The Need-Relevant Instructor Behaviours Scale (NIBS): Development and initial validation

Quested, E.[1], Ntoumanis, N.[1], Stenling A.[2], Thogersen-Ntoumani, C.[1], & Hancox, J.E.[3]

[1] School of Psychology, Curtin University, Australia

[2] Department of Psychology, Umeå University, Sweden

[3] Division of Rehabilitation and Ageing, Queen's Medical Centre, University of Nottingham, UK

Please address all correspondence to:

Dr Eleanor Quested, School of Psychology, Curtin University, GPO Box U1987, 6845, Perth, Western Australia, Australia, Tel: +61 8 9266 5693, email:

eleanor.quested@curtin.edu.au

Abstract

Purpose. This article outlines the development and validation of the Need-Relevant Instructor Behaviors Scale (NIBS). Drawing from self-determination theory, the NIBS is the first observation tool designed to code the frequency and the intensity of autonomy, competence and relatedness relevant behaviours of exercise instructors. The scale also captures the frequency of need indifferent behaviors. ***Methods.*** The behaviors of 27 exercise instructors were coded by trained raters on two occasions, before and after they received training in adaptive motivational communication. ***Results.*** Findings supported the structural validity and reliability of the scale. The scale's sensitivity to detect changes in frequency and intensity of need-relevant behaviors was also evidenced. ***Conclusions.*** The NIBS is a new tool that offers a unique, tripartite assessment of need relevant behaviors of leaders in the physical activity domain.

Keywords: motivation, observation, self-determination theory, need support, need thwarting, exercise

Despite the prevalence and variety of fitness activities available to individuals in western societies, the vast majority of adults remain insufficiently physically active (Kohl et al., 2012), and as a result, many suffer physical and psychological ill-health consequences (Hamer & Chida, 2009). While many adults initiate a fitness regime more than once in their life, few sustain this behavior consistently or for the long term; about 50% drop out within the first six months (Marcus et al., 2006). Numerous studies have highlighted the important role of the motivational environment created by the exercise instructor in determining whether an exerciser sustains or drops out from regular exercise (for reviews, see Ntoumanis, Quested, Reeve, & Cheon, 2018; Teixeira, Carraça, Markland, Silva, & Ryan, 2012). Yet, very few studies have attempted to employ independent raters to assess the motivationally relevant characteristics of such an environment. The development of such a methodological approach is critical if advances are to be made in the evaluation of training programmes designed to help exercise professionals employ more motivationally adaptive communication styles. The aim of this study was twofold: 1) to develop an objective measure of the motivational environment created by exercise instructors, and subsequently, 2) to test the sensitivity of this tool to change by using it to code instructor behavior before and after exercise instructors received training in adaptive motivational communication (see Hancox, Quested, Thøgersen-Ntoumani, & Ntoumanis, 2015; Ntoumanis, Thøgersen-Ntoumani, Quested, & Hancox, 2017)

Theoretical Underpinnings

Research aiming to identify the factors that differentiate between individuals who maintain regular physical activity and those who drop out has highlighted quality of motivation as a distinguishing factor (Edmunds, Ntoumanis, & Duda, 2007). Much of this research pulls from self-determination theory (Ryan & Deci, 2017). This theory, and the associated body of research evidence, posits that exercisers who sustain long term engagement hold autonomous motives for participation in exercise (Hancox, Ntoumanis,

Thogersen-Ntoumani, & Quested, 2015; Teixeira et al., 2012). With such motives, exercisers'

engagement in the activity is underpinned by factors such as enjoyment and personal

satisfaction (i.e., intrinsic regulation), and/or the valuing of the benefits of the activity (i.e.,

identified regulation). On the contrary, those who drop out or experience regular (re)lapses in

exercise engagement tend to be regulated by more controlled forms of motivation, such as

internal contingencies or pressures (i.e., introjected regulation), or external drivers such as the

demands of another person, avoidance of punishment or seeking of rewards (i.e., external

regulation). A substantial body of research has supported the SDT-based premise that the

communication style of the exercise instructor accounts for differences in exercisers' quality

of motivation (autonomous vs controlled) to exercise (for a review see Teixeira et al., 2012).

SDT posits that the degree to which social environments nurture versus deprive

satisfaction of the three basic psychological needs will determine whether individuals within

these environments will develop autonomous or controlled motivation (Ryan & Deci, 2017).

For example, when an exercise instructor behaves in a way that is supportive of the

exercisers' need to experience autonomy (i.e., feeling volitional and self-directed, behaving in

accordance with their values), competence (i.e., feeling capable to meet challenges) and

relatedness (i.e., feeling connected, respected, cared for), exercisers will experience higher

quality (i.e., self-determined) and more sustained motivation ensues. However, when the

social environment is void of need supportive features or includes characteristics that

undermine/thwart the needs, then exercisers will experience need frustration a lower quality

and less sustainable (i.e., more controlled) motivation. Research in the exercise domain has

shown that the satisfaction of the three needs provides the nutriments for higher quality

motivation and adoption of exercise behaviors (Standage & Ryan, 2012). For example,

Wilson, Mack, Muon, and LeBlanc (2007) found that exercise adherence to a 12-week

programme was predicted by moderate increases in competence and relatedness need

satisfaction, as well as autonomous motivation. In a longitudinal investigation of exercise adherence, Duda et al. (2014) showed that perceptions of need support provided by health and fitness advisors at the end of a 3-month exercise programme were positively linked to exercisers' psychological need satisfaction at 3 months. The latter variable positively predicted physical activity at 6 months (i.e., 3 months after end of the programme) via intentions for physical activity at 3 months.

The evidence linking need supportive instructing styles with adaptive outcomes for exercisers is substantial in volume (Standage & Ryan, 2012), yet predominantly reliant on self-report data and correlational analyses. Few intervention studies have attempted to manipulate the communication style used by exercise instructors to create a more need supportive atmosphere; such studies have predominantly taken place in clinical settings (e.g., Mildestvedt, Meland, & Eide, 2008; Rahman, Hudson, Thøgersen-Ntoumani, & Doust, 2015). Very few studies have targeted behaviors of instructors working in commercial fitness settings designed for healthy adults. Two exceptions were the investigations by Edmunds and colleagues (Edmunds, Ntoumanis, & Duda, 2008) and Fortier and colleagues (Fortier, Sweet, O'Sullivan, & Williams, 2007). However both of the aforementioned trials utilised the same instructor for both the intervention (i.e., an SDT instructing style) and control (i.e., a 'typical' instructing style) conditions, and relied upon exercisers' perceptions of instructor behavior to gauge change in provision of need support. As a result, there is limited evidence to explain if or how SDT-based training for exercise instructors can be effective in changing objective instructor behaviors to create more need supportive environments.

Assessing Frequency and Intensity of Leader Behaviors in SDT-Based Observational Tools

Observation is a methodological approach that involves having trained observers follow a specified protocol to record observed dialogue or behavior (Darst, Zakrajsek, & Mancini, 1989). In the sport and education context, there have been several attempts to

employ SDT-based observation methodologies to assess motivationally relevant behaviors of the coach or teacher (for a recent review, see Smith, Quested, Appleton, & Duda, 2016). However there has been a relative dearth of such approaches in the exercise domain.

SDT-based observation scales have been designed to assess the frequency of particular behaviors (e.g., Reeve, Jang, Carrell, Jeon, & Barch, 2004; Webster et al., 2013), and such frequency scores have been used to assess the effectiveness of training interventions. For example, Haerens and colleagues (Haerens et al., 2013; Van den Berghe et al., 2013) developed an observation scale to rate the degree to which motivational behaviors derived from SDT are used on a continuum ranging from 0 ('*not at all*') to 3 ('*all of the time*'). An alternative approach developed by Smith and colleagues (2015) is to gauge the 'potency' (or intensity) of motivation-related behaviors using a scale of 0 (*not at all*), 1 (*weak potency*), 2 (*moderate potency*), and 3 (*strong potency*). Smith and colleagues described the potency rating as capturing the pervasiveness, intensity and expression of the behavior. The latter perspective does not consider the number of behaviors as most relevant; instead, the focus is on the extent of the psychological impact of the behavior upon the basic needs of the individual (Quested, Ntoumanis, Thøgersen-Ntoumani, Hagger, & Hancox, 2017; Smith et al., 2015). Via this approach it is possible for a high potency score for a specific behavior to be achieved even if the frequency rating is low (e.g., instructors may not belittle or devalue exercisers very often, but using such behaviors can still have a strong impact on the receiver). This potency score provides an indication of the 'psychological meaning' of certain behaviors that could be highly impactful, even if infrequently implemented (Smith et al., 2016). In the present study we developed a rating system to assess both the intensity of instructor behaviors (i.e., the perceived strength of the impact of such instructor behaviors upon the need satisfaction and need frustration of exercisers), as well as the frequency of their occurrence.

With regard to the theory-based content of SDT observational measures, there has been some diversity in the environmental dimensions coded. Most commonly, the need supportive features of the environment have been targeted (Tessier, Sarrazin, & Ntoumanis, 2010). Reeve and colleagues (Reeve et al., 2004) created a scale that assesses both positive and negative dimensions of the motivational environment using bipolar pairs (autonomy support vs control, structure vs chaos, interpersonal involvement vs hostility). However, more recent approaches (e.g., Haerens et al., 2013; Van den Berghe et al., 2013) measure positive and negative dimensions of the environment (need supportive vs thwarting) independently, in line with the premise that these dimensions operate as independent constructs, and therefore should be assessed as such (Bartholomew, Ntoumanis, & Thøgersen-Ntoumani, 2009, 2010). In the present study we designed an observation scale aligned with this perspective.

Making a Case for Need Indifferent Behaviors

Our scale includes items that represent need supportive and need thwarting behaviors of exercise instructors (see Hancox, Quested, Ntoumanis, & Thøgersen-Ntoumani, 2017). In addition, we also created a set of items to represent 'need indifferent' behaviors, defined as those that are void of any need supportive or need thwarting characteristics, but are relevant to group exercise (e.g., talking in ways that are motivationally 'empty', e.g., shouting "keep going" with little warmth or specificity, with regard to what, or whom must "keep going"). While phrases such as 'keep pushing' or 'go go go' may be interpreted as encouraging to some exercisers, unless they are directed towards a particular activity or goal, said with warmth and evident interest in the exerciser's strivings, or linked with informational feedback, there is no reason to expect them to have a direct impact on feelings of autonomy, relatedness or competence at performing the activity at hand. In other words, we consider that it is not the words *per se* that are motivationally relevant or irrelevant. Rather, it is a combination of words, tone, context and delivery that gives need relevant or indifferent

meaning to what instructors say. The new instrument was specifically designed to assess this. There have been limited previous attempts to incorporate need indifferent (or neutral) behaviors in SDT research (e.g., Kinnafick, Thøgersen-Ntoumani, & Duda, 2016; Tessier, Sarrazin, & Ntoumanis, 2008). For example, Tessier et al. (2008) observed PE teachers during gymnastics sessions and rated the teachers' behaviors as autonomy supportive, controlling, or neutral. However, Tessier et al. did not provide a theoretical rationale for the inclusion of need indifferent behaviors.

The rationale for including need-indifferent behaviors in the present study was both practical and theoretical. From a practical standpoint, the authors had personally experienced and also observed exercise class instructors employing ongoing rhetoric throughout a class that lacked qualities that would support or thwart autonomy, competence or relatedness. In intervention studies it is, therefore, important to code these motivationally empty (or indifferent) behaviors. This makes it possible to test whether an instructor changes these behaviors so that need supportive content more often characterises what the instructor choses to say when trying to positively impact the participant's experience as a result of SDT training. From a theoretical perspective, Bartholomew and colleagues (Kimberley Bartholomew, Ntoumanis, Ryan, Bosch, & Thøgersen-Ntoumani, 2011) emphasised the importance of differentiating between need frustration and need dissatisfaction at the personal level. The former involves an individual experiencing active frustration of their needs whereas the latter signifies a condition of unmet psychological needs. This distinction was further supported by Costa and colleagues (Costa, Ntoumanis, & Bartholomew, 2015) who demonstrated that need frustration was a better predictor of negative outcomes than need dissatisfaction. It follows then, that the social environment may also be characterized by need indifferent behaviors that neither support nor frustrate the three needs, in addition to those behaviors that support or thwart these needs. The provision of these need indifferent

behaviors represent 'missed opportunities' for need support. Such behaviors may also undermine attempts to be need supportive, by creating 'noise' that dilutes the salience of need supportive behaviors. A thorough assessment of the motivational characteristics of an exercise environment should therefore evaluate all three types of behavior: need supportive, need indifferent, and need thwarting. Existing SDT measures of the perceived or objective motivational environment do not assess need indifferent behaviors. In the present study we addressed this gap in the literature by creating items to objectively assess need indifferent behaviors in the exercise domain.

Study Aims and Hypotheses

In sum, the present study set out to create and evaluate a new observational tool, the Need-Relevant Instructor Behaviors Scale (NIBS), designed to assess the communication style of exercise instructors before and after they completed SDT-based motivation training, as part of a larger study (Hancox, Quested, Thøgersen-Ntoumani, & Ntoumanis, 2015; Ntoumanis, Thøgersen-Ntoumani, Quested, & Hancox, 2017). The motivation training included group education sessions focusing on theory and practical application and which were supported by online materials/discussion and 'homework' activities. The training program was designed to support instructors in increasing their provision of need supportive instruction and reducing the extent to which they employed need indifferent and thwarting behaviours. We hypothesized that from pre- to post intervention, the exercise instructors would decrease the number (frequency) of need thwarting behaviors and need indifferent behaviors during class, and increase the number of need supportive behaviors. We also predicted a decrease in the intensity (potency) of need thwarting behaviors and an increase in the intensity of need supportive behaviors.

Method

Participants and Procedure

Participants were an opportunity sample of 27 exercise instructors (2 male, 25 female) with a mean age of 37.23 years (SD = 8.23; age range = 26-58). Instructors had on average 3.97 years of experience working as a certified exercise instructor (SD = 3.32; range = 6 months - 14 years). Film and audio footages of 2 group cycling classes per instructor (54 classes in total) were rated by two trained raters using the NIBS (described below).

In between the two classes, instructors attended workshops that focused on how to create a more need supportive environment (and reduce need thwarting behaviors) in indoor cycling classes. In brief, the training was designed in collaboration with experienced cycling class instructors and the materials were designed specifically to suit this context. The nine-hour program (delivered in three 3-hour sessions) was grounded in SDT and also incorporated behaviour change techniques. The workshops were interactive and included small group work, activities, exemplar video clips, brain storming, and planning. In the practical sessions in the studio instructors had the opportunity to put into practice what they had learnt and to receive feedback on their use of motivational strategies. See Hancox, Quested, Thøgersen-Ntoumani, & Ntoumanis, 2015; Ntoumanis, Thøgersen-Ntoumani, Quested, & Hancox, 2017 for further details about intervention design, delivery, and content. Raters were instructed to code footage of each exercise class in 20 minute 'chunks' of time and to complete one rating sheet per 20 minute time chunk. Previous studies have supported the approach of coding 'chunks' of filmed footage in observational research in sport settings (Cheon & Reeve, 2013; Smith et al., 2015). Raters were blinded as to whether they were coding pre- or post-motivation training classes. They were instructed to not begin coding (and timing) until the instructor said the first word directed towards participants that was audible on the film. Raters finished coding when the instructor was no longer audible to the camera and/or when the last participant left the room. Prior to completing the ratings, the two raters (who were undergraduate psychology students) completed training led by the first author. The training

procedure was developed from that employed by Smith et al. (2015), and was approximately eight hours in total per coder. The training included education sessions on SDT and its application in exercise settings, group coding, and independent coding and feedback sessions in which the raters' coding was compared with a 'gold standard' rating from the lead researcher. Specifically, coder training involved the following steps: (1) Read papers that contributed to the design of the intervention strategies (Edmunds et al, 2009; Hancox et al., 2015; Deci et al., 1994; Haarens et al., Ntoumanis & Mallet, 2014; Mageau & Vallerand, 2003; Bartholomew et al., 2009; Reeve & Jang, 2006); (2) Session to discuss the underlying theory and the strategies the tool is designed to code; (3) Session to introduce the coding tool and the coding process; (4) Collaborative coding activity – coders and researchers work together to rate the strategies used in a videoed class during 3 music tracks (or more if deemed appropriate) (5) Independent coding #1 – coders worked as a pair to rate 3 tracks and then subsequently discuss coding with the researchers; (6) Independent coding #2 – coders independently and individually code 3 tracks and then compare ratings with 'gold standard' coding; (7) Coders rate one full class; (8) Coders re-code the same class rated in step 7 one week later; (9) Steps 4-8 repeated as necessary until required standard is reached.

Training continued until the interclass correlation co-efficient surpassed 0.70 indicating that an acceptable level (i.e., at least 0.7 Vincent, 1999) had been reached.

Measures

We created a new tool, the NIBS, for the purpose of this study. The tool comprises 17 items tapping common motivationally relevant exercise instructor behaviors (see Table 1.). Items were written by the research team or were taken from previous SDT literature (Bartholomew et al., 2009; Mageau & Vallerand, 2003; Reeve & Jang, 2006; Van den Berghe et al., 2013). These were presented to raters on a grid grouped into need supportive (9 items, e.g., "acknowledging the participants' feelings and responding appropriately"), need

indifferent (4 items, e.g., "talking in ways that are motivationally "empty" (e.g., "keep going")), and need thwarting (4 items, e.g., "criticising, belittling, devaluing or dismissing participant") behaviors. The intention was to produce a sufficient number items to represent a wide array of motivationally relevant behaviors that a group fitness instructor could employ, without at the same time resulting in an unwieldly list that would be impractical for the purpose of coding. The research team also cross-referenced the items with existing tools in the literature to ensure that the tool was comprehensive from a theoretical perspective.

The observation grid included condensed explanations of the types of behavior each item represented. In-depth descriptions of each item were also provided to raters to ensure conceptual clarity of what was being coded. The raters were instructed to use tallies to determine the frequency of use of each of the need supportive, thwarting, and need indifferent behaviors. The tallies are not grouped by need, as each behavior has the potential to impact all three needs, depending on how the behavior described in the item is employed. In contrast, the purpose of the intensities rating is to probe further as to the degree to which each need is impacted by the enactment of need supportive and need thwarting behaviors, based on the way in which the instructor has implemented the behavior. Hence, for need supportive and thwarting behaviours, raters were also asked to make a judgement as to the intensity of the behavior's impact three times, one for each of the three needs (i.e., an intensity score record). In the case of the nine items for supportive behaviors, intensity gauged the degree to which the use of that behavior in the time period may have led participants to experience satisfaction of autonomy, competence and/or relatedness. In the case of four items tapping thwarting behaviours however, the intensity rating provided a gauge of the degree to which autonomy, competence and relatedness were thwarted. Hence, intensity was rated three times (once per need) for each behaviour. Intensity and frequency scores were independent, however if a supportive or thwarting behaviour occurred (i.e., score of 1 or higher), then the intensity

rating for at least one of the three needs would have an intensity score of at least 1. Thorough

descriptions of how to interpret and apply the intensity ratings were also provided to raters

(see Table 2). For each rater, the frequency scores recorded for each chunk of time in each

exercise class were summed to create 'whole class' scores for the frequency of use of each

item on the rating grid. Totals were calculated to create one score for need supportive

behaviors (frequency), need thwarting behaviors (frequency), and need indifferent behaviors

(frequency). Using a similar procedure, a mean score was calculated to represent the overall

intensity of need supportive and need thwarting instructor behaviors in each class. Average

scores for frequency and intensity were then calculated from the two raters' scores to produce

one set of scores for each class.

Data Analysis

All analyses were conducted with Mplus, version 7.4 (Muthén & Muthén, 1998 –

2015). Interrater reliability (indicated by intraclass correlation coefficients; ICC) were

calculated for 10% of the frequency and intensity ratings.

To examine the intensity ratings we estimated a Bayesian exploratory factor analysis

(EFA) in Mplus (see Asparouhov & Muthén, 2012), using an oblique rotation (quartimin)

with the six intensity rating categories at time 1 as indicators. The EFA was performed on the

need-supportive and need thwarting intensities. Given that no intensity ratings are available

for the need indifferent behaviors (as they represent a void of need relevant content) they

could not be included in the EFA. EFA has been found to perform well in terms of factor

recovery also in very small samples, particularly when few factors are expected, model error

is low, and communalities are high (Preacher & MacCallum, 2002). The potential scale

reduction factor (PSRF; Brooks & Gelman, 1998) was used to assess convergence of the two

Markov chain Monte Carlo (MCMC) chains. A low (e.g., < 1.1) and stable PSRF is

considered as evidence of convergence (B. Muthén & Asparouhov, 2012). Model estimation

was performed with 100,000 iterations using the MCMC algorithm and the Gibbs sampler; the first half of the iterations (i.e., 50,000) was discarded as burn-in iterations. Overall data-model fit was evaluated using the posterior predictive p (PPP) value. A PPP value close to 0.50 with a 95% confidence interval (CI) centered on zero indicates a good data-model fit (B. Muthén & Asparouhov, 2012). We compared increasingly complex models from one to three factors and examined whether the inclusion of additional factors improved data-model fit. To compare the different models and decide upon the final solution, we inspected factor loading patterns, the deviance information criterion (DIC), and the Bayes Factor. The DIC is a relative measure of fit; a model with lower DIC value indicates a better model fit compared to a model with higher DIC value (Asparouhov, Muthén, & Morin, 2015). The Bayes factor (BF; Kass & Raftery, 1995) is a summary of evidence provided by the data in favour of one hypothesis (H_0) compared to another (H_1) and is calculated using the formula:

$$BF = \frac{P(H1)}{P(H0)} = \frac{Exp(-0.5BIC_{H1})}{Exp(-0.5BIC_{H0})}$$

A BF larger than 3 is considered evidence in support of H_1. In this study, a one factor model (H_0) was compared to a two factor model (H_1), and a two factor model (H_0) was compared to a three factor model (H_1).

Dependent samples t-tests were used to estimate the mean changes from pre-test to post-test. Following Lakens (2013) guidelines, Hedges g_{av} effect sizes were calculated to indicate the magnitude of the pre-test to post-test differences. The pre-test to post-test difference scores ($\Delta Y_{21i} = Y_{2i} - Y_{1i}$; Coman et al., 2013) of the tally and intensity ratings were correlated. Correlations between variables within and across the two time points are provided in the supplementary material (Table 5). To assess the magnitude of the pre-test to post-test differences at the instructor level, we used the reliable change index (RCI; Jacobson & Truax, 1991; Maassen, 2004). We calculated the RCI based on the formula:

$$RCI = \frac{x^2 - x^1}{SE}$$

In the numerator, x^1 and x^2 are the individual's score at the two time points. In the

denominator, SE is the standard error calculated using the formula (Maassen, Bossema, &

Brand, 2009):

$$SE = \sqrt{(S_X^2 + S_Y^2)(1 - r_{XY})}$$

In the equation S_X^2 is the pre-test variance, S_Y^2 is the post-test variance, and r_{XY} is the test-

retest reliability. The RCI is a measure of change in standardized units; it indicates the

direction of that change and whether it is reliable. A RCI greater than 1.96 in either direction

indicates a reliable change that is statistically significant at $p = .05$, whereas values between -

1.96 and 1.96 indicate that no reliable change has occurred (Jacobson & Truax, 1991). To

obtain a complete data set to perform the RCI calculations, missing values at the post-test

were imputed at the individual level using the expectation maximization algorithm (EM) in

IBM SPSS statistics, version 23.

Results

Coder reliability coefficients were acceptable (Vincent, 1999): frequency need support

(ICC =0.94), frequency need indifferent (ICC =0.71), frequency need thwarting (ICC =0.90),

intensity need support (ICC =0.72), intensity need thwarting (ICC =0.83).

The Bayesian EFA showed that the one-factor model had a poor data-model fit (PPP =

0.001, 95% CI [12.54, 56,64]) and the highest DIC value (137.78). The two-factor model

showed an acceptable data-model fit (PPP = 0.362, 95% CI [-18.87, 27.90]) and the lowest

DIC value (111.99); the BF indicated very strong evidence in favour of the two-factor model

compared to the one-factor model (BF = 12925.49). The factor loading pattern of the two-

factor model showed that the need thwarting behavior intensities loaded strongly (0.83 to

0.97) onto the first factor, whereas the need supportive behavior intensities loaded strongly

(0.68 to 0.90) onto the second factor. The cross-loadings were low and ranged from -0.06 to 0.17 (see Table 5 in supplementary material). The three-factor model showed an acceptable data-model fit (PPP = .347, 95% CI [-18.90, 29.84]), but the DIC value (116.99) was higher compared to the one of the two-factor model; the BF did not favour the three-factor model over the two-factor model (BF = 0.0056). Adding a third factor did not substantively change the factor loading pattern, which remained similar to the pattern of the two-factor model. Collectively, the EFA results clearly support a two-factor model representing need thwarting behavior intensities and need supportive behavior intensities.

Descriptive statistics and the pre-test to post-test differences are displayed in Table 3. The exercise instructors decreased their need thwarting behaviors and need indifferent behaviors, and increased their need supportive behaviors. A similar tendency was displayed for the intensities of these behaviors, with a substantive decrease in the intensities of autonomy, competence and relatedness thwarting behaviors, and an increase in the intensities of autonomy, competence and relatedness support in these behaviors. In general, effect sizes indicated substantive effect sizes (Hedges' g_{av} range from 0.53 to 1.65) of the intervention on the exercise instructors' frequencies of behaviors, as well as the intensities of those behaviors. Bivariate correlations among the difference scores (post- to pre-intervention) showed positive correlations between the frequency of need thwarting behaviors and the intensities of the need thwarting behaviors, and positive correlations between the frequency of need supportive behaviors and the intensities of the need supportive behaviors (Table 4). Frequency of need thwarting behaviors were negatively associated with autonomy support intensity.

To further understand the impact of the intervention at the instructor level, we calculated a RCI value for each of the exercise instructors and have summarized these results in Table 3. More than half (59.3%) of the exercise instructors reliably decreased their need thwarting behaviors and 22.2% reliably increased their need supportive behaviors. Regarding

need indifferent behaviors, most participants were consistent in the use of these behaviors across time, none of the instructors increased such behaviors, whereas a small percentage (11.1%) decreased their need indifferent behaviors. With regard to the intensity scores, the percentage of exercise instructors that showed a reliable increase in the intensity of need supportive behaviors ranged from 11.1% to 55.6% and those showing a reliable decrease of need thwarting behaviors ranged from 14.8% to 40.7%. At the individual need level, the intervention seemed to have the largest impact on the intensity of competence support behaviors (55.6% reliably increased).

Discussion

The development and testing of interventions designed to improve the quality of communication skills of leaders, including exercise instructors, has become an important focus of research among SDT-based scholars in the physical domain. Yet, there are very few observation tools developed to objectively code both the frequency and intensity of the motivational characteristics of communication strategies, a void that is particularly notable in exercise research. The purpose of the present study was to address this gap in the literature by developing and testing the NIBS, an observational tool designed to aid researchers in assessing need supportive, need thwarting, and need indifferent exercise instructor behaviors. Extending existing observation tools grounded in SDT (e.g., Reeve et al., 2004; Smith et al., 2015; Webster et al., 2013), the NIBS was developed to provide a more complete assessment of the motivational environment. The tool included items tapping need indifferent behaviors; that is motivationally relevant instructor behaviors that would be expected to neither support nor thwart exerciser's needs which can co-occur alongside active supportive and thwarting behaviors. Overall, the findings support the utility of the tool as a method to provide a reliable assessment of the frequency and intensity of motivationally relevant communication of exercise instructors.

Our findings provide support for the psychometric properties of the NIBS. The inter-rater reliability between the two trained raters was very good, indicating consistency in the raters' interpretation of the behaviors observed in the videos. Examination of the factor structure of the NIBS intensity showed strong support favouring the two-factor model, with need supportive and need thwarting behavior intensities loading strongly onto two different factors. This highlights that the NIBS can clearly differentiate between the need supportive and need thwarting characteristics of the different behaviors included in the tool.

The sensitivity of the NIBS as a means to detect changes in the frequency of behaviors used by instructors from pre- to post- completion of a motivation training program was also demonstrated in the present study. Specifically, findings showed that the exercise instructors decreased the number of times they employed need thwarting behaviors and need indifferent behaviors during classes, and increased their use of need supportive behaviors. This suggests that the tool was sensitive to detect change in frequency of use of such behaviors. The changes observed by the raters were comparable to the changes in perceptions of instructor behaviors reported by the exercisers (cf. Ntoumanis et al., 2017).

Our study contributes to the SDT literature and existing observational measures by including items to tap the frequency of use of common need indifferent exercise instructor behaviors that lack any need supportive or thwarting characteristics. Although a small percentage of instructors (11.1%) decreased their need indifferent behaviors and such a reduction was large in terms of effect size, when the data were examined at the instructor level the findings suggested that the intervention had a limited effect on instructors' frequency of use of such behaviors. This finding may be because these need indifferent behaviors were the least frequently employed of the three types of behaviors we studied. It is also possible that need supportive and need thwarting behaviors require more cognitive effort and attention, whereas indifferent behaviors are less cognitively demanding and perhaps more habitual in

nature. For example, saying "well done" or "keep going" with little specificity is not cognitively taxing and can be almost an automatic utterance. As such, they may be harder behaviors to modify.

Every use of a need indifferent behavior could be considered as a 'missed opportunity' to support exercisers' needs. The video footage supported this point; many instructors had ongoing dialogue throughout a cycling class, but much was motivationally "empty".

Assessing need indifferent behaviours may be very useful in intervention work, given the regularity with which many instructors use a communication style that does not fully support the needs, but also does not actively thwart these needs either. That is not to say that instructors should necessarily increase the frequency of need supportive communications if their communication style is already rich in need support, as silence is also important to give participants time to reflect, think, and cognitively engage in the activity. At this stage we do not hypothesise a 'magic', quantifiable formula of what would be the optimal frequency of need supportive communication. However, we predict that high intensity of need support throughout the class is most optimal, and this is not dependent on frequency of communication. Future research will be necessary to further explore the unique and co-dependent role that intensity and frequency of need-relevant behaviors may play in predicting need satisfaction and frustration reports of exercisers. Research is also warranted that explores whether participants perceive specific instructor behaviours coded as need indifferent to be void of motivational meaning.

Future interventions could address this issue by developing methods to help instructors become more aware of when and why they use behaviors that, despite intentions to motivate exercisers, lack motivational underpinning. Training instructors to code videos of their own classes, or providing individual feedback could help them to develop better self-awareness.

Technological advances in the ease of recording, viewing, and coding footage could facilitate the incorporation of this behavior change strategy into future interventions.

The behavioral intensity findings provide evidence that the NIBS is sensitive to detect changes over time in how strategies are implemented. This finding suggests that the tool can be used not only to detect increases in use of need supportive strategies (captured by the frequency ratings) that may occur following an intervention, but also changes in the quality of implementing the strategies in such a way as to support the three basic needs. It is noteworthy that a larger percentage of instructors increased the intensity of need supportive behaviors compared to those showing a reliable decrease in the intensity of need thwarting behaviors. This suggests the intervention was more effective in making good practice even better, than it was in changing bad practices to become less detrimental. However, it is also worth noting that the mean ratings for intensity were higher for need thwarting than need satisfaction prior to the intervention, but this switched after the intervention.

It is noteworthy that at the individual level, the intervention had the largest impact on intensity of competence support behaviors (74.1% reliably increased). In the intervention design, particular attention was paid to how all three needs could be equally supported by instructors. It is possible that competence support changed most because instructors focused more on competence-related actions, as group cycling is a physical task in which competence may be the more salient need. In future research it would be of interest to examine whether changes in intensity of autonomy and relatedness related behaviors are greater in interventions targeting instruction of classes that lend themselves more to one-to-one dialogue and group interaction (e.g., aerobics, circuit training).

An important characteristic of interventions grounded in SDT is that need support does not operate as a 'dose-response' relationship (Quested et al., 2017). That is, it is not just important to target *how many* need supportive behaviors an instructor uses, but it is also

critical to address *how well* these behaviors target the three needs. The present study clarifies that it is possible to reliably measure quality of exercise instructors' behaviors, alongside the quantity of such behaviors, using the NIBS. In the future, the tool could be used to examine which features of intervention are relevant to changes in what instructors do (i.e., how many times), and the quality of implementation (i.e., intensity) of such behaviors. Moreover, future research should also address whether participants' experiences of need satisfaction or frustration during an exercise class are related to the use of need supportive and need thwarting behaviors employed by the instructor in the class.

Despite its strengths, as with all observation research this study is also limited by a number of factors. Firstly, it is possible that the instructors' behaviors were not typical, due to their awareness that they were being observed (i.e., Hawthorne effect, McCambridge, Witton & Elbourne, 214). We do not feel this effect had a strong impact on our findings as across both time points we witnessed a range of need supportive, thwarting, and indifferent behaviours. The instructors were also instructors are used to being observed as part of their general instructor training so may not be as reactive to being filmed. Our study was also limited to a relatively small number of indoor cycling instructors. Future research involving larger samples and a wider array of class types being instructed would be needed to test the replicability of our findings.

The NIBS makes a significant contribution to the literature concerning motivational features of the exercise environment. The scale complements existing measures designed to objectively assess the motivationally relevant content of the physical activity environment. For example, the dual coding of intensity and frequency of need supportive and thwarting behavior enactment provides a more meaningful evaluation of the motivational environment manifested in physical activity contexts, than could be ascertained from those measures designed to assess only the frequency of particular behaviors (e.g., Reeve et al., 2004;

Webster et al., 2013) or only the potency of the motivational environment (Smith et al., 2016). With regard to frequency ratings, in recognition of the fact that it is very rare that any behavior could be enacted 100% of the time, the frequency scale in the proposed measure does not take an 'all or nothing' approach. This extends the approach developed by Haerens and colleagues (Haerens et al., 2013; Van den Berghe et al., 2013) in which observers are asked to rate the frequency of specified behaviors on a continuum ranging from 0 ('*not at all'*) to 3 ('*all of the time'*). Our approach also extends beyond existing SDT observational measures that have adopted a bipolar assessment of positive and negative dimensions of the motivational environment (Reeve et al., 2004). While more recent approaches (e.g., Haerens et al., 2013; Van den Berghe et al., 2013) measure positive and negative dimensions of the motivational environment (need supportive vs need thwarting) independently, our tripartite measure also assess those behaviors lacking need supportive or need thwarting characteristics. This approach builds on limited previous research that attempted to assess need indifferent behaviors (e.g., Kinnafick et al., 2016; Tessier et al., 2008).

Conclusions

The present study's findings support the utility of the NIBS as a means to observe and reliably code the need supportive, need thwarting, and need indifferent behaviors of exercise instructors. Our findings also support the tool's sensitivity to detect changes in intensity and quality of these behaviors and have implications for intervention design and measurement of the social environment in exercise and other physical activity contexts.

References

Asparouhov, T., & Muthén, B. (2012). *Comparison of computational methods for high dimensional item factor analysis*. Retrieved from http://www.statmodel.com:

Asparouhov, T., Muthén, B., & Morin, A. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances: Comments on Stromeyer et al. *Journal of Management, 41*, 1561-1577. doi:10.1177/0149206315591075

Bartholomew, K., Ntoumanis, N., Ryan, R., Bosch, J., & Thøgersen-Ntoumani, C. (2011). Self-determination theory and diminished functioning: The role of interpersonal control and psychological need thwarting. *Personality and Social Psychology Bulletin, 37*(11), 1459-1473. doi:10.1177/0146167211413125

Bartholomew, K., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2009). A review of controlling motivation strategies from a self-determination theory perspective: Implications for coaches. *International Review of Sport and Exercise Psychology, 2*, 215-233. doi:doi.org/10.1080/17509840903235330

Bartholomew, K., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2010). The controlling interpersonal style in a coaching context: Development and initial validation of a psychometric scale. *Journal of Sport and Exercise Psychology, 32*(2), 193-216. doi:doi.org/10.1123/jsep.32.2.193

Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics, 7*(4), 434-455. doi:10.1080/10618600.1998.10474787

Cheon, S. H., & Reeve, J. (2013). Do the benefits from autonomy-supportive PE teacher training programs endure?: A one-year follow-up investigation. *Psychology of Sport and Exercise, 14*(4), 508-518.

Coman, E. N., Picho, K., McArdle, J. J., Villagra, V., Dierker, L., & Iordache, E. (2013). The paired t-test as a simple latent change score model. *Frontiers in psychology, 4*, 738. doi:10.3389/fpsyg.2013.00738

Costa, S., Ntoumanis, N., & Bartholomew, K. (2015). Predicting the brighter and darker sides of interpersonal relationships: Does psychological need thwarting matter? *Motivation and Emotion, 39*(1), 11-24.

Darst, P. W., Zakrajsek, D., & Mancini, V. H. (1989). *Analyzing physical education and sport instruction*. Champaign, IL: Human Kinetics.

Duda, J. L., Williams, G. C., Ntoumanis, N., Daley, A., Eves, F. F., Mutrie, N., . . . Jolly, K. (2014). Effects of a standard provision versus an autonomy supportive exercise referral programme on physical activity, quality of life and well-being indicators: a cluster randomised controlled trial. *International Journal of Behavioral Nutrition and Physical Activity, 11*(1), 10. doi:10.1186/1479-5868-11-10

Edmunds, J., Ntoumanis, N., & Duda, J. L. (2007). Perceived autonomy support and psychological need satisfaction as key psychological constructs in the exercise domain. In M. S. M. Hagger & & N. L. D. Chatzisarantis (Eds.), *Self-determination in exercise and sport* (pp. 35-51). Champaign, IL: Human Kinetics.

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. doi:10.1002/ejsp.463

Fortier, M. S., Sweet, S. N., O'Sullivan, T. L., & Williams, G. C. (2007). A self-determination process model of physical activity adoption in the context of a randomized controlled trial. *Psychology of Sport and Exercise, 8*(5), 741-757. doi:10.1016/j.psychsport.2006.10.006

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.

Hamer, M., & Chida, Y. (2009). Physical activity and risk of neurodegenerative disease: a systematic review of prospective evidence. *Psychological medicine, 39*(1), 3-11. doi:10.1017/s0033291708003681

Hancox, J. E., Ntoumanis, N., Thogersen-Ntoumani, C., & Quested, E. (2015). Self-Determination Theory. In J. Middelkamp (Ed.), *Essentials of motivation and behaviour change* (pp. 68–85). Brussels: Europe Active.

Hancox, J. E., Quested, E., Ntoumanis, N., & Thøgersen-Ntoumani, C. (2017). Putting self-determination theory into practice: application of adaptive motivational principles in the exercise domain. *Qualitative Research in Sport, Exercise and Health*, 1-17. doi:10.1080/2159676x.2017.1354059

Hancox, J. E., Quested, E., Thøgersen-Ntoumani, C., & Ntoumanis, N. (2015). An intervention to train group exercise class instructors to adopt a motivationally adaptive communication style: a quasi-experimental study protocol. *Health Psychology and Behavioral Medicine, 3*(1), 190-203. doi:10.1080/21642850.2015.1074075

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of consulting and clinical psychology, 67*(3), 300-307. doi:10.1037//0022-006x.67.3.300

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology, 59*(1), 12-19. doi:10.1037//0022-006x.59.1.12

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. Journal of the American Statistical Association, 90(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kinnafick, F.-E., Thøgersen-Ntoumani, C., & Duda, J. (2016). The effect of need supportive text messages on motivation and physical activity behaviour. *Journal of behavioral medicine, 39*(4), 574-586. doi:10.1007/s10865-016-9722-1

Kohl, H., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J., Leetongin, G., & Kahlmeier, S. (2012). The pandemic of physical inactivity: global action for public health. *The Lancet, 380*(9838), 294-305. doi:10.1016/s0140-6736(12)60898-8

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863. doi:10.3389/fpsyg.2013.00863

Mageau, G. A., & Vallerand, R. J. (2003). The coach–athlete relationship: A motivational model. *Journal of sports science, 21*(11), 883-904.

Marcus, B. H., Williams, D. M., Dubbert, P. M., Sallis, J. F., King, A. C., Yancey, A. K., . . . Claytor, R. P. (2006). Physical activity intervention studies. *Circulation, 114*(24), 2739-2752. doi:10.1161/CIRCULATIONAHA.106.179683

Maassen, G. (2004). The standard error in the Jacobson and Truax Reliable Change Index: The classical approach to the assessment of reliable change. J*ournal of the International Neuropsychological Society*, 10(6), 888-893. doi:10.10170S1355617704106097

Maassen, G. H., Bossema, E., & Brand, N. (2009). Reliable change and practice effects: Outcomes of various indices compared. *Journal of Clinical and Experimental Neuropsychology*, 31(3), 339-352. doi:10.1080/13803390802169059

Mildestvedt, T., Meland, E., & Eide, G. E. (2008). How important are individual counselling, expectancy beliefs and autonomy for the maintenance of exercise after cardiac

rehabilitation? *Scandinavian Journal of Social Medicine, 36*(8), 832-840.

doi:10.1177/1403494808090633

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible

representation of substantive theory. *Psychological methods, 17*(3), 313-335.

doi:10.1037/a0026802

Muthén, L. K., & Muthén, B. O. (1998 –2015). *Mplus user's guide* (7 ed.). Los Angeles, CA:

Muthén & Muthén.

Ntoumanis, N., Quested, E., Reeve, J., & Cheon, S. H. (2018). Need supportive

communication: Implications for motivation in sport, exercise, and physical activity.

In B. Jackson, J. A. Dimmock, & J. Compton (Eds.), *Persuasion and communication

in sport, exercise, and physical activity*. Abingdon, UK: Routledge.

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. doi:10.1111/sms.12713

Preacher, K. J., & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics

research: Factor recovery with small sample sizes. *Behavior genetics, 32*(2), 153-161.

doi:10.1023/A:1015210025234

Quested, E., Ntoumanis, N., Thøgersen-Ntoumani, C., Hagger, M. S., & Hancox, J. E. (2017).

Evaluating quality of implementation in physical activity interventions based on

theories of motivation: current challenges and future directions. *International Review

of Sport and Exercise Psychology, 10*(1), 252-269.

doi:10.1080/1750984X.2016.1217342

Rahman, R. J., Hudson, J., Thøgersen-Ntoumani, C., & Doust, J. H. (2015). Motivational

processes and well-being in cardiac rehabilitation: a self-determination theory

perspective. *Psychology, health & medicine, 20*(5), 518-529.

doi:10.1080/13548506.2015.1017509

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.

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. doi:10.1023/B:MOEM.0000032312.95499.6f

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*: Guilford Publications.

Smith, N., Quested, E., Appleton, P. R., & Duda, J. L. (2016). A review of observational instruments to assess the motivational environment in sport and physical education settings. *International Review of Sport and Exercise Psychology, 9*(1), 134-159. doi:10.1080/1750984X.2015.1132334

Smith, N., Tessier, D., Tzioumakis, Y., Quested, E., Appleton, P., Sarrazin, P., . . . Duda, J. L. (2015). Development and validation of the multidimensional motivational climate observation system. *Journal of Sport and Exercise Psychology, 37*(1), 4-22. doi:10.1123/jsep.2014-0059

Standage, M., & Ryan, R. M. (2012). Self-determination theory and exercise motivation: Facilitating self-regulatory processes to support and maintain health and well-being. In G. C. Roberts & D. C. Treasure (Eds.), *Advances in motivation in sport and exercise* (3rd ed., pp. 233-270). Champaign: Human Kinetics.

Teixeira, P. J., Carraça, E. V., Markland, D., Silva, M. N., & Ryan, R. M. (2012). Exercise, physical activity, and self-determination theory: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 9*(1), 78. doi:10.1186/1479-5868-9-78

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. doi: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. doi:10.1016/j.cedpsych.2010.05.005

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.

Vincent, W. J. (1999). *Statistics in Kinesiology* Champaign, IL: Human Kinetics.

Webster, C. A., Wellborn, B., Hunt, K., LaFleche, M., Cribbs, J., & Lineberger, B. (2013). MPOWER: An observation system for assessing coach autonomy support in high school varsity boys' soccer practices. *International Journal of Sports Science & Coaching, 8*(4), 741-754. doi:10.1260/1747-9541.8.4.741

Table 1

Items of the Need-Relevant Instructor Behavior Scale

Higher Order Factor	Item
Need Thwarting	Using a language or a tone that is pressurising or could induce feelings of guilt or shame when communicating commands or goals
	Criticising, belittling , devaluing or dismissing participants
	Showing disregard or rejection for participants' feelings, preferences, opinions and feedback
	Comparing participants against each other or being overly competitive
Need indifferent	Providing goals and rules with no explanations, or explanations that are unclear or confusing
	Talking in ways that are motivationally "empty" (e.g., "keep going")
	Offering choices that are not meaningful
	Appearing to talk to a "camera"
Need supportive	Taking time to listen and be responsive to the participants' needs
	Encouraging questions and feedback from the participants about their goals, problems or preferences
	Giving meaningful and appropriate explanations
	Giving specific and constructive feedback
	Using an inclusive language (e.g., "we could try…")
	Acknowledging the participants' feelings and responding appropriately
	Offering meaningful praise which is unconditional
	Create opportunities for the participants to have input, choice and make decisions about the workout
	Creating opportunities to interact with all participants

Table 2

Coding Rubric for Rating Intensity of Need Supportive/Thwarting Behaviors

Score	Anchor	Description
0	Not at all	If the behavior has not occurred (i.e., there is no frequency count) then intensity will be 0. If there is a value for frequency count, then intensity must be at least 1.
1	Slightly	Overall in this time period, the way in which this behavior was delivered (i.e., specific language used, specific content of the message conveyed, strength of expressiveness/tone) was <u>a little</u> supporting/thwarting of participants' autonomy/competence/relatedness. However, there would be <u>a considerable</u> number of ways the instructor could have been even more need supportive/thwarting in the delivery of this behavior during this time period by some <u>considerable</u> changes to his/her phrasing or choice of words, content of the message and/or tone or level of expressiveness in delivery of the behavior.
2	Moderately	Overall in this time period, the way in which this behavior was delivered (i.e., specific language used, specific content of the message conveyed, strength of expressiveness/tone) was <u>moderately</u> supporting/thwarting of participants' autonomy/competence/relatedness. However, there would be <u>a few</u> ways that the instructor could have been even more need supportive/thwarting in the delivery of this behavior during this time period by some <u>a few</u> changes to his/her phrasing or choice of words, content of the message and/or tone or level of expressiveness in delivery of the behavior.
3	Highly	Overall in this time period, the way in which this behavior was delivered (i.e., specific language used, specific content of the message conveyed, strength of expressiveness/tone) <u>was highly </u>supporting/thwarting of participants' autonomy/competence/relatedness. There is <u>nothing or almost nothing</u> that the instructor could have done to be even more need supportive/thwarting in the delivery of this behavior during this time period.

Table 3

Descriptive Statistics, Paired Samples t-tests, Effect Sizes of the Pre-test and Post-test Differences, and Reliable Change of the Exercise

Instructors' Behaviors

	Pre-test		Post-test					Reliable Change*		
	M	*SD*	*M*	*SD*	*t*	*p* value	Hedges' g_{av}	*% decrease*	*% stable*	*% increase*
Tallies										
Need thwarting	16.70	12.29	5.56	6.30	7.80	.000	1.11	48.1	51.9	0.0
Need support	25.56	8.85	42.62	14.82	6.49	.000	1.36	0.0	55.6	44.4
Need indifferent	13.07	7.76	7.43	5.55	4.16	.000	0.81	11.1	88.9	0.0
Intensities										
Need thwarting	0.77	0.69	0.26	0.34	6.54	.000	0.91	22.2	77.8	0.0
Autonomy thwarting	1.22	0.71	0.47	0.44	8.76	.000	1.23	0.0	100.0	0.0
Competence thwarting	0.82	0.98	0.21	0.46	5.09	.000	0.77	22.2	77.8	0.0
Relatedness thwarting	0.26	0.48	0.06	0.16	3.02	.0052	0.53	11.1	88.9	0.0
Need support	0.65	0.22	1.24	0.44	7.65	.000	1.65	0.0	44.4	55.6
Autonomy support	0.50	0.23	1.01	0.41	6.26	.000	1.47	0.0	66.7	33.3
Competence support	0.53	0.25	1.20	0.51	8.92	.000	1.63	0.0	25.9	74.1
Relatedness support	0.92	0.29	1.51	0.58	5.07	.000	1.26	0.0	59.3	40.7

Note: *A decrease or increase indicate a critical value lower than -1.96 or higher than 1.96.

Table 4

Correlations of Difference Scores

Tallies	1	2	3	4	5	6	7	8	9	10
1. Need thwarting										
2. Need support	-0.170									
3. Need indifferent	0.144	0.302								
Intensities										
4. Need thwarting	0.800***	-0.087	0.124							
5. Autonomy thwarting	0.762***	0.013	0.141	0.889***						
6. Competence thwarting	0.718***	-0.136	-0.024	0.949***	0.745***					
7. Relatedness thwarting	0.668***	-0.073	0.371	0.834***	0.635***	0.734***				
8. Need support	-0.321	0.677***	0.087	-0.246	-0.226	-0.272	-0.098			
9. Autonomy support	-0.397*	0.501**	-0.094	-0.214	-0.152	-0.200	-0.219	0.780***		
10. Competence support	-0.308	0.387*	-0.060	-0.290	-0.329	-0.260	-0.140	0.849***	0.601**	
11. Relatedness support	-0.169	0.731***	0.246	-0.151	-0.130	-0.227	0.036	0.875***	0.463*	0.611***

Note. $p < .05$, ** $p < .01$, *** $p < .001$