A moral intervention reduces doping likelihood in British and Greek athletes

Kavussanu, Maria; Hurst, Philip; Yukhymenko-Lescroart, Mariya; Galanis, Evangelos; King, Ailish; Hatzigeorgiadis, Antonis; Ring, Chris

DOI:
10.1123/jsep.2019-0313

A Moral Intervention Reduces Doping Likelihood in UK and Greek Athletes: Evidence from a Cluster Randomized Control Trial

Maria Kavussanu[1], Philip Hurst[2], Mariya Yukhymenko-Lescroart[3],

Evangelos Galanis[4], Ailish King[5], Antonis Hatzigeorgiadis[4], and Christopher Ring[1]

[1] University of Birmingham

[2] Canterbury Christ Church University

[3]California State University, Fresno

[4] University of Thessaly

[5]Loughborough University

In Press: Journal of Sport & Exercise Psychology

Author Notes

This research was funded by the World Anti-Doping Agency (WADA), Social Science Research program (2015 – 2018). The authors would like to express their appreciation to WADA for funding this project.

Philip Hurst and Ailish King were research associates at the University of Birmingham when this research was conducted.

Correspondence concerning this article should be addressed to Maria Kavussanu, School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, B15 2TT, UK. E mail: m.kavussanu@bham.ac.uk

Abstract

Objectives. We aimed to develop a moral intervention and determine whether it is more effective than an educational (i.e., knowledge-based) intervention; our primary outcome was doping likelihood, and our secondary outcomes were moral identity, moral disengagement, moral atmosphere, and anticipated guilt.

Methods. Eligible athletes ($N = 303$) in the UK and Greece took part in the study. We randomly assigned 33 clubs to either the moral or the educational intervention. We measured our outcomes pre and postintervention and at a 3- and 6-month follow up.

Results. Athletes in both interventions in both countries reported lower doping likelihood and moral disengagement and higher guilt from pre to postintervention. These effects were maintained at the 3- and 6-month follow up. There were no effects on moral identity or moral atmosphere.

Conclusions. As well as disseminating information about doping, doping prevention programs should include content that focuses on moral variables.

Keywords: moral disengagement; moral identity, moral atmosphere, anticipated guilt

A Moral Intervention Reduces Doping Likelihood in UK and Greek Athletes: Evidence from a Cluster Randomized Control Trial

The use of prohibited performance-enhancing substances and methods - also known as doping - can have serious adverse health consequences for athletes, compromises fair play, and contributes to a negative image of sport in society (e.g., Quaglio et al., 2009). Despite significant financial investment in detection methods by governments worldwide (see WADA 2018 Annual Report), research evidence suggests that doping occurs at alarming levels, particularly in elite populations. For example, in their review of the relevant literature, de Hon et al. (2015) concluded that based on a combination of models of biological parameters and questionnaires, 14-39% of adult elite athletes intentionally use doping. In addition, a recent survey of elite athletes taking part in two international sporting events using the "randomized response technique" estimated that the prevalence of doping in the previous year ranged from 44% to 57% (Ulrich et al., 2018). Thus, the need to develop effective interventions that prevent doping in sport is clear.

One approach to developing effective anti-doping interventions is to intervene on psychological factors that have been strongly asso*ci*ated with doping likelihood[1] in empirical studies (e.g., Kavussanu & Ring, 2017; Kavussanu et al., 2020; Ntoumanis et al., 2014). However, to date, no study has examined whether an intervention focusing exclusively on such factors is effective in reducing doping likelihood, particularly in the long term. Moreover, as doping is a universal phenomenon, there is a need to develop and evaluate interventions that are effective in preventing this behavior in athletes from different countries.

Current Anti-Doping Interventions

Several anti-doping interventions have been developed and evaluated in the past two decades. The first anti-doping interventions were the ATLAS (Athletes Training and Learning to

Avoid Steroids) and ATHENA (Athletes Targeting Healthy Exercise & Nutrition Alternatives) programs (Goldberg & Elliot, 2005). These programs convey knowledge about a range of unhealthy behaviors (e.g., use of anabolic steroids, nutritional supplements and alcohol) to participants during multiple sessions. However, studies examining these programs showed a small reduction in reported doping in the intervention group compared to the control group (Elliot et al., 2008; Goldberg et al., 1996; Goldberg et al., 2000) and a non-significant reduction in the number of reported cases of doping over a season or a school year compared to a control group (Goldberg et al., 2003; Goldberg et al., 2000; Ranby et al., 2009).

One explanation for the limited effectiveness of these interventions is that they are not sufficiently focused, as they aim to influence athletes' overall health-related behaviors (Ntoumanis et al., 2014). For example, ATHENA includes information and activities about healthy and unhealthy eating, as well as the use of alcohol, tobacco, marijuana and anabolic steroids (Elliot et al., 2004; Elliot et al., 2008; Ranby et al., 2009). A second explanation is that participants had low initial intentions to dope (i.e., a floor effect), and therefore, there was not much room for a reduction in their doping intentions and behaviors (see Ntoumanis et al., 2014). Importantly, none of these interventions manipulated psychological variables that have been associated with doping likelihood or behavior in empirical research (Backhouse et al., 2016; Kavussanu et al., 2020; Ntoumanis et al., 2014).

More recent interventions have educated athletes about the health risks of prohibited substances and healthy alternatives and also included content relevant to psychological factors that have been empirically associated with doping. For example, Barkoukis et al. (2016) educated participants on the moral, social and psychological aspects of doping; Sagoe et al. (2016) informed participants about the ethics of doping, and how to resist peer pressure to dope; and Lucidi et al. (2017) discussed the way the media may disregard or minimize the moral

implications of doping. However, although two of these interventions were successful in significantly reducing doping attitudes from pre to postintervention (Barkoukis et al., 2016; Lucidi et al., 2017), doping intention either did not change in the intervention group (Lucidi et al., 2017; Sagoe et al., 2016) or was not measured (Barkoukis et al., 2016). Although doping attitudes are an important outcome, it could be argued that doping intention, likelihood or susceptibility should better predict athletes' actual doping behavior. Finally, grounded on ethical decision making, Elbe and Brand (2016) developed an online intervention with sport-specific moral dilemmas aimed at changing moral reasoning. Unexpectedly, the intervention increased athletes' doping attitudes.

The Need for Evidence-Based Interventions

Anti-doping interventions are more likely to be effective if they focus on changing factors that have been associated with doping likelihood in empirical research. A growing literature suggests that moral variables are particularly relevant for doping prevention (e.g., Boardley et al., 2017; Kavussanu, 2019; Kavussanu et al., 2016; Kavussanu & Ring, 2017; Ntoumanis et al., 2014). These findings make sense, as doping is behavior that breaks the rules of sport and compromises fair play. Three variables that have been shown in past research to be particularly important for doping are moral identity, moral disengagement, and moral atmosphere. We discuss each of these variables next.

Moral identity is a self-conception organized around a set of moral traits, such as being fair, honest, hardworking, friendly, and kind (Aquino & Reed, 2002). At the heart of the moral identity construct are the values of honesty and fairness, two of the values encompassed in the Spirit of Sport, an important part of the Code of the World Anti-Doping Agency (WADA Code, 2015). People with a strong moral identity endorse these values and are motivated to behave in an

ethical manner, due to their desire to maintain consistency between conceptions of their moral self and their actions (Aquino et al., 2009). Several studies have shown that moral identity may be a protective factor for doping. For example, in a study of British athletes from a variety of team sports (Kavussanu & Ring, 2017) the stronger the athletes' moral identity, the lower their likelihood to use banned substances to enhance their performance or recover from injury. These findings, which were corroborated in a second study of British, Danish, and Greek footballers (Kavussanu et al., 2020), underline the importance of moral identity for doping prevention.

Moral disengagement pertains to the cognitive mechanisms that individuals use to minimize negative emotions, such as guilt, which typically arise when people engage in unethical behavior that violates their moral standards (Bandura, 1999). For instance, athletes may absolve themselves of responsibility by thinking that "everybody does it" or that their coach, medical personnel or team captain told them to do it (i.e., diffusion and displacement of responsibility). They can ignore, distort, or minimize the consequences of their transgressive behavior for others (i.e., distortion of consequences) and contrast doping with worse behaviors, such as using illegal drugs (i.e., advantageous comparison) thereby making doping appear less serious. Finally, they may refer to doping as "juice" or "vitamins" (i.e., euphemistic labelling), so that the behavior does not sound as bad. Numerous studies have shown strong positive relationships between moral disengagement and proxies of doping behavior such as doping likelihood, intention, and susceptibility (Boardley et al., 2017; Kavussanu et al., 2016; Kavussanu et al., 2020; Ntoumanis et al., 2017), while in more recent research, manipulating moral disengagement (via scenarios) led to increased doping likelihood (Ring & Hurst, 2019; Stanger & Backhouse, 2020). Thus, moral disengagement is another important factor that should be targeted in our doping-prevention efforts.

The social context within which doping takes place can facilitate doping behavior. Over

time, sport teams develop a shared understanding of what is acceptable behavior within the team. This shared understanding which is reflected on the behavior of the sport team, has been labelled moral atmosphere (Stephens & Bredemeier, 1996). Thus, moral atmosphere refers to the collective group norms regarding moral action, that is, the type of behavior considered acceptable in a group by its group members. In sport research, this construct has been typically measured via athlete perceptions of team norms, that is, by asking athletes to indicate the number of teammates they perceived would cheat or engage in other transgressive acts in hypothetical situations (e.g., Kavussanu et al., 2002; Ommundsen et al., 2003). Using this approach, basketball and football players who perceived a team moral atmosphere condoning cheating and aggression were more likely to report that they would also engage in these behaviors (Kavussanu et al., 2002; Kavussanu & Spray, 2006). A recent meta-analysis (Spruit et al., 2019) showed a strong relationship between the moral climate of the team - a term used to refer to moral atmosphere - and self-reported moral behavior.

Strengthening athletes' moral identity, reducing their moral disengagement, and weakening a team moral atmosphere that condones doping, should lead them to anticipate experiencing more guilt if they doped, which should in turn deter doping. Guilt is a self-conscious moral emotion that plays an important role in regulating moral behavior (Tangney et al., 2007). Anticipated guilt has been inversely associated with doping likelihood in previous studies (e.g., Boardley et al., 2017; Kavussanu & Ring, 2017; Ring et al., 2019); in some of these studies, moral identity and moral disengagement were indirectly related to doping likelihood via anticipated guilt. Thus, by targeting moral identity and moral disengagement, an intervention can also increase anticipated guilt, with subsequent effects on doping.

The Present Research

In sum, despite efforts to prevent doping in sport, the problem of doping remains, and there

is an urgent need to develop and evaluate effective anti-doping programs. Although a stronger focus on moral variables, which have been linked to doping likelihood in empirical research is needed, no study has intervened on such variables. In addition, intervention studies are limited in that they have been conducted in a single country. Doping is a global issue and research is needed to determine whether interventions can be used effectively with athletes from different countries and cultures. Replicating findings across cultures would help us establish general laws of intervention effectiveness across cultures. Importantly, such interventions would greatly contribute to the efforts of WADA, the international organization that aims to prevent doping globally.

Researchers wishing to understand doping behavior and evaluate anti-doping interventions face a big challenge: the measurement of intentional doping. This behavior is very difficult to measure, as essentially, athletes are asked to be honest about breaking the rules of sport. Therefore, actual doping behavior is rarely studied or used as outcome variable for the evaluation of the effectiveness of anti-doping programs. The challenge of accurately measuring doping behavior has led to the development of indirect measures of doping, such as doping likelihood, intention, and susceptibility (e.g., Gucciardi et al., 2010; Kavussanu & Ring, 2017; Lazuras et al., 2015; Lucidi et al., 2008). In this study, our primary outcome is doping likelihood, which is a proxy for doping behavior.

In this research, our aim was to develop and evaluate a moral intervention and determine whether it is more effective than an educational (i.e., knowledge based) intervention in preventing doping in young British and Greek athletes. To this end, in the first phase of this research, we developed the two interventions. In the second phase, we evaluated the moral intervention by examining whether it was effective in reducing doping likelihood. Our secondary outcomes were moral identity, moral disengagement, moral atmosphere, and anticipated guilt. As

these variables have been consistently associated with doping likelihood in previous research (e.g., Boardley et al., 2017; Kavussanu et al., 2016; Kavussanu et al., 2020) an intervention that influences them as well as influencing doping likelihood, should be more effective in preventing actual doping behavior. We hypothesized that the moral intervention would lead to changes in these variables and would be more effective than the educational intervention in reducing doping likelihood, moral disengagement and (pro-doping) moral atmosphere, and increasing moral identity and anticipated guilt. We also expected that these changes would be maintained at three and six months after the completion of the intervention.

A limitation that characterizes the anti-doping literature is the lack of follow-up measures to determine whether the effects of anti-doping interventions are sustained. In a recent evaluation of the United Kingdom (UK) Athletics' anti-doping education program, although doping likelihood was reduced from pre to posttest, this reduction was not maintained at the 3-month follow up (Hurst et al., 2020). Investigating the long-term effectiveness of the intervention is important because ultimately an intervention that is effective only in the short term, is not of much utility in our doping-prevention efforts. We expected that the moral intervention would lead to longer-lasting changes in these variables than the educational intervention. We did not form differential hypotheses for the two countries due to lack of empirical evidence on cross-cultural comparison in doping prevention.

Method

Phase 1

Our main interest was the development and evaluation of the moral intervention. We also developed an educational (i.e., knowledge-based) intervention, which acted as a comparison group, thus allowing us to evaluate the relative effectiveness of the moral intervention. Each intervention consisted of six one-hour sessions, and each session focused on one theme. The two

interventions were identical in duration, mode of delivery, and type of activities. Both included videos and stories of real athletes and their doping-related experiences to facilitate participant engagement and learning (Singler, 2015); they also encouraged group discussion, and allowed participants to engage in role play, and problem solving, which are considered important in anti-doping education (Backhouse et al., 2016). The two interventions differed only in their content: The moral intervention was devoid of any content included in the educational intervention, and the educational intervention was devoid of any content included in the moral intervention.

The two interventions were designed to be delivered by a facilitator to small groups of athletes. Prior to the start of the main trial, each intervention was pilot tested with two groups of student athletes in the UK ($n = 6$) and Greece ($n = 5$), and participant feedback was used to improve and refine each session. During pilot testing, we also evaluated the facilitators' ability to explain concepts, communication, and management of group dynamics.

The Moral Intervention

A detailed description of each session of the moral intervention can be seen in Table 1. Briefly, this intervention included activities and content that targeted moral identity, moral disengagement, and moral atmosphere. In devising these activities, we balanced the need to develop content that was in line with the targeted constructs, with the need for this content to be suitable for our participants (i.e., young athletes). For moral identity[2], the intervention aimed to strengthen the three values of hard work, honesty, and fairness. Being hardworking, honest, and fair, are three of the nine validated traits of a moral person (Aquino & Reed, 2002) and are also relevant to competitive sport. To reduce moral disengagement, we educated participants about the justifications athletes use for doping, presented stories of athletes who doped and used these justifications, asked participants to challenge these justifications, and highlighted the consequences of doping for others (which is relevant to the distortion of consequences

mechanism). Similar approaches have been used in previous interventions aimed to reduce moral disengagement (e.g., Bustamante & Chaux, 2014).

For moral atmosphere, we contrasted a team that plays the "right way" versus a team that cheats, and through discussions with members of the same team, we aimed to influence teams to view doping as unacceptable behavior. This was based on the idea that modelling and the consequences of the modeled behavior can influence one's own behavior (Bandura, 1997). The teams presented were real cases, rather than hypothetical examples, and were used to illustrate the different ways one can play sport and their pros and cons. Participants were asked to reflect on their own experience as a team and consider how they could emulate the team that played the "right way". In addition, we expected that moral atmosphere (which reflects individual team members' perceptions of the team norms regarding acceptable behavior) would be influenced by taking part in the sessions as a group and inferring fellow team members' changes in their views about the use of doping and their subsequent likelihood to use banned substances.

The Educational Intervention

A detailed description of the content of the educational intervention appears in Table 2. In the educational intervention, we used information from a variety of sources, such as the German NADA, UKAD, ASADA, and WADA, and created an intervention that was similar in content to the interventions typically used by these anti-doping organizations. Specifically, we conveyed information about the WADA governance, the rules and regulations, the doping control process, the potential health consequences of banned substances, the risks of supplements, the need for healthy and appropriate nutrition, and whistleblowing.

Phase 2

Research Design

We used a cluster randomized control trial, delivered in the UK and Greece with allocation

ratio 1:1. We collected data at four time points: pre, post, 3-month follow up, 6-month follow up. Thus, this was a mixed factorial design with Intervention (Moral, Educational) and Country (UK, Greece) as between-participant factors and Time (pre, post, 3-month follow up, 6-month follow up) as a within-participant factor. Groups of athletes (rather than individuals) were randomly assigned to the two interventions.

Participants

Participants were identified via a screening survey which included all the measures used in this study and was administered to a large sample of athletes in the two countries (n = 599 in the UK; n = 381 in Greece) to ensure we do not observe floor effects (see Ntoumanis et al., 2014). The criteria for taking part in the study were that: (a) participants were active competitive athletes and (b) clubs had an average score of at least 2 on the item of the first scenario used to measure doping likelihood (described below). This was the first measure included in our screening survey, to ensure participants' responses to this measure were not influenced by their responses to the other measures. A total of 33 clubs[3] and 303 athletes were recruited in the UK (n_{clubs} = 12; $n_{athletes}$ = 121) and Greece (n_{clubs} = 21, $n_{athletes}$ = 182). A detailed description of their characteristics as a function of intervention and country can be seen in Table 3.

Due to the novelty of our moral intervention, the magnitude of the intervention effect size was uncertain. Nevertheless, we expected a large effect size because we targeted variables that have been strongly and consistently associated with doping likelihood in past research (for reviews see Kavussanu, 2019; Ntoumanis et al., 2014). Our view was that if we intervene successfully on these variables, our moral intervention should show a large effect in comparison with a standard educational intervention; we anticipated that the latter would show a small or no change from pre to post intervention based on previous research (Lucidi et al., 2017; Ntoumanis et al., 2014; Sagoe et al., 2016).

Randomization

In each country, clubs whose athletes agreed to take part in the main study were randomly assigned to either the moral or the educational intervention, by a member of the research team. We assigned clubs rather than individual athletes to avoid contamination of the intervention (see Campbell et al., 2012), which would have occurred if athletes from the same club, assigned to different intervention groups, spoke to each other about what is happening in their group. We incorporated stratification by sport type and sex in the randomization process to balance these variables across the two groups. The allocation ratio to the two groups was 1:1. In allocating clubs to the two groups, we used minimization, a process that minimizes the imbalance on important participant characteristics between groups (Moher et al., 2012). In addition to the stratification by sex and sport type, we tried to match the two groups on doping likelihood using participant responses on the screening questionnaire. Minimization has the advantage of making small groups closely similar in terms of participant characteristics, and trials that use minimization are considered methodologically equivalent to randomized trials, even when a random element is not incorporated (Moher et al., 2012).

Measures

Doping Likelihood. We measured doping likelihood using two scenarios adapted from previous research (Kavussanu et al., 2016; Kavussanu & Ring, 2017; Kavussanu et al., 2020). The two scenarios described a situation, where athletes had the opportunity to use a banned substance to enhance their performance (Scenario 1) or speed up recovery from injury (Scenario 2). The performance-enhancement scenario is presented below:

> "It's the week before the most important competitive game/event of your season. Lately, your performance has been below your best. You don't feel you have the necessary fitness for this competition, and you're concerned about how you will

perform. You mention this to a team/club-mate, who tells you that they use a new substance that has enhanced their fitness and performance. The substance is banned for use in sport, but there's no chance that you will be caught."

After participants read each scenario, they were asked to indicate how likely it was that they would use the banned substance, if they were in this hypothetical situation. Responses were made on a Likert scale anchored by 1 (*not at all likely*) and 7 (*very likely*). The average of these two items was used as our doping likelihood variable (Kavussanu et al., 2016; Kavussanu et al., 2020). The approach of computing the average of all relevant items was used for all measures.

Moral Identity. We assessed moral identity using the 5-item internalization subscale of the moral identity scale, which taps the degree to which moral traits are central to individuals' self-concept (Aquino & Reed, 2002). Participants were presented with nine words representing moral traits (i.e., hardworking, honest, fair, helpful, compassionate, caring, friendly, generous, and kind) and were asked to respond to statements concerning these traits (e.g., "It would make me feel good to be a person who has these characteristics"). Responses were made on a 7-point scale anchored by 1 (*strongly disagree*) and 7 (*strongly agree*). The scale has demonstrated very good internal consistency in previous research with an alpha coefficient of .85 (Aquino & Reed, 2002).

Moral Disengagement. We measured moral disengagement in doping with the Moral Disengagement in Doping Scale (Kavussanu et al., 2016). Participants were asked to read six statements and indicate their level of agreement using a Likert scale anchored by 1 (*strongly disagree)* and 7 (*strongly agree)*. Example items are "Doping does not really hurt anyone" and "Players/athletes cannot be blamed for doping if their teammates pressure them to do it". The scale has shown very good levels of internal consistency (α range = .82 - .86) as well as factorial, convergent, concurrent, and discriminant validity (Kavussanu et al., 2016).

Moral Atmosphere. The team moral atmosphere condoning doping was measured with six items developed specifically for this study, based on previous research that has measured moral atmosphere or team norms in sport (Kavussanu & Ntoumanis, 2003; Kavussanu & Spray, 2006; Ommundsen et al., 2003). Participants were asked to imagine that there was an opportunity to use an undetectable banned substance to significantly enhance their performance in a very important competition. Then, they responded to six items indicating "In this hypothetical situation…" how many of their teammates would "approve the use of the substance", "be tempted to use it", "favor its use" etc. Responses were made on a 7-point Likert scale with anchors of 1 (*no-one*) and 7 (*everyone*). The items were in line with the definition of the construct of moral atmosphere as the type of behavior that is considered acceptable in a group by its group members. As measured in this study, moral atmosphere refers to athletes' perceived moral atmosphere. A similar approach has been used in studies that have measured perceived antisocial practice norms (e.g., Benson et al., 2017). Results of Confirmatory Factor Analysis (CFA) of the moral atmosphere measure at the three time points can be seen in the supplementary material (Table S1).

Anticipated Guilt. We measured anticipated guilt with the five-item guilt subscale from the State Shame and Guilt Scale (Marschall et al., 1994). Participants were asked to imagine that they had used a banned substance to significantly enhance their performance in a very important competition and indicate how they think they would feel. The stem for each item was "If I had used a banned substance…" and sample items are "I would feel remorse, regret" and "I would feel bad about what I had done". Participants indicated their responses on a Likert scale anchored by 1 (*not at all*) and 7 (*very strongly*). Marschall et al. (1994) reported very good internal consistency for this measure ($\alpha = .82$).

Procedure

Unless otherwise stated, the same procedure was used in Greece and UK. First, the study was approved by the first author's University Research Ethics committee. Then coaches of eligible clubs were contacted and asked for their assistance in recruiting their athletes for the study, and participants were informed about the study requirements. They were given broad information about the content of the intervention and were told that all information obtained would be anonymous and used only for research purposes.

In each country, the two interventions were delivered by the same trained facilitator over a period of six-to-eight weeks. The two facilitators had experience with teaching small classes of sport science university students and seminars/interventions in young athletes (Greece) or delivering psychological interventions in small groups of prisoners (UK). Prior to the start of the delivery, to ensure the facilitators were competent in delivering the interventions, the material was discussed with them, and they were observed delivering some sessions in small groups of student athletes.

Data were collected using a questionnaire, which included all measures detailed above. Participants completed the questionnaire at four time points at club venues in the UK and in a classroom setting provided by the local authority in Greece, or via an online link if they were not present at their club during our visit. The questionnaire was completed before the start of the first session (preintervention), immediately after the last session (postintervention), and three and six months after the sixth session. Recruitment was conducted over a 12-month period. There were no unintended consequences or harms. Questionnaires at the four time points were linked to each other by a password, which was provided by each participant using information known only to them. This ensured participant anonymity, thus minimizing socially desirable responses.

Data Analysis

Data analyses were conducted in Mplus, version 8.4 (Muthén & Muthén, 2012-2019), using multilevel piecewise linear growth modeling to test changes in study outcomes (i.e., doping likelihood, moral identity, moral disengagement, moral atmosphere, and anticipated guilt), with measurements nested in athletes and athletes nested in clubs. Specifically, the four measurements were modeled in multivariate fashion, and athletes were modeled to be nested within clubs to account for the nonindependence of observations (i.e., three-level analysis, Muthén & Muthen, 1998-2017, p. 303) with Bayesian estimation method. Because we expected to see changes in outcomes immediately after the intervention (from pre to post) and no changes thereafter (from post to 6-month follow up), we specified a piecewise growth model with three slopes. Slope 1 tested changes in outcomes from preintervention to postintervention, slope 2 tested changes from postintervention to the 3-month follow up, and slope 3 tested changes from the 3-month follow up to the 6-month follow up. These were random intercepts and random slopes models at both the athlete and club levels.

In the models, we entered condition (0 = educational, 1 = moral), country (0 = UK, 1 = Greece), and country by condition as predictors of the intercepts (which reflect differences in baseline scores) and of the three slopes (which reflect changes in outcomes). Additionally, because sex differences in doping likelihood have emerged in previous research (e.g., Ntoumanis et al., 2014), we included sex (0 = male, 1 = female) as a factor in our models to control for these differences. Following CONSORT 2010 guidelines related to tests of baseline differences (Schulz et al., 2010), we do not present or discuss the baseline differences because they were not of interest; however, full results for the models are available in the supplementary materials.

We estimated the models under missing data theory (Rubin, 1976) using all available data, and our analyses included all participants who underwent randomization, which reflect the intention to treat (ITT) analysis recommendations for treatment of missing data in RCTs (Moher

et al., 2012, p. 14). In addition, we conducted sensitivity analysis using responses only from those athletes who attended all six sessions and completed all assessments points (i.e., listwise deletion), which reflects the per protocol (PP) analysis. Changes in study outcomes were determined using Bayesian credibility intervals and effect sizes are presented in baseline standard-deviation units (Feingold, 2009). Values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and .80 or greater a large effect.

Results

Figure 1 shows the flow of the progress through the phases of the trial. Retention rates of athletes were good at the end of the 6-week interventions (90%), 3-month follow up (76%), and 6-month follow up (78%). Missing data out of 303 cases were as follows: 0.32% at preintervention, 10.44% at postintervention, 23.67% at the 3-month follow up, and 21.80% at the 6-month follow up. Adherence to the protocol was also satisfactory with most participants having attended six sessions (71.9%). Overall, more Greek than British athletes (64.3% versus 46.3%, adjusted residual = 3.1) and more female than male athletes (69.6% versus 50.7%, adjusted residual = 3.1) completed the study per protocol.

Reliability estimates for all scales can be seen in Table 4. This table also shows descriptive statistics for all outcomes by country and intervention group at the four time points. The results of pre-post changes in multilevel piecewise linear growth models for all outcomes are presented in Table 5. Table 6 shows the results for pre-post changes in all outcomes by country and intervention, adjusted for the variables in the models (i.e., simple slopes analyses). Detailed results for multilevel piecewise linear growth models for each outcome can be found in Tables S2 – S6 in supplementary material. These tables include the results from post to 3-month follow up and from 3- to 6-month follow up, which address the long-term effects of our interventions; they also include sex, which was the control variable. In order to aid understanding of how the two

intervention groups changed across the four time points, we have also presented the means of all outcomes as a function of time and intervention in Figure 2. Below we discuss the results of the ITT and PP analyses as they pertain to the intervention effects on our primary and secondary outcomes. An effect was considered significant if the credibility interval did not include zero.

Primary Outcome

As can be seen in Table 5, ITT analysis showed that athletes reported lower *doping likelihood* at postintervention compared to preintervention, Δ*M* = -0.65, 95% CI [-1.19, -0.13], *SE* = 0.27, ES = -.46. No changes emerged from post to 3-month follow up or from 3- to 6-month follow up (Table S2) indicating that both interventions were effective in maintaining the reduction in doping likelihood over six months. These effects were similar across the two countries (Table 6).

PP analysis (Table 5) also revealed a Time main effect for doping likelihood, but this was superseded by an intervention by country interaction of a large magnitude, Δ*M* = -1.31, 95% CI [-2.38, -0.26], *SE* = 0.54, ES = -.94. Tests of simple slopes (Table 6) showed that, holding other variables constant, in the UK, athletes in the moral intervention showed no change in their doping likelihood, Δ*M* = -0.20, 95% CI [-0.83, 0.43], *SE* = 0.32, ES = -.14, but athletes in the educational intervention reported lower doping likelihood at postintervention, Δ*M* = -0.91, 95% CI [-1.60, -0.28], *SE* = 0.33, ES = -.66. However, in Greece, athletes in the moral intervention reported lower doping likelihood, Δ*M* = -0.94, 95% CI [-1.39, -0.49], *SE* = 0.23, ES = -.68, whereas athletes in the educational intervention did not change their doping likelihood from pre to post, Δ*M* = -0.35, 95% CI [-0.82, 0.11], *SE* = 0.24, ES = -.25. These results are illustrated in Figure 3.

Secondary Outcomes

The results for our secondary outcomes are presented in Tables 5 and 6 and S3 to S6. As can be seen in Table 5, we found no intervention effects for *moral identity*, from pre to post, based on the credibility intervals in either the ITT or the PP analyses. There were also no changes from post to 3-month follow up or from 3- to 6-month follow up (Table S3). For *moral disengagement*, ITT analysis showed that athletes reported lower moral disengagement at postintervention compared to preintervention, ΔM = -0.51, 95% CI [-0.84, -0.18], SE = 0.17, ES = -.56. No changes emerged from post to 3-month follow up or from 3- to 6-month follow up (Table S4) indicating that the two interventions had long-term effects on moral disengagement. These effects were similar across the two interventions and the two countries and were confirmed in PP analysis. No intervention effects emerged for *moral atmosphere* based on the credibility intervals in the ITT or PP analyses. There were also no changes from post to 3-month follow up or from 3- to 6-month follow up, and these effects were similar in the two interventions and the two countries (Table S5).

ITT analysis showed that both interventions increased *anticipated guilt*, from pre to post, ΔM = 0.49, 95% CI [0.03, 0.98], SE = 0.24, ES = .35 (Table 5). No changes emerged from post to 3-month follow up or from 3- to 6-month follow up, and these effects were similar across the two interventions and the two countries (Table S6). In PP analysis, even though the effect size was small-to-moderate for pre-post changes in anticipated guilt, the credibility interval did include zero (ΔM = 0.43, 95% CI [-0.26, 1.13], SE = 0.35, ES = .31).

Discussion

Doping is a behavior that can have significant adverse health consequences for the user (e.g., Quaglio et al., 2009). Despite considerable investment in doping prevention by governments and international organizations worldwide, recent reports suggest that doping is widespread (Ulrich et al., 2018). Therefore, developing interventions that prevent doping is an

important endeavor. In this research, we developed and evaluated an intervention that targeted variables, which have been empirically associated with doping likelihood (e.g., Boardley et al., 2017; Kavussanu, 2019; Kavussanu et al., 2016; Kavussanu & Ring, 2017; Ntoumanis et al., 2014). We also developed an educational intervention, which acted as our comparison group.

In line with the CONSORT recommendations (Schulz et al., 2010) we conducted two sets of analyses. Our primary analysis was ITT, where we included all participants, regardless of the number of sessions they attended and the number of measurement points they completed. Our secondary analysis was PP, in which we included only those participants who attended all six sessions of each intervention and completed measures at all time points. Below we consider the results of both sets of analyses, and discuss our findings as they pertain to our primary and secondary outcomes.

Primary Outcome

Our ITT analysis showed that both interventions, in both countries, were effective in reducing doping likelihood from pre to postintervention. Importantly, in both countries, no significant changes emerged from postintervention to the 6-month follow up indicating that both interventions were effective in producing sustained effects in doping likelihood. Our study is the first to show that an intervention that targets moral variables can significantly reduce doping likelihood not only in the short but also in the long term. This highlights the importance of integrating moral variables in the development of doping prevention programs. A recent study evaluating the UK Athletics anti-doping program, a one-off session with content that is primarily knowledge based, showed that although doping likelihood was reduced immediately following the program (d = .20), these changes were not maintained three months later (Hurst et al., 2020). Nicholls et al. (2020) have recently shown sustained effects of a face-to-face anti-doping intervention on doping susceptibility, however, to our knowledge, ours is the first study to show

sustained reduction in doping likelihood over a 6-month period, after completion of the program in two countries.

Our PP analysis revealed an interaction between country and intervention, such that the educational intervention in the UK, and the moral intervention in Greece, were effective in reducing doping likelihood from pre to post. However, the moral intervention in the UK and the educational intervention in Greece did not show a significant reduction in doping likelihood. Inspection of the adjusted means (see Figure 3) suggests that, in the UK sample, doping likelihood at baseline was lower in the moral compared to the educational intervention, whereas the opposite pattern was observed in the Greek sample. This may explain the lack of a significant reduction in doping likelihood in the moral intervention in the UK and in the educational intervention in Greece, revealed in our PP analysis.

Secondary Outcomes

Our primary (ITT) analysis showed that athletes in both interventions and both countries reported lower moral disengagement and higher anticipated guilt from pre to postintervention, and these changes were maintained at the 3- and 6-month follow up. Both of these variables have been strongly linked to doping likelihood or susceptibility in several studies (e.g., Boardley et al., 2017; Kavussanu et al., 2016; Kavussanu & Ring, 2017; Kavussanu et al., 2020; Ntoumanis et al., 2014) in the opposite direction, that is moral disengagement has been positively, and anticipated guilt has been negatively, associated with doping likelihood. That our interventions influenced these two variables in the expected direction is an important finding that attests to the effectiveness of not only the moral but also our educational intervention. That the two interventions produced sustained changes in these variables across the two countries suggests that they contained elements that cut across cultures and are relevant to athletes from different countries.

Changes in moral disengagement and anticipated guilt were hypothesized only for the moral intervention. There are several explanations for the observed changes in these variables in the educational intervention. First, elements contained in the latter intervention may have unintentionally influenced these variables. For example, the topic of the last session of the educational intervention was whistleblowing, which represents a moral dilemma (see Erickson et al., 2019). By listening to stories of athletes who chose to "blow the whistle" because this was "the right thing to do", participants may have received the message that "doping is not right", thus triggering feelings of anticipated guilt - a moral emotion that is elicited when people engage in a transgression (Tangney et al., 2007) - in the hypothetical situation of using a banned substance. Second, by discussing doping-related issues over six sessions in the educational intervention and learning about the doping sanctions and risks to health, participants may have been sensitized to the "doping problem" and this could have led to a reduction in moral disengagement and an increase in anticipated guilt. It would be interesting for future research evaluating doping-prevention programs to examine the effects of these programs on moral disengagement and anticipated guilt.

It is also worth noting that doping moral disengagement has been positively and strongly associated with doping attitudes in past research (e.g., Kavussanu et al., 2016). It is possible that there is some overlap between the two constructs. Doping attitudes have been shown to change in anti-doping interventions (e.g., Nicholls et al., 2020), which do not specifically target moral disengagement. It may be that typical interventions that focus on delivering information about doping, have the potential to change not just doping attitudes but also moral disengagement, which is strongly associated with doping attitudes (Kavussanu et al., 2016). Future research should attempt to shed light on this issue by measuring moral disengagement in interventions that do not target this construct.

Results of the PP analysis confirmed the ITT findings for moral disengagement, thus attesting to the robustness of the effect of the two interventions on this variable. However, the PP analysis did not reveal significant effects on anticipated guilt from pre to postintervention. This may be explained by the substantial reduction in our sample size from the ITT ($N = 303$) to the PP ($n = 172$) analysis.

We found no effects on moral identity from pre to postintervention, or during the 6-month follow-up period in either set of analyses. Thus, contrary to our hypotheses, the moral intervention did not influence moral identity. Some tentative explanations could be offered for this null finding. First, moral identity may need a much longer intervention to change. As conceptualized in this research, this construct refers to the importance individuals place on being moral or the centrality of being a moral person in one's self-concept (Aquino & Reed, 2002). In our intervention, in two of our six sessions, we focused on three of the nine traits which have been identified as essential characteristics of a moral person: hardworking, fair, and honest. Perhaps an intervention that focuses on the other traits of moral identity over more sessions may be needed for this variable to change. Although previous studies have successfully primed moral identity and observed acute effects in experimental settings (e.g., Aquino et al., 2009), ours was the first study that attempted to produce longer-lasting changes in this variable.

Contrary to our hypotheses, the moral intervention had no effect on moral atmosphere. One explanation for this finding is that the part of the intervention that focused on moral atmosphere was not sufficiently strong to change this construct. The intervention contrasted teams who valued winning at all costs with teams who played "the right way" (i.e., respecting others, playing fair), and participants were asked to identify ways that they, as a team, can take part in sport "the right way". Thus, participants engaged critically in discussions and activities to reflect on and revise the norms within their group. Perhaps the importance of playing fair as a team

should have been further emphasized. A second explanation is that our measure did not capture group changes. In line with past sport research (e.g., Kavussanu & Ntoumanis, 2003; Ommundsen et al., 2003) moral atmosphere was measured as individual athletes' perceptions of the acceptability of doping within the team, as manifested in their interactions within the group. Our participants may not have had sufficient information to accurately assess team norms, based on their interaction with their teammates over the course of the intervention. Thus, our measure may not have been sensitive enough to capture changes in the moral atmosphere of the team.

Limitations of the Study and Directions for Future Research

Our research has some limitations and our findings need to be interpreted in light of these limitations. First, we did not have a no-intervention control group, thus we do not know whether the observed changes in our outcomes would have naturally occurred over time. Our comparison group also received an intervention, which was also effective in producing changes in our outcomes. However, that doping likelihood, moral disengagement and anticipated guilt remained reduced at the 3- and 6-month follow up and these effects were consistent across the two countries, attests to the overall effectiveness of our interventions. Nevertheless, future research should replicate the present work employing a no-intervention control group. Second, due to the clustered nature of our data, our study was powered to detect only a large effect. We used this design (i.e., assigning clubs rather than participants to the intervention) because by delivering the intervention in the clubs we avoided contamination of the intervention and minimized our dropout rate. Future research using a similar design should employ a larger number of teams, or deliver the intervention in groups of athletes from different clubs.

Third, as is typical of this type of research, we had missing data that ranged from .32% to 23.67% across the four time points. Missing data occurred because some participants were not present on the day of the visit, or they had left the club. However, our retention rates are much

higher than those reported in similar studies. For example, Nicholls et al. (2020) reported completion rates of 27% to 69% across three time points. Fourth, in line with previous research assessing team norms (e.g., Benson et al., 2017; Ommundsen et al., 2003), we measured moral atmosphere as an individual- rather than a group-level via athlete perceptions. However, moral atmosphere is best conceptualized as an emergent concept that is the result of interactions among individuals within groups. Kozlowski et al. (2013) have provided a set of recommendations for advancing research on emergent phenomena, which should be considered in future research, in light of the importance of group processes in sport teams. Finally, we did not assess the protective factors of the educational intervention, for example knowledge of the health risks of doping substances or understanding of the anti-doping rule violations. Researchers seeking to replicate this work should include such measures (see Hurst et al., 2020).

Conclusion

In conclusion, this is the first research to develop and evaluate an intervention targeting moral variables and compare it with an educational intervention of equal duration. Our primary analysis showed that both interventions were effective in reducing doping likelihood and moral disengagement in two countries, and these effects were sustained six months after their completion. The findings, which can be generalized to young British and Greek athletes, suggest that alongside their typical content (e.g., providing information about the harms of banned substances, doping control process, anti-doping rule violations) anti-doping education programs should consider targeting moral variables. Ultimately, the aim of these programs is to produce long-term changes in the likelihood of athletes to use prohibited substances. Incorporating elements included in both of our interventions (moral and educational) should facilitate this goal.

Endnotes

[1]Studies have used various proxies of doping behavior, for example, doping intention, likelihood, and susceptibility. For simplicity, we use the term doping likelihood when we discuss studies measuring doping intention and susceptibility, given the large overlap among these measures.

[2]For the intervention to be effective, it was important to develop content that was not only in line with the targeted constructs but also relevant to the participants' experiences. Linking the value of hard work to the feeling of success when athletes reach their athletic potential *through hard work* was considered the best way to balance these two needs. To this end, some elements of Session 1 were inspired by achievement goal theory, while ensuring that the content reflected the "hardworking" characteristic, which is one of the nine validated characteristics of the moral person, used to measure moral identity (Aquino & Reed, 2002).

[3]In the UK, in addition to clubs, some sport colleges took part in the study. However, for simplicity we refer to clubs. All participants in these colleges were competitive athletes. The results of power analysis – which was conducted post hoc - using Optimal Design Software for cluster RCT with person-level outcomes and repeated measures showed that with 33 clusters, an average cluster size of 10, and intraclass cluster correlation of .10, we had power of .80 to detect a minimum detectable effect size (MDES) of .45 (that is, a mean difference equivalent to .45 in units of the population standard deviation of the outcome). Bloom (1995) defines the MDES as "the smallest true effect that can be detected for a specified level of power and significance level for any given sample size."

References

Aquino, K., Freeman, D., Reed II, A., Lim, V. K., & Felps, W. (2009). Testing a social-cognitive model of moral behavior: the interactive influence of situations and moral identity centrality. *Journal of Personality and Social Psychology*, *97*(1), 123-141. https://doi.org/10.1037/a0015406

Aquino, K., & Reed, I. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, *83*(6), 1423-1440. https://doi.org/10.1037/0022-3514.83.6.1423

Backhouse, S., Whitaker, L., Patterson, L., Erickson, K., & McKenna, J. (2016). *Social psychology of doping in sport: A mixed studies narrative synthesis*. Montreal: World Anti-Doping Agency.

Bandura, A. (1995). Modeling. In A. S. R. Manstead & M. Hewstone (Eds.), *Blackwell encyclopedia of social psychology*. Oxford: Blackwell.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, *3*(3), 193-209. https://doi.org/10.1207/s15327957pspr0303_3

Barkoukis, V., Kartali, K., Lazuras, L., & Tsorbatzoudis, H. (2016). Evaluation of an anti-doping intervention for adolescents: Findings from a school-based study. *Sport Management Review*, *19*(1), 23-34. https://doi.org/10.1016/j.smr.2015.12.003

Benson, A. J., Bruner, M. W., & Eys, M. (2017). A social identity approach to understanding the conditions associated with antisocial behaviors among teammates in female teams. *Sport, Exercise, and Performance Psychology*, *6*(2), 129-142. https://doi.org/10.1037/spy0000090

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review*, *19*(5), 547-556. https://doi.org/10.1177/0193841X9501900504

Boardley, I. D., Smith, A. L., Mills, J. P., Grix, J., & Wynne, C. (2017). Empathic and self-regulatory processes governing doping behavior. *Frontiers in Psychology*, *8*, 1495. https://doi.org/10.3389/fpsyg.2017.01495

Bustamante, A., & Chaux, E. (2014). Reducing moral disengagement mechanisms: A comparison of two interventions. *Journal of Latino/Latin American Studies*, *6*(1), 52-63.

Campbell, M. K., Piaggio, G., Elbourne, D. R., & Altman, D. G. (2012). CONSORT 2010 statement: extension to cluster randomised trials. *British Medical Journal, 328,* 702-708. doi: https://doi.org/10.1136/bmj.e5661

de Hon, O., Kuipers, H., & van Bottenburg, M. (2015). Prevalence of doping use in elite sports: a review of numbers and methods. *Sports Medicine*, *45*(1), 57-69. https://doi.org/10.1007/s40279-014-0247-x

Elbe, A.-M., & Brand, R. (2016). The effect of an ethical decision-making training on young athletes' attitudes toward doping. *Ethics & Behavior*, *26*(1), 32-44. https://doi.org/10.1080/10508422.2014.976864

Elliot, D. L., Goldberg, L., Moe, E. L., Defrancesco, C. A., Durham, M. B., & Hix-Small, H. (2004). Preventing substance use and disordered eating: initial outcomes of the ATHENA (athletes targeting healthy exercise and nutrition alternatives) program. *Archives of Pediatrics and Adolescent Medicine*, *158*(11), 1043-1049. https://doi.org/10.1001/archpedi.158.11.1043

Elliot, D. L., Goldberg, L., Moe, E. L., DeFrancesco, C. A., Durham, M. B., McGinnis, W., & Lockwood, C. (2008). Long-term outcomes of the ATHENA (Athletes Targeting Healthy Exercise & Nutrition Alternatives) program for female high school athletes. *Journal of Alcohol and Drug Education*, *52*(2), 73-92.

Erickson, K., Patterson, L. B., & Backhouse, S. H. (2019). "The process isn'ta case of report it and stop": Athletes' lived experience of whistleblowing on doping in sport. *Sport Management Review*, *22*(5), 724-735. https://doi.org/10.1016/j.smr.2018.12.001

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods*, *14*(1), 43-53. https://doi.org/10.1037/a0014699

Goldberg, L., Elliot, D., Clarke, G. N., MacKinnon, D. P., Moe, E., Zoref, L., Green, C., Wolf, S. L., Greffrath, E., & Miller, D. J. (1996). Effects of a multidimensional anabolic steroid prevention intervention: The Adolescents Training and Learning to Avoid Steroids (ATLAS) Program. *JAMA*, *276*(19), 1555-1562.

Goldberg, L., & Elliot, D. L. (2005). Preventing substance use among high school athletes: The ATLAS and ATHENA programs. *Journal of Applied School Psychology*, *21*(2), 63-87. https://doi.org/10.1300/J370v21n02_05

Goldberg, L., Elliot, D. L., MacKinnon, D. P., Moe, E., Kuehl, K. S., Nohre, L., & Lockwood, C. M. (2003). Drug testing athletes to prevent substance abuse: background and pilot study results of the SATURN (Student Athlete Testing Using Random Notification) study. *Journal of Adolescent Health*, *32*(1), 16-25. https://doi.org/10.1016/S1054-139X(02)00444-5

Goldberg, L., MacKinnon, D. P., Elliot, D. L., Moe, E. L., Clarke, G., & Cheong, J. (2000). The adolescents training and learning to avoid steroids program: preventing drug use and promoting health behaviors. *Archives of Pediatrics and Adolescent Medicine*, *154*(4), 332-338. doi:10.1001/archpedi.154.4.332

Gucciardi, D. F., Jalleh, G., & Donovan, R. J. (2010). Does social desirability influence the relationship between doping attitudes and doping susceptibility in athletes? *Psychology of Sport and Exercise*, *11*(6), 479-486. https://doi.org/10.1016/j.psychsport.2010.06.002

Hurst, P., Ring, C., & Kavussanu, M. (2020). Evaluation of UK Athletics Clean Sport programme in preventing doping in junior elite athletes. *Performance Enhancement & Health*. https://doi.org/10.1016/j.peh.2019.100155

Kavussanu, M. (2019). Toward an understanding of transgressive behavior in sport: Progress and prospects. *Psychology of Sport and Exercise*, *42*, 33-39. https://doi.org/10.1016/j.psychsport.2019.01.009

Kavussanu, M., Hatzigeorgiadis, A., Elbe, A.-M., & Ring, C. (2016). The moral disengagement in doping scale. *Psychology of Sport and Exercise*, *24*, 188-198. https://doi.org/10.1016/j.psychsport.2016.02.003

Kavussanu, M., & Ntoumanis, N. (2003). Participation in sport and moral functioning: Does ego orientation mediate their relationship? *Journal of Sport & Exercise Psychology*, *25*(4), 501-518. (https://doi.org/10.1123/jsep.25.4.501)

Kavussanu, M., & Ring, C. (2017). Moral identity predicts doping likelihood via moral disengagement and anticipated guilt. *Journal of Sport & Exercise Psychology*, *39*(4), 293-301. https://doi.org/10.1123/jsep.2016-0333

Kavussanu, M., Roberts, G. C., & Ntoumanis, N. (2002). Contextual influences on moral functioning of college basketball players. *The Sport Psychologist*, *16*(4), 347-367. https://doi.org/https://doi.org/10.1123/tsp.16.4.347

Kavussanu, M., & Spray, C. M. (2006). Contextual influences on moral functioning of male youth footballers. *The Sport Psychologist*, *20*(1), 1-23. DOI: https://doi.org/10.1123/tsp.16.4.347

Kavussanu, M., Yukhymenko-Lescroart, M. A., Elbe, A.-M., & Hatzigeorgiadis, A. (2020). Integrating moral and achievement variables to predict doping likelihood in football: A cross-cultural investigation. *Psychology of Sport and Exercise*, 101518. https://doi.org/10.1016/j.psychsport.2019.04.008

Kozlowski, S. W., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2013). Advancing multilevel research design: Capturing the dynamics of emergence. *Organizational Research Methods*, *16*(4), 581-615. https://doi.org/10.1177/1094428113493119

Lazuras, L., Barkoukis, V., & Tsorbatzoudis, H. (2015). Toward an integrative model of doping use: an empirical study with adolescent athletes. *Journal of Sport & Exercise Psychology*, *37*(1), 37-50. https://doi.org/10.1123/jsep.2013-0232

Lucidi, F., Mallia, L., Alivernini, F., Chirico, A., Manganelli, S., Galli, F., Biasi, V., & Zelli, A. (2017). The Effectiveness of a New School-Based Media Literacy Intervention on Adolescents' Doping Attitudes and Supplements Use. *Frontiers in Psychology*, *8*, 749. https://doi.org/10.3389/fpsyg.2017.00749

Lucidi, F., Zelli, A., Mallia, L., Grano, C., Russo, P. M., & Violani, C. (2008). The social-cognitive mechanisms regulating adolescents' use of doping substances. *Journal of Sports Sciences*, *26*(5), 447-456. https://doi.org/10.1080/02640410701579370

Marschall, D., Sanftner, J., & Tangney, J. P. (1994). The state shame and guilt scale. *Fairfax, VA: George Mason University*.

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P., Elbourne, D., Egger, M., & Altman, D. G. (2012). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *International Journal of Surgery*, *10*(1), 28-55. https://doi.org/ https://doi.org/10.1136/bmj.c869

Muthén, L., & Muthén, B. (2012-2019). Mplus (Version 7.11) Muthén & Muthén. *Los Angeles*.

Muthén, L. K., & Muthen, B. (1998-2017). *Mplus user's guide: Statistical analysis with latent variables, user's guide*. Muthén & Muthén.

Nicholls, A. R., Morley, D., Thompson, M. A., Huang, C., Abt, G., Rothwell, M., Cope, E., & Ntoumanis, N. (2020). The effects of the iPlayClean education programme on doping attitudes and susceptibility to use banned substances among high-level adolescent athletes from the UK: A cluster-randomised controlled trial. *International Journal of Drug Policy*, *82*, 102820. https://doi.org/10.1016/j.drugpo.2020.102820

Ntoumanis, N., Barkoukis, V., Gucciardi, D. F., & Chan, D. K. C. (2017). Linking coach interpersonal style with athlete doping intentions and doping use: A prospective study. *Journal of Sport & Exercise Psychology*, *39*(3), 188-198.

DOI: https://doi.org/10.1123/jsep.2016-0243

Ntoumanis, N., Ng, J. Y., Barkoukis, V., & Backhouse, S. (2014). Personal and psychosocial predictors of doping use in physical activity settings: a meta-analysis. *Sports Medicine*, *44*(11), 1603-1624. https://doi.org/10.1007/s40279-014-0240-4

Ommundsen, Y., Roberts, G., Lemyre, P., & Treasure, D. (2003). Perceived motivational climate in male youth soccer: Relations to social–moral functioning, sportspersonship and team norm perceptions. *Psychology of Sport and Exercise*, *4*(4), 397-413. https://doi.org/10.1016/S1469-0292(02)00038-9

Quaglio, G., Fornasiero, A., Mezzelani, P., Moreschini, S., Lugoboni, F., & Lechi, A. (2009). Anabolic steroids: dependence and complications of chronic use. *Internal and Emergency Medicine*, *4*(4), 289-296. DOI 10.1007/s11739-009-0260-5

Ranby, K. W., Aiken, L. S., MacKinnon, D. P., Elliot, D. L., Moe, E. L., McGinnis, W., & Goldberg, L. (2009). A mediation analysis of the ATHENA intervention for female athletes: prevention of athletic-enhancing substance use and unhealthy weight loss

behaviors. *Journal of Pediatric Psychology*, *34*(10), 1069-1083. https://doi.org/10.1093/jpepsy/jsp025

Ring, C., Kavussanu, M., Lucidi, S., & Hurst, P. (2019). Effects of personal and situational factors on self-referenced doping likelihood. *Psychology of Sport and Exercise*, *41*, 29-35. https://doi.org/10.1016/j.psychsport.2018.11.003

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581-592. https://doi.org/10.1093/biomet/63.3.581

Sagoe, D., Holden, G., Rise, E. N. K., Torgersen, T., Paulsen, G., Krosshaug, T., Lauritzen, F., & Pallesen, S. (2016). Doping prevention through anti-doping education and practical strength training: The Hercules program. *Performance Enhancement & Health*, *5*(1), 24-30. https://doi.org/10.1016/j.peh.2016.01.001

Schulz, K. F., Altman, D. G., Moher, D., & Group, C. (2010). CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Trials*, *11*(1), 32.

Singler, A. (2015). Why education of athletes is not enough. In V. Moller, I. Waddington, & J. Hoberman (Eds.), *Routledge Handbook of Drugs and Sport*. Routledge.

Spruit, A., Kavussanu, M., Smit, T., & IJntema, M. (2019). The relationship between moral climate of sports and the moral behavior of young athletes: A multilevel meta-analysis. *Journal of Youth and Adolescence*, *48*(2), 228-242. https://doi.org/10.1007/s10964-018-0968-5

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*, 345-372.

Ulrich, R., Pope, H. G., Cléret, L., Petróczi, A., Nepusz, T., Schaffer, J., Kanayama, G., Comstock, R. D., & Simon, P. (2018). Doping in two elite athletics competitions assessed

by randomized-response surveys. *Sports Medicine*, *48*(1), 211-219. https://doi.org/10.1007/s40279-017-0765-4

World Anti-Doping Agency. (2019). *Annual Report 2018.* Montreal, Canada. World Anti-Doping Agency.

World Anti-Doping Code (2015). *World Anti-Doping Code 2015*. Montreal, Canada, World Anti-Doping Agency.

Table 1

Title, Aims, and Content of the Six Sessions of the Moral Intervention

Title	Aim	Content
Success in sport (moral identity)	To promote the value of hard work and an appreciation of feelings of success when athletes develop their talent and reach their athletic potential	Participants are presented with two approaches to success: (a) winning at all costs and (b) being the best you can be. They are asked to compare the pros and cons of each approach. The value of working hard (moral identity trait) is emphasized. The link between the be the best you can be and the Spirit of Sport is made, highlighting the values of excellence in performance, dedication, and commitment.
Values in sport (moral identity)	To highlight the importance of honesty and fair play in sport and discuss the role of doping in undermining these values	Cases of elite athletes displaying honesty and fairness (two moral identity traits) at the cost of their own victory are discussed, and contrasted with athletes who dope. Participants reflect on the importance of fair play. A debate of whether "doping is cheating" illustrates the fact that doping undermines fair play.
Justifications for doping (moral disengagement)	To make participants aware of the justifications athletes use for doping and to challenge these justifications	The justifications athletes use for doping are presented (e.g., comparing doping with worse acts, blaming others for our behavior, etc) using real-athlete stories. Participants are asked to identify justifications in athletes who have doped, challenge these justifications, and reflect on their own experiences of justifying transgressive behavior.
Consequences of doping for others (moral disengagement)	To highlight the consequences of doping for others, thus challenging the distortion of consequences mechanism of moral disengagement	The devastating consequences of doping for others are discussed. Stories of athletes who have been awarded medals retrospectively (e.g., Kelly Sotherton, Adam Nelson, Valerie Adams) are presented, drawing attention to the emotions they experience. Consequences for one's teammates and family are also discussed.
The culture of the team (moral atmosphere)	To get participants to understand that taking part in sport as a team "the right way" is the best way forward.	Two types of teams are presented and contrasted using real life examples: teams who value winning at all costs and teams who play "the right way" by respecting others and forfeiting winning for the sake of fair play. The different team cultures are discussed and participants are asked to identify ways they can play sport "the right way".
Course conclusion	To summarize the main points of each session	Participants are asked to present the main points discussed in the previous five sessions and to come up with a slogan that accurately represents the content of each session. The facilitator also summarizes the main points.

Note. The moral intervention targeted three variables, which have been associated with doping likelihood in previous research: moral identity (-), moral disengagement (+), and moral atmosphere of the team (+). Each session focused on one of these variables.

Table 2

Title, Aims, and Content of the Six Sessions of the Educational (Knowledge-Based) Intervention

Title	Aim	Content
Introduction to doping	To introduce participants to the WADA and its role in regulating doping in sport	The role of WADA as an international organization that regulates doping is discussed. The WADA Code is presented and the 10 anti-doping rule violations are explained giving examples of athletes who have violated these rules.
Doping control	To introduce participants to the doping control process	Participants are informed that they can be drug tested at any time and place. The anti-doping drug-testing procedure is explained and participants role play each step of the procedure using official anti-doping bottles and documentation. They are also informed about ADAMS and the athlete biological passport.
Banned substances	To introduce banned substances and the consequences they can have on athletes' health	The risks associated with the most common types of banned performance-enhancing substances (e.g., anabolic steroids, stimulants, erythropoietin) are explained, and participants watch a video of the East German shot putter, Heidi Krieger. They are also introduced to the process of obtaining a Therapeutic Use Exemption (TUE).
Sport supplements	To inform participants of the risks associated with sport supplements (e.g., protein, energy drinks, creatine)	Participants are informed about the potential contamination of sport supplements with banned substances and are instructed to check sport supplements using the Informed-Sport website. Cases of athletes failing a drug test due to contamination of sport supplements are presented. Participants are also asked to assess the need of sport supplements and consider if the benefits are the result of a placebo effect.
Nutrition	To discuss the role of nutrition and its benefits for performance and recovery	Information about carbohydrates, proteins, fats, vitamins, and minerals is presented and how to use these pre, during, and post competition is discussed. Participants are asked to examine their own nutrition using the MyFitnessPal app and identify the areas of their diet that could be improved.
Whistleblowing	To discuss the importance of whistleblowing in protecting clean athletes	Whistleblowing is explained, and examples of athletes who blew the whistle (e.g., Yuliya Stepanova) are presented. Participants are informed how to use the WADA Speak-Up website, and they are asked to test their knowledge about anti-doping rules and regulations using the WADA Play-True quiz.

Note. The aim of the educational intervention was to introduce doping and the doping control process, and to provide information about the health consequences of banned substances, the risks of sport supplements, and healthy nutrition. Whistleblowing was also covered in this intervention.

Table 3

Participant Characteristics by Country and Intervention, k = 33, N = 303

		Intervention – UK		Intervention - Greece	
		Moral	Educational	Moral	Educational
Variable		(*n* = 66)	(*n* = 55)	(*n* = 102)	(*n* = 80)
		Clubs			
Number of Clubs		6	6	10	11
Sex	Male	3	3	4	4
	Female	0	2	2	3
	Mixed	3	1	4	4
Sport type	Individual	0	1	5	5
	Team	4	5	5	6
	Mixed	2	0	0	0
		Athletes			
Sex	Male	50 (16.5%)	34 (11.2%)	66 (21.8%)	51 (16.8%)
	Female	16 (5.3%)	21 (6.9%)	36 (11.9%)	29 (9.6%)
Sport type	Individual	6 (2.0%)	6 (2.0%)	41 (13.5%)	41 (13.5%)
	Team	59 (19.5%)	49 (16.2%)	61 (20.1%)	39 (12.9%)
		M / M_{dn} (*SD*)			
Age		16.61 / 17 (0.68)	18.00 / 17 (1.83)	18.19 / 19 (2.49)	19.16 / 19 (1.69)
Years training		8.83 / 9 (2.95)	8.31 / 10 (3.45)	5.89 / 5 (3.90)	8.29 / 8.5 (3.49)
Hours/week training		11.11 / 14 (4.94)	6.89 / 6 (3.12)	7.84 / 6 (3.68)	8.64 / 8 (4.33)

Table 4

Reliability Estimates for all Scales and Descriptive Statistics by Country and Intervention

		Intervention - UK						Intervention - Greece					
		Moral			Educational			Moral			Educational		
Outcome	ω	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*	*N*	*M*	*SD*
Doping likelihood (pre)	.86	66	2.37	1.49	55	2.88	1.69	102	2.36	1.40	80	2.30	1.17
Doping likelihood (post)	.88	61	1.97	1.22	51	2.37	1.56	83	1.73	0.97	77	1.94	1.13
Doping likelihood (3-month follow up)	.89	50	1.83	1.13	41	2.00	1.39	76	1.68	0.92	64	1.96	1.21
Doping likelihood (6-month follow up)	.89	55	1.68	1.13	34	1.85	1.12	76	1.54	0.82	72	2.13	1.45
Moral identity (pre)	.76	65	5.65	1.16	55	5.86	1.08	102	5.76	1.19	80	5.91	0.95
Moral identity (post)	.85	60	5.62	1.21	51	6.24	0.84	83	6.23	0.96	77	6.00	0.96
Moral identity (3-month follow up)	.78	50	5.66	1.12	41	6.15	1.05	76	6.18	0.82	64	6.00	0.84
Moral identity (6-month follow up)	.80	55	5.81	1.26	34	6.12	1.01	76	6.27	0.70	72	5.72	1.12
Moral disengagement (pre)	.75	66	2.37	0.87	55	2.47	0.88	102	1.91	0.78	80	2.23	1.05
Moral disengagement (post)	.74	61	2.15	0.88	51	2.07	0.66	83	1.71	0.75	77	1.73	0.68
Moral disengagement (3-month follow up)	.81	51	2.09	1.10	41	2.09	0.89	76	1.62	0.64	64	1.85	0.83
Moral disengagement (6-month follow up)	.77	55	1.92	0.93	34	1.86	0.74	76	1.49	0.60	72	1.81	0.79
Moral atmosphere (pre)	.92	65	2.34	1.01	55	2.78	1.37	102	2.61	1.11	80	2.77	1.06
Moral atmosphere (post)	.90	61	2.30	1.08	51	2.45	1.09	83	2.25	0.89	77	2.53	0.99
Moral atmosphere (3-month follow up)	.93	51	1.94	0.92	41	2.06	0.98	76	2.11	0.88	64	2.65	1.19
Moral atmosphere (6-month follow up)	.92	55	1.96	1.00	34	1.90	0.87	76	2.02	0.81	72	2.37	1.00
Anticipated guilt (pre)	.90	65	5.34	1.45	55	5.24	1.51	102	5.37	1.28	80	5.07	1.45
Anticipated guilt (post)	.91	60	5.56	1.48	51	5.80	1.19	83	5.68	1.28	77	5.32	1.41
Anticipated guilt (3-month follow up)	.93	50	5.84	1.42	41	5.94	1.41	76	5.74	1.24	64	5.32	1.40
Anticipated guilt (6-month follow up)	.93	55	5.83	1.62	34	5.99	1.34	76	5.54	1.46	72	5.41	1.43

Table 5

Results for Pre-Post Changes in Multilevel Piecewise Linear Growth Models for Study Outcomes

	Intention-to-Treat Analysis $N = 303$, $k = 33$			Per-Protocol Analysis $n = 172$, $k = 32$		
	Estimate	95% CI	ES	Estimate	95% CI	ES
Doping Likelihood						
Time	-0.65 (0.27)	[-1.19, -0.13]	-.46	-0.91 (0.33)	[-1.60, -0.28]	-.65
Time*Intervention	0.16 (0.35)	[-0.53, 0.86]	.11	0.72 (0.45)	[-0.16, 1.63]	.52
Time*Country	0.11 (0.33)	[-0.54, 0.74]	.08	0.57 (0.38)	[-0.17, 1.34]	.41
Time*Interv.*Country	-0.41 (0.45)	[-1.30, 0.48]	-.29	-1.31 (0.54)	[-2.38, -0.26]	-.94
Moral Identity						
Time	0.38 (0.21)	[-0.04, 0.78]	.35	0.29 (0.24)	[-0.20, 0.76]	.30
Time*Intervention	-0.37 (0.28)	[-0.92, 0.17]	-.34	-0.47 (0.32)	[-1.12, 0.14]	-.48
Time*Country	-0.25 (0.25)	[-0.75, 0.25]	-.23	-0.10 (0.28)	[-0.65, 0.45]	-.10
Time*Interv.*Country	0.57 (0.35)	[-0.11, 1.25]	.52	0.49 (0.39)	[-0.25, 1.28]	.50
Moral Disengagement						
Time	-0.51 (0.17)	[-0.84, -0.18]	-.56	-0.49 (0.22)	[-0.93, -0.04]	-.53
Time*Intervention	0.23 (0.22)	[-0.20, 0.67]	.25	0.03 (0.30)	[-0.57, 0.60]	.03
Time*Country	-0.03 (0.20)	[-0.43, 0.38]	-.03	-0.10 (0.26)	[-0.62, 0.41]	-.11
Time*Interv.*Country	0 (0.28)	[-0.58, 0.54]	.00	0.22 (0.35)	[-0.47, 0.94]	.24
Moral Atmosphere						
Time	-0.26 (0.23)	[-0.70, 0.19]	-.23	-0.11 (0.30)	[-0.73, 0.46]	-.10
Time*Intervention	0.23 (0.30)	[-0.37, 0.82]	.20	0.32 (0.39)	[-0.43, 1.11]	.30
Time*Country	0.09 (0.28)	[-0.47, 0.63]	.08	0.08 (0.34)	[-0.58, 0.78]	.07
Time*Interv.*Country	-0.51 (0.39)	[-1.28, 0.25]	-.45	-0.70 (0.47)	[-1.64, 0.22]	-.65
Anticipated Guilt						
Time	0.49 (0.24)	[0.03, 0.98]	.35	0.43 (0.35)	[-0.26, 1.13]	.31
Time*Intervention	-0.26 (0.31)	[-0.89, 0.35]	-.19	-0.47 (0.45)	[-1.36, 0.44]	-.34
Time*Country	-0.23 (0.29)	[-0.81, 0.34]	-.16	-0.16 (0.40)	[-0.94, 0.65]	-.12
Time*Interv.*Country	0.35 (0.40)	[-0.42, 1.14]	.25	0.54 (0.54)	[-0.54, 1.59]	.39

Note. Intervention: 0 = educational, 1 = moral; country: 0 = UK, 1 = Greece; Intention-to-treat analyses included all cases available. Per-protocol analysis included cases who completed all six intervention sessions and provided responses to all waves of measurement (i.e., listwise deletion). Posterior standard deviations are in parenthesis. CI = credibility interval; ES = effect size, expressed in the baseline standard deviation units; values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and 0.80 or greater a large effect. All effects (including those for the control variables) are reported in the supplementary file. The full model was adjusted for sex.

Table 6

Adjusted Pre-Post Changes in Study Outcomes by Country and Intervention

	Intention-to-Treat Analysis $N = 303$, $k = 33$			Per-Protocol Analysis $n = 172$, $k = 32$		
	Δ*M* (*SD*)	95% CI	ES	Δ*M* (*SD*)	95% CI	ES
Doping Likelihood						
UK: Moral	-0.49 (0.25)	[-0.99, -0.01]	-.34	-0.20 (0.32)	[-0.83, 0.43]	-.14
UK: Educational	-0.65 (0.27)	[-1.19, -0.13]	-.46	-0.91 (0.33)	[-1.60, -0.28]	-.66
Greece: Moral	-0.79 (0.21)	[-1.23, -0.38]	-.56	-0.94 (0.23)	[-1.39, -0.49]	-.68
Greece: Educational	-0.54 (0.22)	[-0.99, -0.12]	-.38	-0.35 (0.24)	[-0.82, 0.11]	-.25
Moral Identity						
UK: Moral	0 (0.19)	[-0.38, 0.38]	.00	-0.19 (0.23)	[-0.64, 0.25]	-.19
UK: Educational	0.38 (0.21)	[-0.04, 0.78]	.34	0.29 (0.24)	[-0.20, 0.76]	.30
Greece: Moral	0.33 (0.17)	[-0.01, 0.64]	.30	0.21 (0.17)	[-0.12, 0.53]	.22
Greece: Educational	0.13 (0.17)	[-0.21, 0.45]	.12	0.19 (0.17)	[-0.16, 0.52]	.20
Moral Disengagement						
UK: Moral	-0.27 (0.15)	[-0.58, 0.03]	-.30	-0.46 (0.21)	[-0.88, -0.07]	-.50
UK: Educational	-0.51 (0.17)	[-0.84, -0.18]	-.55	-0.49 (0.22)	[-0.93, -0.04]	-.53
Greece: Moral	-0.31 (0.13)	[-0.58, -0.05]	-.33	-0.33 (0.16)	[-0.65, -0.04]	-.36
Greece: Educational	-0.53 (0.13)	[-0.80, -0.28]	-.58	-0.58 (0.16)	[-0.90, -0.28]	-.64
Moral Atmosphere						
UK: Moral	-0.03 (0.21)	[-0.45, 0.39]	-.03	0.22 (0.27)	[-0.31, 0.74]	.20
UK: Educational	-0.26 (0.23)	[-0.70, 0.19]	-.23	-0.11 (0.30)	[-0.73, 0.46]	-.10
Greece: Moral	-0.45 (0.18)	[-0.82, -0.10]	-.40	-0.40 (0.20)	[-0.80, 0.01]	-.37
Greece: Educational	-0.17 (0.18)	[-0.53, 0.19]	-.15	-0.03 (0.21)	[-0.44, 0.39]	-.02
Anticipated Guilt						
UK: Moral	0.23 (0.22)	[-0.20, 0.65]	.16	-0.03 (0.31)	[-0.64, 0.59]	-.02
UK: Educational	0.49 (0.24)	[0.03, 0.98]	.35	0.43 (0.35)	[-0.26, 1.13]	.31
Greece: Moral	0.35 (0.19)	[-0.02, 0.73]	.25	0.35 (0.23)	[-0.11, 0.80]	.25
Greece: Educational	0.26 (0.19)	[-0.12, 0.64]	.19	0.28 (0.24)	[-0.19, 0.75]	.20

Note. Intention-to-treat analyses included all cases available. Per-protocol analysis included cases who completed all intervention sessions (six) and provided responses to all waves of measurement (i.e., listwise deletion). Posterior standard deviations are in parenthesis. CI = credibility interval; ES = effect size, expressed in the baseline standard deviation units, with values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and 0.80 or greater a large effect.

Figure 1

CONSORT Flowchart

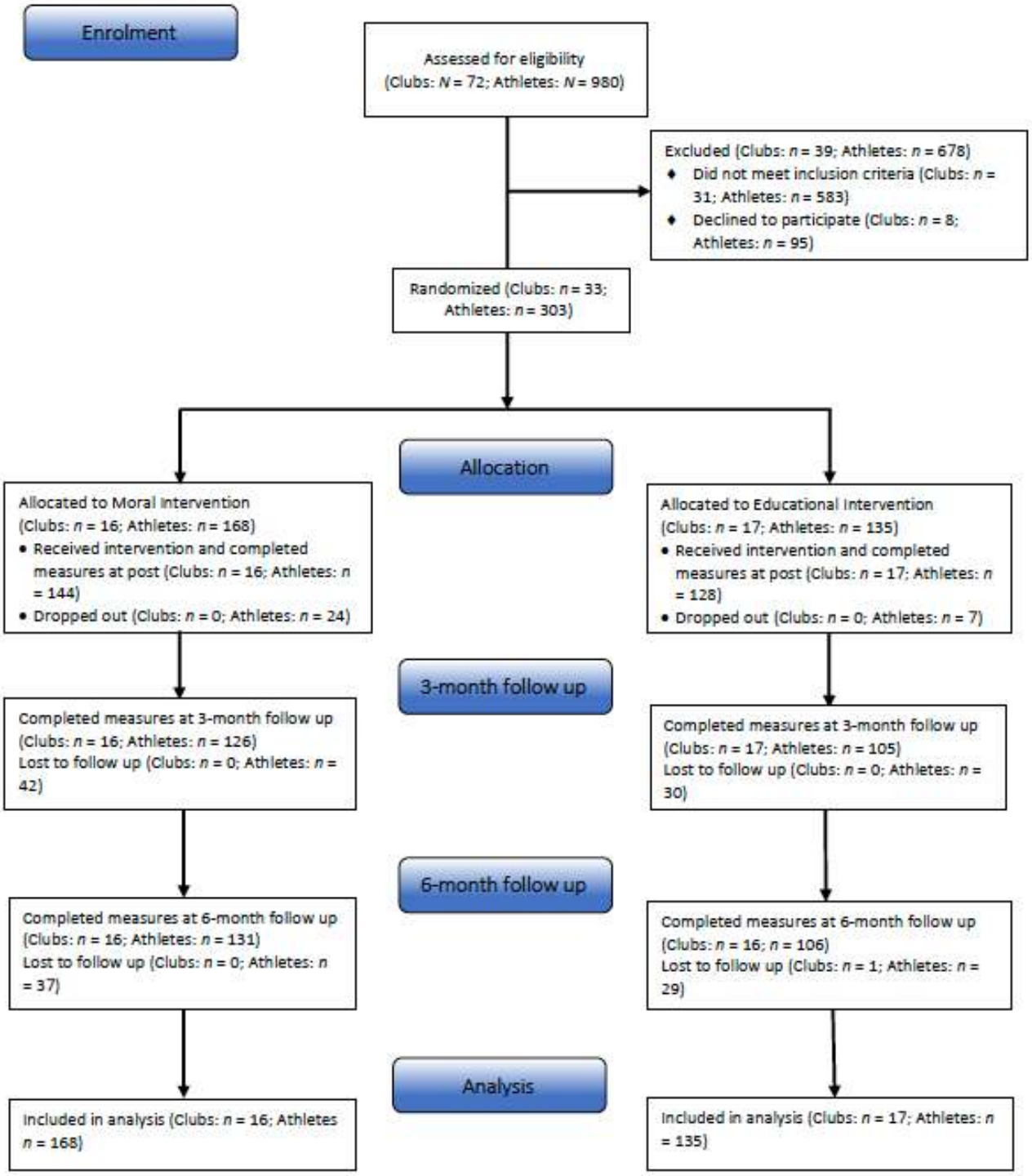
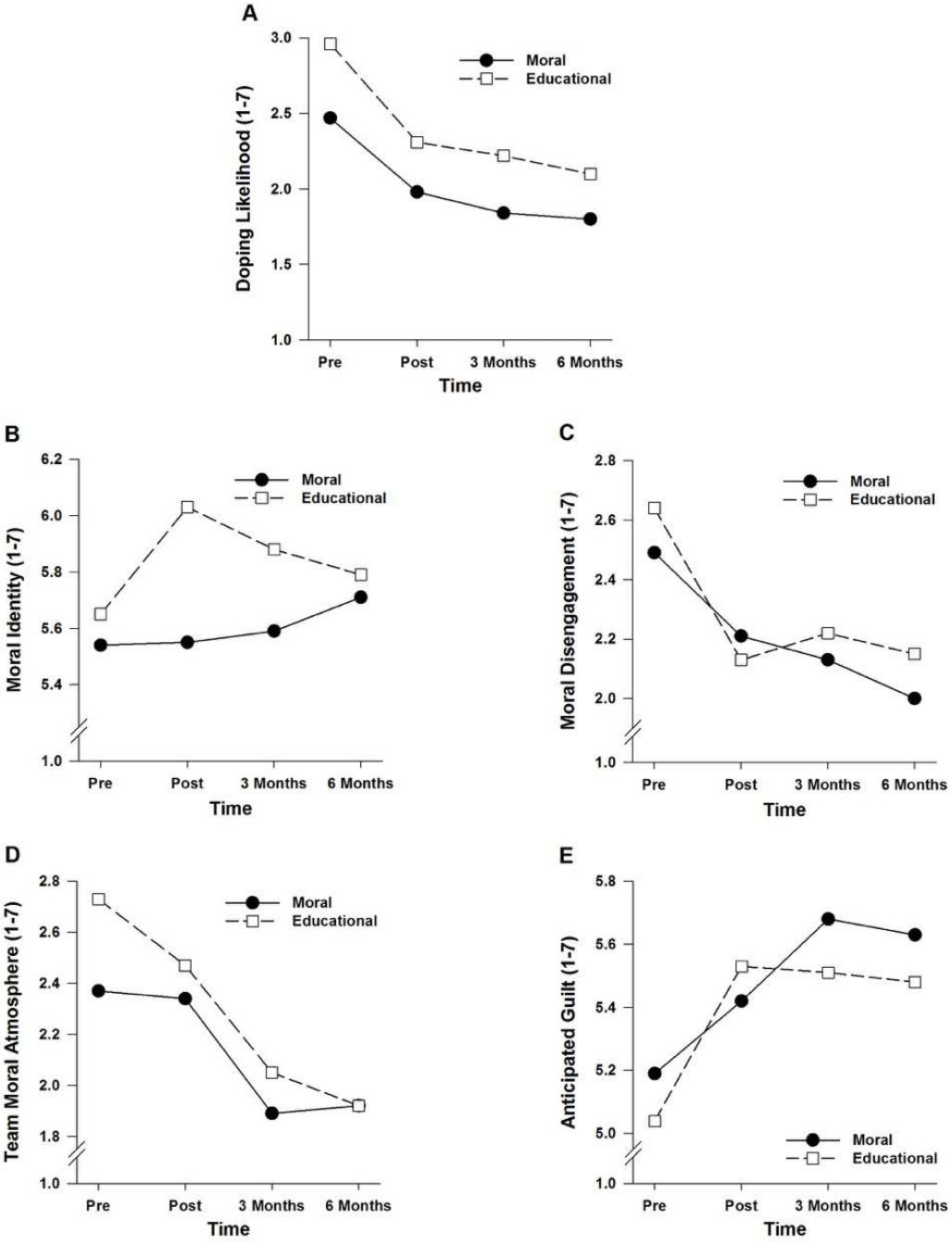


Figure 2

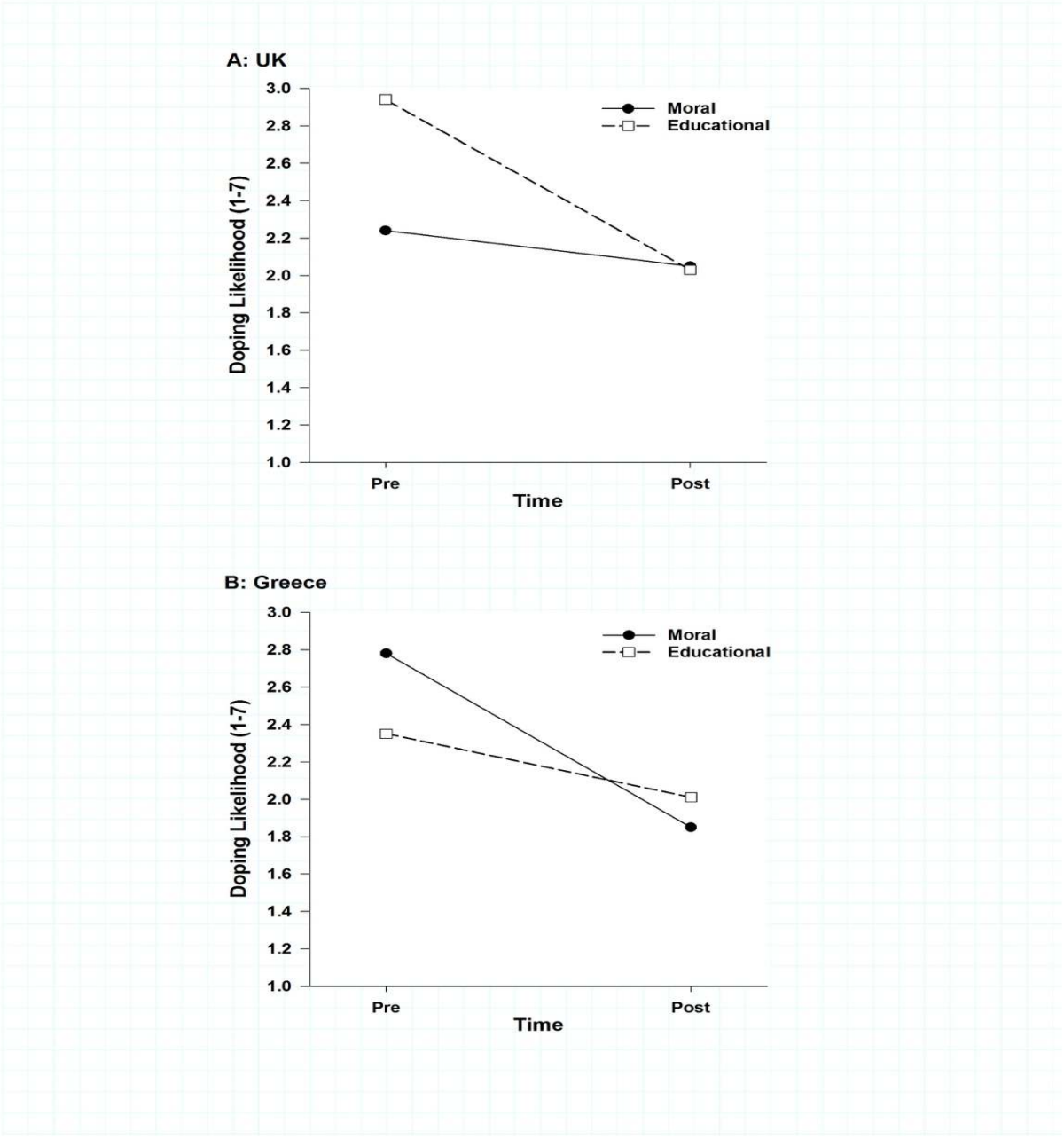
Changes in Outcomes as a Function of Time and Intervention



Note. This figure shows adjusted means in doping likelihood (Panel A), moral identity (Panel B), moral disengagement (Panel C), moral atmosphere (Panel D), and anticipated guilt (Panel E) across the four time points for each intervention group, using all available data (N_{clubs} = 33, $N_{athletes}$ = 303).

Figure 3

Changes in Doping Likelihood as a Function of Intervention in the UK (Panel A) and Greece (Panel B)



Note. This figure illustrates the interaction between Intervention, Country and Time on doping likelihood, revealed in the per-protocol analysis. In the UK, the moral intervention showed a small nonsignificant change from pre to post (ES = -.14), but the educational intervention showed a large significant change (ES = -.68). In Greece, the moral intervention showed a large significant change (ES = -.68) but the educational showed a small nonsignificant change (ES = -.25). ES = Effect size

Supplementary Material

Table S1

Results of Confirmatory Factor Analysis for Measured Variables

	N	χ^2	*df*	CFI	TLI	RMSEA	90% CI	SRMR
Moral Identity (pre)	303	40.4	9	.917	.862	.107	[.075, .142]	.047
Moral Identity (post)	272	39.3	9	.914	.857	.111	[.077, .148]	.053
Moral Identity (3-month follow up)	232	50.0	9	.899	.832	.140	[.104, .179]	.051
Moral Identity (6-month follow up)	237	27.4	9	.944	.906	.093	[.054, .134]	.046
Moral disengagement (pre)	302	82.1	9	.943	.905	.164	[.133, .197]	.034
Moral disengagement (post)	272	96.3	9	.913	.855	.189	[.156, .224]	.044
Moral disengagement (3-month follow up)	232	84.3	9	.928	.880	.190	[.154, .228]	.039
Moral disengagement (6-month follow up)	237	58.3	9	.950	.916	.152	[.116, .190]	.032
Moral atmosphere (pre)	302	112.4	5	.902	.804	.267	[.225, .311]	.062
Moral atmosphere (post)	271	33.5	5	.967	.934	.145	[.101, .193]	.028
Moral atmosphere (3-month follow up)	231	14.0	5	.990	.980	.088	[.035, .145]	.018
Moral atmosphere (6-month follow up)	237	81.7	5	.926	.853	.254	[.208, .304]	.040

Note. CFI = comparative fit index; TLI = Tucker Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual.

Table S2

Results for Multilevel Piecewise Linear Growth Model for Doping Likelihood

	Intention-to-Treat Analysis			Per-Protocol Analysis		
	Estimate	95% CI	ES	Estimate	95% CI	ES
Baseline	2.96 (0.33)	[2.29, 3.59]		2.94 (0.43)	[2.09, 3.78]	
Baseline*Intervention	-0.49 (0.44)	[-1.34, 0.40]	-.34	-0.70 (0.58)	[-1.83, 0.45]	-.50
Baseline*Country	-0.54 (0.40)	[-1.32, 0.25]	-.38	-0.59 (0.49)	[-1.58, 0.37]	-.42
Baseline*Interv.*Country	0.63 (0.56)	[-0.47, 1.74]	.44	1.13 (0.69)	[-0.21, 2.51]	.81
Baseline*Sex	-0.42 (0.23)	[-0.86, 0.04]	-.29	-0.56 (0.27)	[-1.07, -0.03]	-.40
Pre – post						
Time	-0.65 (0.27)	[-1.19, -0.13]	-.46	-0.91 (0.33)	[-1.6, -0.28]	-.65
Time*Intervention	0.16 (0.35)	[-0.53, 0.86]	.11	0.72 (0.45)	[-0.16, 1.63]	.52
Time*Country	0.11 (0.33)	[-0.54, 0.74]	.08	0.57 (0.38)	[-0.17, 1.34]	.41
Time*Interv.*Country	-0.41 (0.45)	[-1.30, 0.48]	-.29	-1.31 (0.54)	[-2.38, -0.26]	-.94
Time*Sex	0.35 (0.21)	[-0.05, 0.77]	.25	0.29 (0.22)	[-0.15, 0.73]	.21
Post – 3-month follow up						
Time	-0.09 (0.24)	[-0.55, 0.39]	-.06	0.14 (0.28)	[-0.40, 0.70]	.10
Time*Intervention	-0.05 (0.31)	[-0.67, 0.55]	-.04	-0.39 (0.37)	[-1.13, 0.33]	-.28
Time*Country	0.17 (0.28)	[-0.39, 0.74]	.12	-0.23 (0.32)	[-0.85, 0.40]	-.17
Time*Interv.*Country	-0.04 (0.39)	[-0.82, 0.74]	-.03	0.46 (0.44)	[-0.41, 1.34]	.33
Time*Sex	-0.16 (0.18)	[-0.52, 0.19]	-.11	-0.10 (0.19)	[-0.47, 0.26]	-.07
3-month – 6-month follow up						
Time	-0.12 (0.24)	[-0.61, 0.35]	-.08	-0.13 (0.33)	[-0.78, 0.53]	-.09
Time*Intervention	0.08 (0.32)	[-0.52, 0.72]	.06	0.32 (0.45)	[-0.52, 1.24]	.23
Time*Country	0.39 (0.29)	[-0.17, 0.99]	.27	0.45 (0.38)	[-0.30, 1.21]	.32
Time*Interv.*Country	-0.43 (0.40)	[-1.24, 0.32]	-.30	-0.75 (0.53)	[-1.82, 0.27]	-.54
Time*Sex	-0.18 (0.17)	[-0.52, 0.15]	-.13	-0.16 (0.18)	[-0.52, 0.19]	-.12

Note. Intervention: 0 = educational, 1 = moral; country: 0 = UK, 1 = Greece; sex: 0 = male, 1 = female. Intention-to-treat analysis included all cases available. Per-protocol analysis included cases who completed all six intervention sessions and provided responses to all waves of measurement (i.e., listwise deletion). Posterior standard deviations are in parenthesis. CI = credibility intervals; ES = effect size, expressed in the baseline standard deviation units; values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and 0.80 or greater a large effect.

Table S3

Results for Multilevel Piecewise Linear Growth Model for Moral Identity

	Intention-to-Treat Analysis			Per-Protocol Analysis		
	Estimate	95% CI	ES	Estimate	95% CI	ES
Baseline	5.65 (0.23)	[5.21, 6.11]		5.67 (0.26)	[5.13, 6.17]	
Baseline*Intervention	-0.11 (0.30)	[-0.71, 0.48]	-.10	-0.36 (0.35)	[-1.04, 0.34]	-.37
Baseline*Country	0.07 (0.28)	[-0.49, 0.62]	.06	0.02 (0.30)	[-0.57, 0.63]	.02
Baseline*Interv.*Country	0.05 (0.38)	[-0.70, 0.82]	.05	0.47 (0.42)	[-0.37, 1.31]	.48
Baseline*Sex	0.49 (0.17)	[0.16, 0.81]	.45	0.49 (0.18)	[0.14, 0.84]	.50
Pre – post						
Time	0.38 (0.21)	[-0.04, 0.78]	.35	0.29 (0.24)	[-0.20, 0.76]	.30
Time*Intervention	-0.37 (0.28)	[-0.92, 0.17]	-.34	-0.47 (0.32)	[-1.12, 0.14]	-.48
Time*Country	-0.25 (0.25)	[-0.75, 0.25]	-.23	-0.10 (0.28)	[-0.65, 0.45]	-.10
Time*Interv.*Country	0.57 (0.35)	[-0.11, 1.25]	.52	0.49 (0.39)	[-0.25, 1.28]	.50
Time*Sex	0.01 (0.16)	[-0.29, 0.33]	.01	0.08 (0.16)	[-0.22, 0.39]	.08
Post – 3-month follow up						
Time	-0.15 (0.24)	[-0.63, 0.32]	-.14	-0.12 (0.28)	[-0.65, 0.43]	-.12
Time*Intervention	0.19 (0.32)	[-0.44, 0.82]	.17	0.53 (0.36)	[-0.19, 1.24]	.54
Time*Country	0.06 (0.29)	[-0.51, 0.63]	.05	-0.04 (0.32)	[-0.66, 0.59]	-.04
Time*Interv.*Country	-0.20 (0.40)	[-0.99, 0.60]	-.18	-0.47 (0.43)	[-1.32, 0.39]	-.48
Time*Sex	0.09 (0.18)	[-0.27, 0.43]	.08	0.05 (0.19)	[-0.32, 0.41]	.05
3-month – 6-month follow up						
Time	-0.09 (0.23)	[-0.54, 0.37]	-.08	-0.05 (0.27)	[-0.57, 0.49]	-.05
Time*Intervention	0.21 (0.29)	[-0.37, 0.77]	.19	0.09 (0.36)	[-0.62, 0.80]	.09
Time*Country	-0.16 (0.27)	[-0.68, 0.37]	-.15	-0.09 (0.31)	[-0.71, 0.52]	-.09
Time*Interv.*Country	0.11 (0.36)	[-0.60, 0.83]	.10	0.17 (0.43)	[-0.69, 1.02]	.17
Time*Sex	0.01 (0.16)	[-0.30, 0.34]	.01	-0.10 (0.18)	[-0.45, 0.27]	-.10

Note. Intervention: 0 = educational, 1 = moral; country: 0 = UK, 1 = Greece; sex: 0 = male, 1 = female. Intention-to-treat analysis included all cases available. Per-protocol analysis included cases who completed all six intervention sessions and provided responses to all waves of measurement (i.e., listwise deletion). Posterior standard deviations are in parenthesis. CI = credibility intervals; ES = effect size, expressed in the baseline standard deviation units; values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and 0.80 or greater a large effect.

Table S4

Results for Multilevel Piecewise Linear Growth Model for Moral Disengagement

	Intention-to-Treat Analysis			Per-Protocol Analysis		
	Estimate	95% CI	ES	Estimate	95% CI	ES
Baseline	2.64 (0.23)	[2.18, 3.09]		2.49 (0.31)	[1.86, 3.10]	
Baseline*Intervention	-0.15 (0.31)	[-0.75, 0.46]	-.16	0.04 (0.42)	[-0.75, 0.90]	.04
Baseline*Country	-0.26 (0.28)	[-0.81, 0.30]	-.28	-0.14 (0.36)	[-0.84, 0.60]	-.15
Baseline*Interv.*Country	-0.09 (0.39)	[-0.84, 0.69]	-.10	-0.26 (0.50)	[-1.28, 0.71]	-.28
Baseline*Sex	-0.53 (0.15)	[-0.82, -0.24]	-.58	-0.60 (0.19)	[-0.98, -0.25]	-.65
Pre – post						
Time	-0.51 (0.17)	[-0.84, -0.18]	-.56	-0.49 (0.22)	[-0.93, -0.04]	-.53
Time*Intervention	0.23 (0.22)	[-0.20, 0.67]	.25	0.03 (0.30)	[-0.57, 0.60]	.03
Time*Country	-0.03 (0.20)	[-0.43, 0.38]	-.03	-0.10 (0.26)	[-0.62, 0.41]	-.11
Time*Interv.*Country	0 (0.28)	[-0.58, 0.54]	.00	0.22 (0.35)	[-0.47, 0.94]	.24
Time*Sex	0.19 (0.11)	[-0.03, 0.42]	.21	0.28 (0.14)	[0.02, 0.56]	.31
Post – 3-month follow up						
Time	0.09 (0.17)	[-0.26, 0.42]	.10	0.06 (0.25)	[-0.44, 0.54]	.07
Time*Intervention	-0.17 (0.23)	[-0.62, 0.30]	-.19	0.15 (0.33)	[-0.49, 0.81]	.16
Time*Country	0 (0.21)	[-0.42, 0.41]	.00	0.01 (0.29)	[-0.56, 0.58]	.01
Time*Interv.*Country	-0.02 (0.29)	[-0.61, 0.56]	-.02	-0.35 (0.40)	[-1.15, 0.42]	-.38
Time*Sex	0.10 (0.13)	[-0.15, 0.36]	.11	0.12 (0.15)	[-0.17, 0.40]	.13
3-month – 6-month follow up						
Time	-0.07 (0.18)	[-0.43, 0.28]	-.08	-0.06 (0.23)	[-0.51, 0.41]	-.07
Time*Intervention	-0.06 (0.23)	[-0.51, 0.39]	-.07	-0.03 (0.31)	[-0.64, 0.57]	-.03
Time*Country	0.08 (0.21)	[-0.32, 0.51]	.09	0.14 (0.27)	[-0.39, 0.67]	.15
Time*Interv.*Country	-0.06 (0.29)	[-0.63, 0.51]	-.07	-0.13 (0.37)	[-0.86, 0.60]	-.14
Time*Sex	-0.09 (0.12)	[-0.34, 0.15]	-.10	-0.14 (0.15)	[-0.44, 0.14]	-.15

Note. Intervention: 0 = educational, 1 = moral; country: 0 = UK, 1 = Greece; sex: 0 = male, 1 = female. Intention-to-treat analysis included all cases available. Per-protocol analysis included cases who completed all six intervention sessions and provided responses to all waves of measurement (i.e., listwise deletion). Posterior standard deviations are in parenthesis. CI = credibility intervals; ES = effect size, expressed in the baseline standard deviation units; values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and 0.80 or greater a large effect.

Table S5

Results for Multilevel Piecewise Linear Growth Model for Moral Atmosphere

	Intention-to-Treat Analysis			Per-Protocol Analysis		
	Estimate	95% CI	ES	Estimate	95% CI	ES
Baseline	2.73 (0.35)	[2.02, 3.42]		2.70 (0.43)	[1.85, 3.55]	
Baseline*Intervention	-0.36 (0.49)	[-1.31, 0.62]	-.32	-0.30 (0.56)	[-1.41, 0.82]	-.28
Baseline*Country	0.07 (0.43)	[-0.79, 0.92]	.06	0 (0.50)	[-0.99, 0.98]	.00
Baseline*Interv.*Country	0.37 (0.61)	[-0.85, 1.58]	.33	0.38 (0.69)	[-0.98, 1.73]	.35
Baseline*Sex	-0.15 (0.19)	[-0.53, 0.24]	-.13	-0.22 (0.23)	[-0.67, 0.24]	-.20
Pre – post						
Time	-0.26 (0.23)	[-0.70, 0.19]	-.23	-0.11 (0.30)	[-0.73, 0.46]	-.10
Time*Intervention	0.23 (0.30)	[-0.37, 0.82]	.20	0.32 (0.39)	[-0.43, 1.11]	.30
Time*Country	0.09 (0.28)	[-0.47, 0.63]	.08	0.08 (0.34)	[-0.58, 0.78]	.07
Time*Interv.*Country	-0.51 (0.39)	[-1.28, 0.25]	-.45	-0.70 (0.47)	[-1.64, 0.22]	-.65
Time*Sex	-0.08 (0.16)	[-0.39, 0.24]	-.07	-0.22 (0.19)	[-0.59, 0.16]	-.20
Post – 3-month follow up						
Time	-0.42 (0.25)	[-0.91, 0.09]	-.37	-0.44 (0.35)	[-1.14, 0.27]	-.41
Time*Intervention	-0.03 (0.33)	[-0.70, 0.61]	-.03	-0.13 (0.46)	[-1.02, 0.78]	-.12
Time*Country	0.46 (0.30)	[-0.14, 1.05]	.41	0.41 (0.40)	[-0.38, 1.21]	.38
Time*Interv.*Country	-0.17 (0.42)	[-1.00, 0.67]	-.15	-0.01 (0.55)	[-1.12, 1.08]	-.01
Time*Sex	0.08 (0.18)	[-0.29, 0.43]	.07	0.17 (0.22)	[-0.28, 0.60]	.16
3-month – 6-month follow up						
Time	-0.20 (0.25)	[-0.68, 0.29]	-.18	-0.21 (0.33)	[-0.86, 0.44]	-.19
Time*Intervention	0.23 (0.32)	[-0.42, 0.87]	.20	0.31 (0.42)	[-0.52, 1.15]	.29
Time*Country	-0.05 (0.29)	[-0.62, 0.53]	-.04	-0.06 (0.38)	[-0.84, 0.70]	-.06
Time*Interv.*Country	-0.12 (0.40)	[-0.91, 0.68]	-.11	-0.24 (0.52)	[-1.28, 0.78]	-.22
Time*Sex	0.05 (0.16)	[-0.26, 0.37]	.04	0.17 (0.20)	[-0.21, 0.58]	.16

Note. Intervention: 0 = educational, 1 = moral; country: 0 = UK, 1 = Greece; sex: 0 = male, 1 = female. Intention-to-treat analyses included all cases available. Per-protocol analysis included cases who completed all six intervention sessions and provided responses to all waves of measurement (i.e., listwise deletion). Posterior standard deviations are in parenthesis. CI = credibility interval; ES = effect size, expressed in the baseline standard deviation units; values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and 0.80 or greater a large effect.

Table S6

Results for Multilevel Piecewise Linear Growth Model for Anticipated Guilt

	Intention-to-Treat Analysis			Per-Protocol Analysis		
	Estimate	95% CI	ES	Estimate	95% CI	ES
Baseline	5.04 (0.35)	[4.35, 5.74]		4.97 (0.52)	[3.94, 5.99]	
Baseline*Intervention	0.15 (0.47)	[-0.81, 1.07]	.11	0.20 (0.67)	[-1.16, 1.5]	.14
Baseline*Country	-0.25 (0.43)	[-1.11, 0.60]	-.18	-0.23 (0.60)	[-1.42, 0.96]	-.17
Baseline*Interv.*Country	0.06 (0.60)	[-1.12, 1.25]	.04	-0.06 (0.80)	[-1.63, 1.57]	-.04
Baseline*Sex	0.68 (0.23)	[0.24, 1.13]	.48	0.82 (0.29)	[0.26, 1.38]	.59
Pre – post						
Time	0.49 (0.24)	[0.03, 0.98]	.35	0.43 (0.35)	[-0.26, 1.13]	.31
Time*Intervention	-0.26 (0.31)	[-0.89, 0.35]	-.19	-0.47 (0.45)	[-1.36, 0.44]	-.34
Time*Country	-0.23 (0.29)	[-0.81, 0.34]	-.16	-0.16 (0.40)	[-0.94, 0.65]	-.12
Time*Interv.*Country	0.35 (0.40)	[-0.42, 1.14]	.25	0.54 (0.54)	[-0.54, 1.59]	.39
Time*Sex	0.10 (0.18)	[-0.26, 0.46]	.07	0.06 (0.22)	[-0.38, 0.50]	.04
Post – 3-month follow up						
Time	-0.02 (0.27)	[-0.56, 0.49]	-.01	-0.33 (0.48)	[-1.41, 0.50]	-.24
Time*Intervention	0.28 (0.35)	[-0.40, 0.99]	.20	0.68 (0.64)	[-0.50, 2.04]	.49
Time*Country	0.01 (0.32)	[-0.60, 0.65]	.01	0.34 (0.56)	[-0.64, 1.58]	.24
Time*Interv.*Country	-0.21 (0.44)	[-1.09, 0.66]	-.15	-0.56 (0.78)	[-2.21, 0.91]	-.40
Time*Sex	0.05 (0.20)	[-0.34, 0.43]	.04	-0.09 (0.26)	[-0.61, 0.40]	-.06
3-month – 6-month follow up						
Time	-0.03 (0.30)	[-0.63, 0.56]	-.02	0.48 (0.60)	[-0.57, 1.79]	.35
Time*Intervention	-0.02 (0.38)	[-0.77, 0.73]	-.01	-0.07 (0.80)	[-1.71, 1.43]	-.05
Time*Country	0.09 (0.36)	[-0.61, 0.79]	.06	-0.46 (0.70)	[-1.98, 0.79]	-.33
Time*Interv.*Country	-0.25 (0.47)	[-1.19, 0.69]	-.18	-0.22 (0.97)	[-2.05, 1.82]	-.16
Time*Sex	0.02 (0.21)	[-0.39, 0.44]	.01	0.10 (0.28)	[-0.43, 0.64]	.07

Note. Intervention: 0 = educational, 1 = moral; country: 0 = UK, 1 = Greece; sex: 0 = male, 1 = female. Intention-to-treat analysis included all cases available. Per-protocol analysis included cases who completed all six intervention sessions and provided responses to all waves of measurement (i.e., listwise deletion). Posterior standard deviations are in parenthesis. CI = credibility interval; ES = effect size, expressed in the baseline standard deviation units; values between 0.20 and 0.49 constitute a small effect, 0.50 to 0.79 a medium effect, and 0.80 or greater a large effect.

CONSORT 2010 checklist of information to include when reporting a randomised trial*

Section/Topic	Item No	Checklist item	Reported on page No
Title and abstract			
	1a	Identification as a randomised trial in the title	1
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)	2
Introduction			
Background and objectives	2a	Scientific background and explanation of rationale	5-7
	2b	Specific objectives or hypotheses	8
Methods			
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio	11
	3b	Important changes to methods after trial commencement (such as eligibility criteria), with reasons	N/A
Participants	4a	Eligibility criteria for participants	11
	4b	Settings and locations where the data were collected	15
Interventions	5	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered	9-11, Tables 2-3, 15
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	12-15
	6b	Any changes to trial outcomes after the trial commenced, with reasons	N/A
Sample size	7a	How sample size was determined	N/A
	7b	When applicable, explanation of any interim analyses and stopping guidelines	N/A
Randomisation:			
Sequence generation	8a	Method used to generate the random allocation sequence	N/A
	8b	Type of randomisation; details of any restriction (such as blocking and block size)	12
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	N/A
Implementation	10	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions	12

Blinding	11a	If done, who was blinded after assignment to interventions (for example, participants, care providers, those assessing outcomes) and how	N/A
	11b	If relevant, description of the similarity of interventions	N/A
Statistical methods	12a	Statistical methods used to compare groups for primary and secondary outcomes	15-16
	12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	17
Results			
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	Figure 1
	13b	For each group, losses and exclusions after randomisation, together with reasons	Figure 1
Recruitment	14a	Dates defining the periods of recruitment and follow-up	16
	14b	Why the trial ended or was stopped	N/A
Baseline data	15	A table showing baseline demographic and clinical characteristics for each group	Table 3
Numbers analysed	16	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups	Figure 1
Outcomes and estimation	17a	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	Tables S2-S6
	17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended	N/A
Ancillary analyses	18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory	N/A
Harms	19	All important harms or unintended effects in each group (for specific guidance see CONSORT for harms)	15
Discussion			
Limitations	20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses	23-24
Generalisability	21	Generalisability (external validity, applicability) of the trial findings	24
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	19-23
Other information			
Registration	23	Registration number and name of trial registry	N/A
Protocol	24	Where the full trial protocol can be accessed, if available	N/A
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders	1

*We strongly recommend reading this statement in conjunction with the CONSORT 2010 Explanation and Elaboration for important clarifications on all the items. If relevant, we also recommend reading CONSORT extensions for cluster randomised trials, non-inferiority and equivalence trials, non-pharmacological treatments, herbal interventions, and pragmatic trials. Additional extensions are forthcoming: for those and for up to date references relevant to this checklist, see www.consort-statement.org.