test (Cattell, 1966) were used to determine the appropriate number of factors to retain. Parsimony and theoretical convergence were also considered. Due to the nature of the constructs, it was assumed that factors would be correlated. Therefore, an oblimin rotation (Jennrich & Sampson, 1966) was employed (Carroll, 1978; Child, 2006). Criteria for determining factor adequacy were established a priori. Pattern coefficients $\geq .37$ were considered salient (i.e., both practically and statistically significant as per Norman & Streiner, 2014). Factors with a minimum of three salient pattern coefficients, internal consistency reliability $\geq .70$, and that were theoretically meaningful were considered adequate. Confirmatory factor analysis (CFA) using Mplus version 22 and 8.4 software. The following goodness-of-fit indices were used: Chi-square ($\chi^2 > .05$), Bentler comparative fit index (CFI; $> .95$) and Tucker-Lewis index (TLI; $> .95$), Root Mean Squared Error of Approximation (RMSEA; $< .06$) (Hu & Bentler, 1999). Finally, a chi-squared difference test ($\Delta\chi^2$) was used to compare the competing hypothesized models (Rigdon, 1999). If the p value associated with the chi-square test is non-significant, H_0 is rejected and an alternative hypothesis H_1 accepted. In other words, a p value $< .001$ for the chi-square difference tests indicates statistically significant differences in the estimated models.

Results

Inter-rater agreement

ICC (2, k) coefficient was used to assess the consistency of item ratings with the mean correlation across all raters and rating dimensions indicating the extent of rating consistency. Between-rater correlation ranged between .89 and .90. Item mean was .28 ranging from minimum .08 to maximum .41. Analyses produced an ICC of .96(6.19), CI 95% (.95, .97) with item ratings ranging from .80 to .95.

Factorial Validity

To explore the factorial validity of the NSTI scale, all nine instructional strategies were subjected to an EFA. KMO verified sampling adequacy for the analysis, KMO = .96. Bartlett’s test of sphericity (36) = 1056.16, $p < .001$ indicated that correlation structure is adequate for factor analyses. Parallel analysis, MAP, and scree all suggested that one-factor structure should be retained, but theory (Deci & Ryan, 1985) suggests that there could be potentially a three-factor structure. Analyses showed that the three-factors solution was not salient as none of the aforementioned analyses did not support the three-factor solution with the factor’s internal consistency

(alpha) reliability being only .17. The one-factor solution was conceptually plausible as instructional strategies oftentimes impact not only say autonomy but relatedness as well. The one-factor solution was next examined for adequacy. Factor loading ranged between .88 and .99. (Table 1), with 92% (CI 95% = .89- .95) of the total variance explained.

Table 1
Factor loadings of the need-supportive and -thwarting teaching behaviors

Items	Factor
Chances to choose drills at an optimal skill level.	.977
Tangible goals of the lesson and lays out reasonable expectations relating to participants' individual development and/or growth	.882
Offers participants a rationale and expected learning outcomes for tasks and exercises.	.957
Positive and informational feedback that signifies competence or improves competence.	.979
Meaningful choice within the skill drills and flexibility with appropriate techniques.	.973
New guidelines, tips and advice for participants lead to more engagement and higher-level learning.	.964
Instructor encourages participants to persist (e.g., "come on, you can do it").	.935
Is close to the participants, moves around, and interacts with participants.	.987
Asks participants questions about their interests, problems, values, or wishes addressing them by their first name.	.948

CFA

CFA was conducted for each of the nine factors extracted through the EFA. The model fit indices for the Default Model showed an acceptable model fit: $\chi^2(27) = 35.45$, $p = .073$, CFI = .99, TLI = .99, RMSEA = .07, CI 90% [.01, .13]. However, the low factor loading for the instructional strategy nine ("*...asks participants questions about their interests, problems, values, or wishes addressing them by their first name*") $\lambda^9 = .490$ ($p = < .001$) was distinctively smaller than other loadings lambda's ranging between .882 and .988 (Table 1). Thus, we tried to omit the ninth instructional strategy to test whether this modified eighth-factor model (Modified Model) improves the model fit. The new Modified Model fit indices showed a weaker model fit: $\chi^2(20) = 31.76$, $p = .046$, CFI = .99, TLI = .98, RMSEA = .100, CI 90% [.01, .12] compared to the Default Model. To test whether the models differed in statistical significance, $\Delta\chi^2$ test was employed. However, the $\Delta\chi^2$ test showed that the Modified Model was not statistically different from the Default Model 1 ($\Delta\chi^2 = 0.013$, $p = .846$). Thus, the final model parameters are presented in Table 2.

Table 2
Final Model factor loadings of the instructional behavior

Items	S.E.	Est./SE	<i>p</i>	<i>R</i> ² (S.E.)	<i>p</i>
Instructor provides tangible goals of the lesson and lays out reasonable expectations relating to participants' individual development and/or growth	.976	.007	.000	.95(.01)	.000
Instructor provides chances to choose drills at an optimal skill level.	.882	.029	.000	.78(.05)	.000
Instructor offers participants a rationale and expected learning outcomes for tasks and exercises.	.956	.012	.000	.92(.01)	.000
Instructor provides positive and informational feedback that signifies competence or improves competence.	.977	.007	.000	.96(.01)	.000
Instructor provides meaningful choice within the skill drills and flexibility with appropriate techniques.	.973	.008	.000	.95(.02)	.000
Instructor provides new guidelines, tips and advice for participants leading to more engagement and higher-level learning.	.962	.010	.000	.94(.02)	.000
Instructor encourages participants to persist (e.g., "come on, you can do it").	.937	.016	.000	.88(.03)	.000
Instructor is close to the participants, moves around, and interacts with participants.	.988	.004	.000	.98(.01)	.000
Instructor asks participants questions about their interests, problems, values, or wishes addressing them by their first name.	.490	.099	.000	.24(.10)	.014

Discussion

The present study aimed to develop and validate the NSTI observation scale for the physical activity context. The psychometric testing supported reliability and validity of the 9-item NSTI scale. Our analyses demonstrated excellent inter-rater reliability consistency with our three raters' scores achieving a 96% agreement. EFA suggested either one or three factor solutions. Parallel analysis, MAP, and scree all suggested that one-factor structure should be retained, but due to the theoretical postulation of the SDT (Deci & Ryan, 1985), also a three-factor structure was tested. Analyses, however, did not support the three-factor solution as internal reliability was weak. It is noteworthy that one-factor solution is plausible as instructional strategies almost always impact simultaneously two or three psychological needs and not only one need. For example, positive and informational feedback nurtures not only participants' competence but can support their autonomy and relatedness as well, and thus are complementary (Tessier et al., 2010). Finally, one-factor model factor loadings were strong, with 92% of the total variance explained by the selected nine factors. The result suggested that nine factors in bipolar format presented instructors' need-supportive and -thwarting instructional behaviors well in the physical activity context.

Next, we conducted a CFA to test the construct validity of this nine-factor solution. Our analyses showed that the factor loadings of the CFA supported the nine-item solution. However, the item nine ("...asks participants questions about their interests, problems, values, or wishes addressing them by their first name") was distinctively smaller than other loadings with lambda being .490. Thus, we tried to omit the ninth instructional strategy to test whether this modified eighth-factor solution would be significantly better than the nine-factor solution. However, the chi-squared difference test showed that the eight-factor solution did not differ from the nine-factor model at the statistically significant level. Thus, we retrieved the nine-factor model as the best model to assess validly need-supportive and -thwarting instructional strategies.

Considering the plethora of experimental studies examining the impact of need-support instruction in physical activity (e.g., Manninen et al., 2022), there are very few published reliable and valid questionnaires to measure SDT-centered concepts, e.g. autonomy-support and, especially, psychological need-support. A need for valid observation scales is imminent as measuring treatment fidelity (e.g., teachers' adherence to intended teaching practices) is a necessary component of rigorously designed and conducted experimental studies. Although the previous scales have demonstrated moderate reliability and validity (Cheon et al., 2012; Haerens et al., 2013) to date there are no validated observation scales that measure need-supportive and -thwarting instructional practices objectively. Most of the time, need-supportive instructional practices have been assessed subjectively by asking participants' perceived experiences of autonomy, competence, and/or relatedness (Abula, 2018; Barkoukis, 2021; Chang, 2016; Edmunds, 2008; Franco, 2017; Gonzalez-Cutre, 2018; Leptokaridou, 2016; Mertens, 2018; Moreno-Murcia, 2016; Moustaka, 2012; Ntoumanis, 2017; Pulido, 2017; S'anchez-Oliva et al., 2017; Sebire, 2016; Tilga, 2019; Tilga, 2020; Ulstad, 2018; Zazo, 2018). Some studies have assessed instructors' need-support analyzing recorded video or audio from the teaching situations, relying on raters' interpretations of teaching guidelines (Fin, 2019; Franco, 2017; Langan, 2015; Lonsdale, 2013, 2019; Moreno-Murcia, 2016). Although Cheon and colleagues (2012) used a bipolar observation scale to assess teachers' autonomy-supportive behavior and controlling behavior on site, this scale targeted only instructors' autonomy-support excluding competence- and relatedness-support. This study responded to the need to have valid SDT-centered observation scale measuring psychological need-supportive and -thwarting instructional practices in organic, real-time physical activity sessions. Our nine-item study showed that the NSTI scale is a valid observation scale to measure all dimensions of need-supportive and -thwarting instructional behaviors.

Limitations and future studies

As a limitation, the sample comprised university physical education students. Future studies should test the validity of the scale in K-12 physical education context, as university physical students are conceptually different from K-12 physical education students. In addition, future studies should further test the reliability of this scale using a different sample. Furthermore, additional comprehensive efforts to determine the validity of the NSTI scale are encouraged.

Conclusions

The main conclusion of this study is that this observation validation fills an important gap in SDT-centered literature and represents an advancement of the scientific literature focusing on how to measure need-supportive instruction in physical activity validly and reliably. The findings of this study suggest desirable validity and reliability of the NSTI scale.

Conflicts of interest None

References

- Abula, K., Beckmann, J., He, Z., Cheong, C., Lu, F., & Gröpel, P. (2020). Autonomy support in physical education promotes autonomous motivation towards leisure-time physical activity: Evidence from a sample of Chinese college students. *Health Promotion International*, 35(1), e1–e10. <https://doi.org/10.1093/heapro/day102>

- Barkoukis, V., Chatzisarantis, N., & Hagger, M. S. (2021). Effects of a School-Based Intervention on Motivation for Out-of-School Physical Activity Participation. *Research Quarterly for Exercise and Sport*, 92(3), 477–491. <https://doi.org/10.1080/02701367.2020.1751029>
- Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., & Thøgersen-Ntoumani, C. (2011a). Psychological need thwarting in the sport context: Assessing the darker sides of athletic experience. *Journal of Sport and Exercise Psychology*, 33(1), 75–102.
- Bartholomew, K. J., Ntoumanis, N., Ryan, R. M., Bosch, J. A., & Thøgersen-Ntoumani, C. (2011b). Self-determination theory and diminished functioning: The role of interpersonal control and psychological need thwarting. *Personality and Social Psychology Bulletin*, 37(11), 1459–1473.
- Bartlett, M. S. (1954). A further note on the multiplying factors for various chi-square approximations in factor analysis. *Journal of the Royal Statistical Society, Series B*, 16, 296–298.
- Belmont, M., Skinner, E., Wellborn, J., & Connell, J. (1988). *Teacher as social context: A measure of student perceptions of teacher provision of involvement, structure, and autonomy support*. Vol. Tech. Rep. No. 102. Rochester, NY: University of Rochester.
- Briggs, N. E., MacCallum, R. C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. *Multivariate Behavioral Research*, 38, 25–56. doi:10.1207/S15327906MBR3801_2
- Carroll, J. B. (1978). How shall we study individual differences in cognitive abilities? Methodological and theoretical perspectives. *Intelligence*, 2, 87–115. doi:10.1016/0160-2896(78)90002-8
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245–276. doi:10.1207/s15327906mbr0102_10
- Chang, Y.-K., Chen, S., Tu, K.-W., & Chi, L.-K. (2016). Effect of autonomy support on self-determined motivation in elementary physical education. *Journal of Sports Science & Medicine*, 15, 460–466.
- Cheon, S. H., Reeve, J., & Moon, I. S. (2012). Experimentally based, longitudinally designed, teacher-focused intervention to help physical education teachers be more autonomy supportive toward their students. *Journal of Sport & Exercise Psychology*, 34(3), 365–396.
- Child, D. (2006). *The essentials of factor analysis*. (3rd ed.). New York, NY: Continuum International Publishing Group.
- Deci, E. L. (1995). *Why we do what we do: The dynamics of personal autonomy*. New York: Penguin Books.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- 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.
- Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, 62, 119–142.
- Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation in education: The self-determination perspective. *Educational Psychologist*, 26, 325–346.
- Edmunds, J., Ntoumanis, N., & Duda, J. L. (2008). Testing a self-determination theory-based teaching style intervention in the exercise domain. *European Journal of Social Psychology*, 38(2), 375–388. <https://doi.org/10.1002/ejsp.463>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272–299. doi:10.1037/1082-989X.4.3.272
- Franco, E., & Coterón, J. (2017). The Effects of a Physical Education Intervention to Support the Satisfaction of Basic Psychological Needs on the Motivation and Intentions to Be Physically Active. *Journal of Human Kinetics*, 59(1), 5–15. <https://doi.org/10.1515/hukin-2017-0143>
- González-Cutre, D., Sierra, A. C., Beltrán-Carrillo, V. J., Peláez-Pérez, M., & Cervelló, E. (2018). A school-based motivational intervention to promote physical activity from a self-determination theory perspective. *The Journal of Educational Research*, 111(3), 320–330. <https://doi.org/10.1080/00220671.2016.1255871>
- Guttman, L. (1956). "Best possible" systematic estimates of communalities. *Psychometrika*, 21, 273–285. doi:10.1007/BF02289137
- Haerens, L., Aelterman, N., Van den Berghe, L., De Meyer, J., Soenens, B., & Vansteenkiste, M. (2013). Observing physical education teachers' need-supportive interactions in classroom settings. *Journal of Sport and Exercise Psychology*, 35(1), 3–17.
- Helmer, O. (1965) *Analysis of the future: the Delphi method*. Santa Monica, CA: Rand Corporation.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185. doi:10.1007/BF02289447
- Hsu CC, Sandford BA (2007) The Delphi technique: making sense of consensus. *Practical Assessment, Research & Evaluation* 12, 1–8.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. doi: 10.1080/10705519909540118

- Jenrich, R. L., Sampson, P. F. (1966). Rotation for simple loading. *Psychometrika*, 31, 313-323. doi:10.1007/BF02289465
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31-36. doi:10.1007/BF02291575
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155-163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Langan, E., Toner, J., Blake, C., & Lonsdale, C. (2015). Testing the Effects of a Self-Determination Theory-Based Intervention with Youth Gaelic Football Coaches on Athlete Motivation and Burnout. *The Sport Psychologist*, 29(4), 293-301. <https://doi.org/10.1123/tsp.2013-0107>
- Leptokaridou, E. T., Vlachopoulos, S. P., & Papaioannou, A. G. (2016). Experimental longitudinal test of the influence of autonomy-supportive teaching on motivation for participation in elementary school physical education. *Educational Psychology*, 36(7), 1138-1159. <https://doi.org/10.1080/01443410.2014.950195>
- Lonsdale, C., Lester, A., Owen, K. B., White, R. L., Peralta, L., Kirwan, M., Diallo, T. M. O., Maeder, A. J., Bennie, A., MacMillan, F., Kolt, G. S., Ntoumanis, N., Gore, J. M., Cerin, E., Cliff, D. P., & Lubans, D. R. (2019). An internet-supported school physical activity intervention in low socioeconomic status communities: Results from the Activity and Motivation in Physical Education (AMPED) cluster randomised controlled trial. *British Journal of Sports Medicine*, 53(6), 341-347. <https://doi.org/10.1136/bjsports-2017-097904>
- Manninen, M., & Yli-Piipari, S. (2021). Ten Practical Strategies to Motivate Students in Physical Education: Psychological Need-Support Approach, *Strategies*, 34:2, 24-30, DOI: 10.1080/08924562.2020.1867270
- Manninen, M., Deng, Y., Hwang, Y., Waller, S., & Yli-Piipari, S. (2022). Psychological need-supportive instruction improves novel skill performance, intrinsic motivation, and enjoyment: A cluster-randomised study. *International Journal of Sport and Exercise Psychology*, 20 (1), 122-146. <https://doi.org/10.1080/1612197X.2020.1826999>
- Norman, G. R., Streiner, D. L. (2014). *Biostatistics: The bare essentials* (4th ed.). Shelton, CT: People's Medical Publishing.
- Ntoumanis, N., Thøgersen-Ntoumani, C., Quested, E., & Hancox, J. (2017). The effects of training group exercise class instructors to adopt a motivationally adaptive communication style. *Scandinavian Journal of Medicine & Science in Sports*, 27(9), 1026-1034. <https://doi.org/10.1111/sms.12713>
- Reeve, J. (1996). *Motivating others: Nurturing inner motivational resources*. Needham Heights, MA: Allyn & Bacon.
- Reeve, J. (1998). Autonomy support as an interpersonal motivating style: Is it teachable? *Contemporary Educational Psychology*, 23, 312-330.
- Reeve, J., & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, 98(1), 209-218. doi:10.1037/0022-0663.98.1.209
- Reeve, J., Jang, H., Carrell, D., Jeon, S., & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, 28(2), 147-169.
- Rigdon, E. E. (1999). Using the Friedman method of ranks for model comparison in structural equation modelling. *Structural Equation Modelling: A Multidisciplinary Journal*, 6(3), 219-232.
- Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3-33). Rochester, NY: University of Rochester.
- Ryan, R. M., & Deci, E. L. (2006). Self-regulation and the problem of human autonomy: does psychology need choice, self-determination, and will? *Journal of Personality*, 74, 1557-1585. <http://dx.doi.org/10.1111/j.1467-6494.2006.00420.x>
- Ryan, R.M., & La Guardia, J. G. (1999). Achievement motivation within a pressured society: Intrinsic and extrinsic motivations to learn and the politics of school reform. In T. Urdan (Ed.), *Advances in motivation and achievement* (Vol. 11, pp. 45-85). Greenwich, CT: JAI Press.
- Ryan, R. M., & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. United States: Guilford Publications.
- Sarrazin, P. G., Tessier, D. P., Pelletier, L. G., Trouilloud, D. O., & Chanal, J. P. (2006). The effects of teachers' expectations about students' motivation on teachers' autonomy-supportive and controlling behaviors. *International Journal of Sport and Exercise Psychology*, 4(3), 283-301. <https://doi.org/10.1080/1612197X.2006.9671799>
- Sheldon, K. M., Garton, B., Orr, R., & Smith, A. (2015). The Advisor Quality Survey: Good College Advisors are Available, Knowledgeable, and Autonomy Supportive. *Journal of College Student Development*, 56(3), 261-273. <https://doi.org/10.1353/csd.2015.0027>
- Tessier, D., Sarrazin, P., & Ntoumanis, N. (2008). The effects of an experimental programme to support students' autonomy on the overt behaviours of physical education teachers. *European Journal of Psychology of Education*, 23(3), 239-253. <https://doi.org/10.1007/BF03172998>

- Tessier, D., Sarrazin, P., & Ntoumanis, N. (2010). The effect of an intervention to improve newly qualified teachers' interpersonal style, students' motivation and psychological need satisfaction in sport-based physical education. *Contemporary Educational Psychology*, 35(4), 242–253. <https://doi.org/10.1016/j.cedpsych.2010.05.005>
- Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In *Advances in Experimental Social Psychology* (Vol. 29, pp. 271-360). New York, NY: Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60019-2](https://doi.org/10.1016/S0065-2601(08)60019-2)
- Van den Berghe, L., Soenens, B., Vansteenkiste, M., Aelterman, N., Cardon, G., Tallir, I. B., & Haerens, L. (2013). Observed need-supportive and need-thwarting teaching behavior in physical education: Do teachers' motivational orientations matter? *Psychology of Sport and Exercise*, 14(5), 650–661. <https://doi.org/10.1016/j.psychsport.2013.04.006>
- Vansteenkiste, M., Zhou, M.M., Lens, W., & Soenens, B. (2005). Experiences of autonomy and control among Chinese learners: vitalizing or immobilizing? *Journal of Educational Psychology*, 97(3), 468–483. doi:10.1037/0022-0663.97.3.468
- Velicer, W. F., Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, 3, 231-251. doi:10.1037/1082-989X.3.2.231
- Yli-Piipari, S. (2017). Motivation as adaptation or maladaptation of personal goals with learning goals. In C. D. Ennis (Ed.), *Routledge Handbook of Physical Education Pedagogies*. Routledge, London. Great Britain, 619-630.