Testing the effects of a preceding self-control task on decision making in soccer refereeing

Chris Englert[1*†], Anna Dziuba[2], Geoffrey Schweizer[3†]

[1]Institute of Education, Department of Educational Psychology, University of Bern, Bern, Switzerland

[2]Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

[3]Institute of Sports Sciences, Department of Sports Psychology, Heidelberg University, Heidelberg, Germany

*** Correspondence:** Christoph.englert@edu.unibe.ch

[†]These authors have contributed equally to this work

1 Abstract

The present study tested the assumption that the momentary level of self-control strength affects the accuracy rates in a sports-related judgement and decision-making task. A total of $N = 27$ participants rated the veracity of 28 video-taped statements of soccer players who were interviewed by a non-visible referee after a critical game-related situation. In half of the videos, the players were lying and in the other half they were telling the truth. Participants were tested twice: once with temporarily depleted self-control strength and once with temporarily available self-control strength (order counterbalanced; measurements separated by exactly seven days). Self-control strength was experimentally manipulated with the Stroop task. In line with two-process models of information processing, we hypothesized that under ego depletion, information is processed in a rather heuristic manner, leading to lower accuracy rates. Contrary to our expectations, the level of temporarily available self-control strength did not have an effect on accuracy rates. Limitations and implications for future research endeavors are discussed.

Keywords: decision making, ego depletion, cognitive fatigue, sports, refereeing, Stroop, effort

2 Introduction

Deception in sports is a critical issue as it might decisively change the outcome of a match (Güldenpenning, et al., 2017). According to Hsu (1997), deception means "making someone believe something that is not true in order to get what you want" (p. 167). For instance, a wrongfully granted penalty kick during overtime in a tied soccer match will likely determine which team wins the game (Sabag et al., 2018). In sports, lying to the referee can be considered a special form of deception. While research on deception has a long tradition in sports (for an overview see Güldenpenning et al., 2017), and the ability to detect deceit and, especially, lies has been center stage in the criminal justice system (e.g., Akehurst et al., 1996) as well as in educational settings (e.g., Marksteiner et al., 2013) for many years, but only recently has the topic of lie detection been addressed in sports-related contexts. This seems rather surprising, given the high potential impact of "successfully" lying to a referee.

Given the impending influence of deceit on the results of a sporting competition, it seems highly important that a referee's judgement and decision making take place as accurately as possible. However, as far as we know, there have been very few systematic, experimental studies on referee accuracy rates regarding deception (e.g., Aragão e Pina et al., 2018; Morris & Lewis, 2010; Renden et al., 2014), as most studies on deception in sports have been correlational and, for instance, asked their participants how they would possibly behave in a certain hypothetical situation (e.g., Kavussanu & Ntoumanis, 2003). A notable exception is a study series by Morris and Lewis (2010), in which they first generated a sequence of video clips in which soccer players were instructed to overstate the effects of a tackle by an opposing player. In a subsequent study, neutral observers rated each video clip whether the respective video-taped player had actually been fouled or not. The results revealed that the neutral observers judged the video-clips very accurately. Another experimental study on lie detection was conducted by Englert and Schweizer (2020). Taking a similar approach, the authors first created 28 video clips in which soccer players were either telling the truth or lying regarding two simulated critical game situations. The veracity of each of the 28 video clips was later rated by neutral observers in a series of three studies. The results were rather mixed, as the statements of some of the interviewed players were rather easy to classify, while other players were fairly good at lying. When looking at the accuracy rates of correctly classifying truths and lies in other domains (e.g., the criminal justice system), recent meta-analyses indicate that, overall, individuals are not very accurate at detecting lies, or more precisely, they are only slightly better than the chance level (i.e., accuracy rate of 54%) (e.g., Bond & DePaulo, 2006).

It remains largely unknown which factors influence the accuracy rates of referees. Previous meta-analyses found no empirical evidence that gender, age, expertise, or certain personality traits significantly impacted the accuracy rates (e.g., Aamodt & Custer, 2006; Bond & DePaulo, 2006). In order to identify potential factors, we must first take a closer look at the actual judgement and decision-making process. Dual-process models of information processing assume that there are two different types of information processing when making a judgement (e.g., Chaiken & Maheswaran, 1994; Chaiken & Trope, 1999; Petty et al., 2005) (for an application of dual-process theorizing to the domain of sports see Furley et al., 2015): *Heuristically* (also called *peripheral route*) or *systematically* (also called *central route*). When processing information and making a judgement in a heuristic manner, individuals focus less carefully on the content of a statement and more so on peripheral cues, such as the likability or trustworthiness of the source or simply the number of arguments presented by the source (Petty et al., 2005). On the contrary, systematic information processing allows a person to carefully pay attention and evaluate the quality of the arguments presented (e.g., Chaiken & Trope, 1999). The importance of dual-process models has also been

shown in other sport- and exercise-related settings (Furley et al., 2015): for instance, a physically inactive person might have the intention to work out in the evening, but has a negative attitude towards physical exercise and tends to avoid straining physical activities (e.g., Bluemke et al., 2010). In the evening, his/her favorite TV program is on and the person has to make a decision on whether to exercise or not. When making the decision heuristically, the person is less likely to exercise as he/she pays less attention to the positive aspects of physical activity. However, when making the decision systematically, he/she weighs the positive and negative aspects of exercising against one another and is more likely to work out (see also, Englert & Rummel, 2016). Taken together, heuristic information processing is less reflective and requires less effort than systematic information processing (Petty et al. 2009; Petty et al., 2005). Previous research from the criminal justice system has reliably shown that judgements are more accurate when taking the systematic information processing route (e.g., Feeley & DeTurck, 1995; Masip et al., 2009; Vrij et al., 2010). This leads to the question: Which factors determine which type of information processing dominates in a given situation? One potential candidate is the level of temporarily available self-control strength, which we will describe in more detail in the following sections (e.g., Davis & Leo, 2012).

According to the strength model, all self-control acts are based on a global metaphorical resource with a limited capacity (e.g., Baumeister et al., 1998; see also André et al., 2019; Audiffren & André, 2015). In this context, self-control means inhibiting certain impulses or response tendencies in order to keep striving for desirable outcomes and to perform at the highest possible level (e.g., Englert, 2017, 2019). Self-control acts include, amongst others, emotion regulation, attention regulation, and most importantly for the present investigation, judgement and decision making (Hagger et al., 2010; Samuel et al., 2018) (for an overview, see also Englert, 2017, 2019). It is assumed that after individuals have worked on a self-control task their self-control resources become temporarily depleted for a certain amount of time. During this so-called state of *ego depletion*, following self-control tasks are executed less efficiently as less cognitive effort is likely to be invested (e.g., Baumeister et al., 1998). Given that self-control strength needs to be exerted in order to process information via the cognitively demanding systematic route, previous empirical research has shown that ego depleted individuals tend to process information in a heuristic manner (e.g., Baumeister et al., 2008; Unger & Stahlberg, 2011; Wheeler et al., 2007). In two studies, Reinhard, Scharmach, and Stahlberg (2013) manipulated ego depletion and found out that ego depleted participants were more likely to process information heuristically and displayed lower lie detection accuracy rates than non-depleted participants (for similar findings, see also Davis & Leo, 2012; Wheeler et al., 2007).

Based on these empirical findings and theoretical assumptions, we assumed that individuals are more likely to process information heuristically if they had been working on a straining self-control task before (i.e., under ego depletion). As systematic information processing is associated with higher accuracy rates during judgement and decision making, we tested the hypothesis that depleted individuals are less accurate in correctly classifying ambiguous situations during a soccer match than non-depleted participants (see also Reinhard et al., 2013). In order to test these assumptions, we adopted Englert and Schweizer's approach (2020) and asked participants at two separate times of measurement to rate the truth of a series of 28 video-taped statements of soccer players, in which they either lied to a referee or told him the truth. At one time of measurement, participants' self-control strength was experimentally depleted, while it remained intact at the other time of measurement (order counterbalanced).

3 Materials and Methods

3.1 Participants

A G*Power (Faul et al., 2007) analysis showed that a sample of $N = 27$ was necessary for detecting at least a medium effect (parameters: $f = 0.30$, $\alpha = 0.05$, $1-\beta = 0.85$, $r_{\text{repeated measures}} = 0.50$, $\varepsilon = 1$). Based on this estimate, a total of $N = 27$ university students from a German university volunteered to partake in the present investigation (16 females, 11 males; $M_{Age} = 27.74$ years, $SD_{Age} = 7.17$). Three participants had soccer refereeing experience ($M = 3.67$ years, $SD = 3.79$). The study was approved by the local ethics committee, and all participants delivered written informed consent.

3.2 Design, procedure, and measures

The participants were tested at two times of measurement exactly seven days apart under standardized conditions in single sessions on a regular computer in a university lab room. All instructions, video clips, and questionnaires were delivered via an online survey program (Unipark). Each participant was wearing regular stereo headphones and the sound was played at a constant volume. At one time of measurement, participants' self-control strength was experimentally depleted (depletion condition), while it remained intact at the other time of measurement (control condition; order counterbalanced). First, participants reported demographic information (i.e., age, sex, refereeing experience).

Then, self-control strength was experimentally manipulated by using the Stroop test, which has been frequently applied in self-control research (e.g., Bray et al., 2012; Englert & Bertrams, 2014). The Stroop test consists of color words which are either displayed in the same font color as the color word (congruent Stroop trial; e.g., the word "red" written in red font color) or in a different font color (incongruent Stroop trial; e.g., the word "red" written in yellow font color); participants need to always name the font color instead of the written color word. It has been reliably shown that in order to ignore the color word and to read the font color instead, self-control needs to be invested, which is why this task has been regularly applied to manipulate self-control strength. In the present study, at both times of measurement, participants first performed a series of 32 practice trials and then worked on 300 incongruent Stroop trials in the depletion condition and on 300 congruent Stroop trials in the control condition. The number of falsely identified Stroop trials and the average response latencies were measured as manipulation checks, assuming that in the depletion condition, participants would make more mistakes and would need longer to answer each trial (in milliseconds) (e.g., Bray et al., 2012; see also, Pageaux et al., 2014).

At both times of measurement, following the Stroop task, the participants were informed that they would be watching a series of video clips. These video clips were taken from Englert and Schweizer's study (2020), in which the authors created 28 video clips in which male soccer players from a club from the sixth highest league in Germany (out of 11 leagues) were either telling the truth or lying regarding two simulated critical game situations. These simulated game situations took place immediately before an interview with a professional soccer referee. In both situations, the player acted as a defender as another player played a long pass toward the goal line for his teammate. Once, the defender was asked to not allow the other player to get to the ball and to instead let the ball cross the goal line, which would lead to a goal kick for his team. In the other situation, the instructions were similar with the only difference being that the defender did actually touch the ball last before it passed the goal line. In this latter case, the correct decision would have been a corner kick. However, in both situations, the defender was asked to tell the referee, who had not seen the critical situation, in the subsequent video interview, that the offensive player had touched the ball last and the correct decision was supposedly a goal kick, meaning that the defender was telling the truth in one interview and was lying in the other. The referee asked each player exactly the same questions and was not

seen in the video. The participants in the current study did not watch the critical situation, but only the subsequent interview. The participants were also told that each player was in a similar critical situation twice during the same game and would thus be interviewed by the same referee at two separate times. However, the participants were not made aware of the fact that each player was lying in one interview and speaking the truth in the other interview, leading to a total of 14 true statements and 14 lies. On average, each video clip lasted roughly 28 seconds (M = 27.5, SD = 6.27) and the player's upper torso, face, and legs could be seen in each clip. The sound quality was the same in all video clips. Participants were further instructed that they would have to rate the veracity of each interview on a continuous scale ranging from 1 (*not at all true*) to 10 (*totally true*) immediately following each video clip (for this procedure, see also Marksteiner et al., 2013). The video clips were displayed in a randomized order immediately after finishing the Stroop task in both conditions. In total, participants rated the veracity of 28 video statements while being ego depleted and the veracity of the same 28 video statements with fully available self-control strength. In order to reduce the likelihood of a learning effect, the two times of measurement were separated by exactly seven days, and the order of the video presentation was randomized.

– Figure 1 about here –

Finally, after the second time of measurement, the participants were debriefed and thanked for their participation.

3.3 Data analysis

Data were analyzed using SPSS (version 27; SPSS Inc., Chicago, IL, United States). We ran paired samples *t*-tests to investigate the assumptions that the depletion condition would perform worse in the Stroop task (i.e., longer response latencies in milliseconds; higher number of Stroop errors) and would be less adept in correctly distinguishing between true and false statements than the control condition. All effect sizes were calculated as Cohen's d (i.e., small effect: d = 0.2; medium effect: d = 0.5; large effect: d = 0.8; Cohen, 1988). For all analyses, statistical significance was accepted as $p < 0.05$.

4 Results

4.1 Preliminary analyses

As expected, the Stroop response latencies in the depletion condition (M = 839.08 ms, SD = 179.68) were significantly longer than in the control condition (M = 717.55 ms, SD = 156.80), $t(26) = 7.02$, $p < .0001$, d = 1.35. Additionally, there was the expected tendency in the number of Stroop errors between the depletion condition (M = 7.96, SD = 6.00) and the control condition (M = 6.59, SD = 5.80), which however failed to reach statistical significance, $t(26) = 1.86$, $p = .075$, d = 0.36. On average, the depletion condition (M = 331633.70 ms, SD = 69442.08) needed significantly longer to finish the 300 Stroop trials than the control condition (M = 295259.67 ms, SD = 53785.51), $t(26) = 3.89$, $p < .0001$, d = 0.75.

4.2 Primary analyses

In line with Englert and Schweizer's approach (2020), for both conditions, we first compared the veracity ratings of the true statements to the veracity ratings of the lies in order to investigate the question of whether participants in both conditions were able to distinguish (on average) between

true and false statements (for descriptive statistics see Table 1). In both groups, false statements were rated significantly lower than true statements, indicating that participants in both conditions were able to distinguish between true and false statements (control: $t(26) = 2.15$, $p = .041$, $d = 0.41$; depletion: $t(26) = 4.34$, $p < .001$, $d = 0.83$).

Next, in order to investigate potential differences between the depletion and the control conditions, we compared the ratings of the true statements between the two times of measurement (control vs. depletion). Contrary to our hypothesis, the veracity ratings did not differ statistically significantly between the depletion condition ($M = 5.99$, $SD = 0.81$) and the control condition ($M = 5.83$, $SD = 0.98$), $t(26) = 0.81$, $p = .426$, $d = 0.16$. There were also no significant differences between the depletion condition ($M = 5.39$, $SD = 0.73$) and the control condition ($M = 5.32$, $SD = 0.98$) in the veracity ratings of the false statements, $t(26) = 0.39$, $p = .703$, $d = 0.07$ (see also Table 1).

4.3 Complementary Bayesian Hypothesis Testing

We ran additional Bayesian paired samples *t*-tests, to further investigate whether the differences in the veracity ratings of true and false statements between the depletion and the control condition do not exist (i.e., that the null hypotheses are more likely to be true; for this approach, see also Dienes, 2014; Wagenmakers et al., 2018a, b). For the true statements, a two-sided analysis revealed a Bayes factor (BF01) suggesting that the data were 3.64 times more likely under the null (i.e., the two conditions do not differ in their veracity statements of the true statements) than the alternative hypothesis (i.e., the two conditions differ) with a median effect size of 0.14, which indicates moderate evidence in favor of the null hypothesis. For the false statements, the results indicate that the observed data are 4.58 times more likely under the null (i.e., the two conditions do not differ in their veracity statements of the false statements) than the alternative hypothesis (i.e., the two conditions differ) with a median effect size of 0.07, which indicates moderate evidence in favor of the null hypothesis.

– Table 1 about here –

5 Discussion

In the present study, we tested the assumption that individuals would be less adept in correctly identifying the veracity of a player's statement following a critical game situation during a soccer match if they had been working on a straining self-control task beforehand. For that reason, participants rated a series of video statements at two times of measurement, once with fully available self-control strength and once in a state of ego depletion (order counterbalanced). According to two-process models, there are two types of information processing, namely a heuristic and a systematic mode. When judging the veracity of a statement in a heuristic manner, individuals tend to focus on rather invalid cues to deception (e.g., number of statements), while a systematic mode is related to an increased focus on valid cues (e.g., actual content of the statement) and a higher likelihood of classifying a statement correctly (DePaulo et al., 2003; Forrest et al., 2004). But, systematic information processing is effortful and, according to several authors, requires self-control strength (e.g., Baumeister et al., 2008; Davis & Leo, 2012; Reinhard et al., 2013; Unger & Stahlberg, 2011; Wheeler et al., 2007). If one's self-control resources had been taxed in a previous task, he/she is less likely to have the necessary self-control strength to process information systematically and will tend to process heuristically instead. However, the results did not support our hypothesis as there were no statistically significant differences in the accuracy rates between the control and the depletion condition.

When investigating why the control and the depletion condition did not differ regarding their veracity ratings, it is important to emphasize that in both conditions, participants actually could differentiate between true and false statements (although not very strongly). This can be considered a necessary prerequisite for testing our main hypothesis: If participants in the control condition cannot distinguish between true and false statements, then they cannot get worse in the depletion condition. Given that this prerequisite was met, how can we then explain that participants in the depletion and the control conditions did not differ, considering that the study was adequately powered and that the depletion manipulation was effective? One potential explanation for this pattern is that participants in the control condition did rely on heuristic processing as well. This would both explain why participants were not able to distinguish more strongly between false and true statements (because doing so would require more systematic processing) and why they did not get worse in the depletion condition. To address this issue, further research might want to employ not only a condition that is supposed to decrease systematic and to increase heuristic processing (such as the depletion condition in the present research), but furthermore a condition that is supposed to increase systematic processing. This might be accomplished by incentivizing participants, for example (see also, Beckmann, 2020).

Another potential explanation might be the low level of expertise/experience of the participants in our study (only three participants had soccer refereeing experience), as one might reason that participants with soccer refereeing experience are better at correctly judging player statements (e.g., MacMahon et al., 2007; Moore et al., 2019). Even though several large-scale studies from the criminal justice system and educational psychology have reliably demonstrated that the raters' expertise does not affect their accuracy rates (e.g., Aamodt & Custer, 2006; Bond & DePaulo, 2006), future studies should investigate whether the same is true in sports-related judgement and decision making situations.

We would also like to address the fact that the depletion condition took significantly longer to finish the Stroop task than the control condition. This matter seems especially important, as a recent study by Boat and colleagues (2020) revealed that longer Stroop task durations were related to lower performance in a subsequent self-control task. However, in the current study we did not find an effect of the different Stroop task durations on the veracity ratings. Future studies should continue to dig deeper into the effects of different self-control task durations on performance (see also Wolff et al., 2021).

Individuals do not only differ in their levels of temporarily available self-control, but also in their general self-control abilities, meaning that some are simply better at regulating themselves than others (i.e., trait self-control; Tangney et al., 2004). In general, individuals with higher levels of trait self-control are more adept at volitionally controlling their impulses and focusing on the task at hand (e.g., De Ridder et al., 2012). In the current study, we did not measure trait self-control strength; however, given the fact that we applied a repeated measures design, we assume that trait self-control strength did not play a major part in our study. It has to be noted that the validity of the ego depletion effect itself has been questioned on theoretical and empirical grounds. On an empirical level, some recent large-scale replication studies did not find reliable statistical evidence for the ego depletion effect (e.g., Blázquez et al., 2017; Hagger et al., 2016). For instance, Vohs and colleagues (in press) conducted a preregistered replication report with over 3,500 participants from 36 labs worldwide. While participants with depleted self-control did not differ significantly from the non-depleted participants in terms of their performance, depleted participants did feel more fatigued than control participants. So why did depleted participants feel fatigued while their actual performance did not

suffer from their depletion? It might be reasonable to assume that the dependent variable in the Vohs et al. study (Cognitive Estimation Test) (Bullard et al., 2004) was not self-control demanding *enough*. If the dependent measure only requires minimal effort, it is highly unlikely to be affected by a straining preceding self-control task (see also, Loschelder & Friese, 2016). In a similar fashion, in our study rating the videos systematically might not place sufficiently high self-control demands on one's self-control resources, thus making it more difficult to find statistically significant differences between the depleted and the non-depleted conditions. Furthermore, while the results of the Stroop test revealed the expected differences between the depletion and the control condition, we did not apply an additional manipulation check measuring the level of perceived depletion following the Stroop task. This notion seems especially important, as for instance Clarkson and colleagues (2010) have demonstrated, that participants who perceived themselves as being more depleted performed worse in following self-control acts than participants who perceived themselves as being less depleted (see also Wright & Mlynski, 2019). Even though previous studies have reliably shown that participants reported significantly higher levels of perceived depletion after the incongruent Stroop task compared to the congruent one (e.g., Hagger et al., 2010), future studies should apply additional manipulation checks to test the effectiveness of the respective ego depletion manipulation.

On a theoretical level, several researchers argue that the assumption of a limited metaphorical self-control resource is not appropriate and cannot be adequately tested empirically (for a discussion, see also Eronen & Bringmann, 2021). For instance, the process model by Inzlicht and Schmeichel (2012, 2016) postulates that a primary self-control act does not deplete limit resources but rather instigates shits in motivation (i.e., the person does not want to work on another straining task), emotions (i.e., the person perceives other straining tasks as rather negative), and attention (i.e., impaired attention regulation), which ultimately affects performance in subsequent self-control tasks. In a similar fashion, according to the behavioral restraint extension of the general fatigue analysis (e.g., Wright Agtarap, 2015; Wright & Mlynski, 2019), the amount of self-control (i.e., restraint intensity) one can or, more precisely, is willing to invest in a given task is not dependent on temporarily available self-control resources. Rather it is a function of perceived fatigue, task difficulty (i.e., the magnitude of an unwanted urge), and success importance (i.e., the importance of resisting the urge), with associated cardiovascular responses following (i.e., changes in systolic and diastolic blood pressure as well as mean arterial pressure; Wright et al., 2012). Therefore, fatigue does not automatically lead to less effort or impaired self-control performance (e.g., Wright et al., 2013). For instance, if a fatigued person thinks that success in an upcoming task is highly unlikely and that success is not especially important, he or she is unlikely to invest high amounts of effort which will eventually lead to impaired performance. However, if the same person views success in the upcoming task as likely and important, he or she will be willing to invest more effort and perform at a higher level. Assessing these additional psychological and physiological parameters specified in the process model as well as the behavioral restraint extension of the general fatigue analysis might shed some light on the actual mechanisms contributing to our present pattern of results. Taken together, even though we did not find statistically significant differences between the control and the depletion condition in accuracy rates, we do consider the present findings to be highly informative. First, they suggest that participants are not necessarily worse at detecting lies in sports when in a state of ego-depletion. Second, the present findings suggest fruitful avenues for further research (e.g., different manipulations for systematic and heuristic processing). Third, it adds to the recent discussion surrounding the ego depletion effect, indicating that systematic information processing might be less prone to be affected by states of ego depletion. Fourth, it highlights the necessity to dig deeper into the psychological and physiological mechanisms potentially affecting self-control performance.

6 Tables

6.1 Table 1

Mean veracity ratings for the true and false statements, separated by condition (depletion vs. control)

Statement	Depletion condition		Control condition	
	M	*SD*	*M*	*SD*
True	5.99	0.81	5.83	0.98
False	5.39	0.73	5.32	0.98

Note. $N = 27$. Each video was rated on a continuous scale ranging from 1 (*not at all true*) to 10 (*totally true*).

7 Figures

Figure 1. Illustration of the experimental setup for the generation of the stimulus material. The player wearing the jacket is a confederate acting as an attacking player, the player wearing the white jersey is a confederate acting as the teammate of the attacking player, and the player wearing the black jersey is the target player acting as the defender. The referee is standing on the right, observing the scene.

8 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

9 Author Contributions

CE, AD, and GS equally contributed to the conceptualization of the study and review of relevant related work. CE, GS, and AD analysed and interpreted the data. CE and GS prepared the draft

manuscript, while AD provided the critical revisions. All authors approved the final version of the manuscript and agreed with the order of presentation of the authors.

10 Funding

The authors did not receive any specific funding for this research.

11 References

Aamodt, M. G., and Custer, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examin. 15*, 6–11.

Akehurst, L., Koehnken, G., Vrij, A., and Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behavior. *Appl. Cogn. Psychol. 10*, 461–471. doi:10.1002/(SICI)1099-0720(199612)10:6<461::AID-ACP413>3.0.CO;2-2

André, N., Audiffren, M., & Baumeister, R. F. (2019). An integrative model of effortful control. *Frontiers in Systems Neuroscience*, *13*, 79. doi:10.3389/fnsys.2019.00079

Aragão e Pina, J., Passos, A., Araújo, D., and Maynard, M. T. (2018). Football refereeing: an integrative review. *Psychol. Sport Exerc. 35*, 10–26. doi:10.1016/j.psychsport.2017.10.006

Audiffren, M., & André, N. (2015). The strength model of self-control revisited: Linking acute and chronic effects of exercise on executive functions. *Journal of Sport and Health Science*, *4*, 30-46. doi:10.1016/j.jshs.2014.09.002

Baumeister, R. F., Bratslavsky, E., Muraven, M., Tice, D. M. (1998). Ego depletion: is the active self a limited resource? *J. Pers. Soc. Psychol. 74*, 1252–1265. doi:10.1037/0022-3514.74.5.1252

Baumeister, R. F., Sparks, E. A., Stillman, T. F., and Vohs, K. D. (2008). Free will in consumer behavior: self-control, ego depletion, and choice. *J. Consum. Psychol. 18*, 4–13. doi:10.1016/j.jcps.2007.10.002

Beckmann, J. (2020). Achievement motivation and motivational and volitional processes in sports. *Motiv. Sci. 6*, 192–194. https://doi.org/10.1037/mot0000197

Blázquez, D., Botella, J., and Suero, M. (2017). The debate on the ego-depletion effect: evidence from meta-analysis with the p-uniform method. *Fron. Psychol. 8*, 197. doi:10.3389/fpsyg.2017.00197

Bluemke, M., Brand, R., Schweizer, G., and Kahlert, D. (2010). Exercise might be good for me, but I don't feel good about it: Do automatic associations predict exercise behavior? *Journal of Sport and Exercise Psychology 32*, 137–153. doi:10.1123/jsep.32.2.137

Boat, R., Hunte, R., Welsh, E., Dunn, A., Treadwell, E., and Cooper, S. B. (2020). Manipulation of the duration of the initial self-control task within the sequential-task paradigm: effect on exercise performance. *Frontiers in Neuroscience 14*, 1093. doi:10.3389/fnins.2020.571312

Bond, C. F., and DePaulo, B. M. (2006). Accuracy of deception judgments. *Pers. Soc. Psychol. Rev. 10*, 214–234. doi:10.1207/s15327957pspr1003_2

Bray, S. R., Graham, J. D., Ginis, K. A. M., and Hicks, A. L. (2012). Cognitive task performance causes impaired maximum force production in human hand flexor muscles. *Biol. Psychol. 89*, 195–200. doi:10.1016/j.biopsycho.2011.10.008

Bullard, S. E., Fein, D., Gleeson, M. K., Tischer, N., Mapou, R. L., and Kaplan, E. (2004). The Biber cognitive estimation test. *Arch. Clin. Neuropsych. 19*, 835–846. doi:0.1016/j.acn.2003.12.002

Chaiken, S., and Maheswaran, D. (1994). Heuristic processing can bias systematic processing: effects of source credibility, argument ambiguity, and task importance on attitude judgment. *J. Pers. Soc. Psychol. 66*, 460−473. doi:10.1037/0022-3514.66.3.460

Chaiken, S., and Trope, Y. (1999). *Dual-process Theories in Social Psychology.* New York: Guilford Press.

Clarkson, J. J., Hirt, E. R., Jia, L., and Alexander, M. B. (2010). When perception is more than reality: The effects of perceived versus actual resource depletion on self-regulatory behavior. *Journal of Personality and Social Psychology 98*, 29–46. doi:10.1037/a0017539

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Routledge Academic.

Davis, D., and Leo, R. A. (2012). Interrogation-related regulatory decline: ego depletion, failures of self-regulation, and the decision to confess. *Psychol. Public Policy, Law. 18*, 673–704. doi:10.1037/a0027367

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., and Cooper, H. (2003). Cues to deception. *Psychol. Bull. 129*, 74–118. doi:10.1037/0033-2909.129.1.74

De Ridder, D. T., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., and Baumeister, R. F. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors. *Personality and Social Psychology Review*, *16*, 76-99. doi:10.1177/1088868311418749

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*:781. doi:10.3389/fpsyg.2014.00781

Englert, C. (2017). Ego depletion in sports: highlighting the importance of self-control strength for high-level sport performance. *Curr. Opin. Psychol. 16*, 1–5. https://doi.org/10.1016/j.copsyc.2017.02.028

Englert, C. (2019). The self-regulation of human performance: a critical discussion and future directions for self-control research. *Perform. Enhanc. Health, 6*, 156–157. doi:10.1016/j.peh.2019.04.001

Englert, C., and Bertrams, A. (2014). What is self-control depleting in sports? Effects of vicarious experiences on performance. *International Journal of Sport Psychology 45*, 1-10.

Englert, C., and Rummel, J. (2016). I want to keep on exercising but I don't: The negative impact of momentary lacks of self-control on exercise adherence. *Psychology of Sport and Exercise*, *26*, 24-31. doi:10.1016/j.psychsport.2016.06.001

Englert, C., and Schweizer, G. (2020). "Are You Telling the Truth?"—Testing individuals' ability to differentiate between truth and deceit in soccer. *Front. Psychol. 11*, 1082. doi:10.3389/fpsyg.2020.01082

Eronen, M. I., and Bringmann, L. F. (2021). The theory crisis in psychology: How to move forward. *Perspectives on Psychological Science.* doi:10.1177/17456916209705

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods. 39*, 175–191. doi:10.3758/BF03193146

Feeley, T. H., and DeTurck, M. A. (1995). Global cue usage in behavioral lie detection. *Commun. Q.* 43, 420–430. doi:10.1080/01463379509369989

Forrest, J. A., Feldman, R. S., and Tyler, J. M. (2004). When accurate beliefs lead to better lie detection. *J. Appl. Soc. Psychol. 34*, 764–780. doi:10.1111/j.1559-1816.2004.tb02569.x

Furley, P., Schweizer, G., and Bertrams, A. (2015). The two modes of an athlete: dual-process theories in the field of sport. *Int. Rev. Sport Exerc. Psychol. 8*, 106-124. doi:10.1080/1750984X.2015.1022203

Güldenpenning, I., Kunde, W., and Weigelt, M. (2017). How to trick your opponent: a review article on deceptive actions in interactive sports. *Front. Psychol. 8*, 917. doi:10.3389/fpsyg.2017.00917

Hsu, L. (1997). C*heating and sports rules. International Olympic Academy: 7th Postgraduate Session.* Greece.

Hagger, M. S., Chatzisarantis, N. L., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... and Calvillo, D. P. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspect. Psychol. Sci. 11*, 546–573. doi:10.1177/1745691616652873

Hagger, M. S., Wood, C., Stiff, C., and Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychol. Bull. 136*, 495–525. doi:10.1037/a0019486

Inzlicht, M., and Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science 7*, 450–463. doi:10.1177/1745691612454134

Inzlicht, M., and Schmeichel, B. J. (2016). Beyond limited resources: Self-control failure as the product of shifting priorities. In K. D. Vohs and R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (3rd ed., pp. 165–181). Guilford Press.

Kavussanu, M., and Ntoumanis, N. (2003). Participation in sport and moral functioning: does ego orientation mediate their relationship? *J. Sport Exerc. Psychol. 25*, 501–518. doi:10.1123/jsep.25.4.501

Loschelder, D. D., and Friese, M. (2016). "Moderators of the ego depletion effect," in *Self-regulation and ego control*, eds. E. R. Hirt, J. J. Clarkson, and L. Jia (San Diego, CA: Elsevier Academic Press), 21–42. 10.1016/B978-0-12-801850-7.00002-0

MacMahon, C., Helsen, W. F., Starkes, J. L., & Weston, M. (2007). Decision-making skills and deliberate practice in elite association football referees. *Journal of Sports Sciences*, *25*, 65-78.

Marksteiner, T., Reinhard, M. A., Lettau, F., and Dickhäuser, O. (2013). Bullying, cheating, deceiving: teachers' perception of deceitful situations at school. *Int. J. Educ. Psychol. 2*, 193–220. doi:10.4471/ijep.2013.24

Masip, J., Garrido, E., and Herrero, C. (2009). Heuristic versus systematic processing of information in detecting deception: questioning the truth bias. *Psychol. Rep. 105*, 11–36. doi:10.2466/PR0.105.1.11-36

Moore, L. J., Harris, D. J., Sharpe, B. T., Vine, S. J., & Wilson, M. R. (2019). Perceptual-cognitive expertise when refereeing the scrum in rugby union. *Journal of Sports Sciences*, *37*, 1778–1786. doi:10.1080/02640414.2019.1594568

Morris, P. H., and Lewis, D. (2010). Tackling diving: the perception of deceptive intentions in association football (soccer). *J. Nonverbal Behav. 34*, 1–13. doi:10.1007/s10919-009-0075-0

Pageaux, B., Lepers, R., Dietz, K. C., and Marcora, S. M. (2014). Response inhibition impairs subsequent self-paced endurance performance. *Eur. J. Appl. Physiol. 114*, 1095–1105. doi:10.1007/s00421-014-2838-5

Petty, R. E., Barden, J., and Wheeler, S. C. (2009). "The elaboration likelihood model of persuasion: developing health promotions for sustained behavioral change," in *Emerging theories in health promotion practice and research* (2nd ed.), eds. R. J. DiClemente, R. A. Crosby and M. C. Kegler (San Francisco, CA: Jossey-Bass), 185–214.

Petty, R. E., Cacioppo, J. T., Strathman, A. J., and Priester, J. R. (2005). "To think or not to think: Exploring two routes to persuasion," in *Persuasion: psychological insights and perspectives* (2nd ed.), eds. T. C. Brock & M. C. Green (Thousand Oaks, CA: Sage), 81–116.

Reinhard, M. A., Scharmach, M., and Stahlberg, D. (2013). Too exhausted to see the truth: ego depletion and the ability to detect deception. *Brit. J. Soc. Psychol. 52*, 618–630. doi:10.1111/j.2044-8309.2012.02113.x

Renden, P. G., Kerstens, S., Oudejans, R. R., and Cañal-Bruland, R. (2014). Foul or dive? Motor contributions to judging ambiguous foul situations in football. *Eur. J. Sport Sci. 14* (sup1), S221–S227. doi:10.1080/17461391.2012.683813

Sabag, E., Lidor, R., Morgulev, E., Arnon, M., Azar, O., and Bar-Eli, M. (2018). To dive or not to dive in the penalty area? The questionable art of deception in soccer. *Inter. J. Sport Exerc. Psychol.* 1–12. doi:10.1080/1612197X.2018.1462100

Samuel, R. D., Englert, C., Zhang, Q., and Basevitch, I. (2018). Hi ref, are you in control? Self-control, ego-depletion, and performance in soccer referees. *Psychol. Sport Exerc. 38*, 167–175. doi:10.1016/j.psychsport.2018.06.009

Tangney, J. P., Baumeister, R. F., and Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality 72*, 271–324. Doi: 10.1111/j.0022-3506.2004.00263.x

Unger, A., and Stahlberg, D. (2011). Ego-depletion and risk behavior. *Soc. Psychol. 42*, 28–38. doi:10.1027/1864-9335/a000040

Vrij, A., Granhag, P. A., and Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychol. Sci. Public Interest 11*, 89–121. doi:10.1177/1529100610390861

Vohs, K. D., Schmeichel, B. J., Gronau, Q. F., Finley, A., Wagenmakers, E. J., … and Albarracín, D. A. (in press). Multi-site preregistered paradigmatic test of the ego depletion effect. *Psychol. Sci.*

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2018a). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review 25*, 58–76. doi:10.3758/s13423-017-1323-7

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., … Morey, R. D. (2018b). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review 25*, 35–57. doi:10.3758/s13423-017-1343-3

Wheeler, S. C., Brinol, P., and Hermann, A. D. (2007). Resistance to persuasion as self-regulation: ego-depletion and its effects on attitude change processes. *J. Exp. Soc. Psychol. 43*, 150–156. doi:10.1016/j.jesp.2006.01.001

Wolff, W., Sieber, V., Bieleke, M., and Englert, C. (2021). Task duration and task order do not matter: no effect on self-control performance. *Psychological Research* 85, 397–407. 10.1007/s00426-019-01230-1

Wright, R., and Agtarap, S. (2015). The intensity of behavioral restraint: Determinants and cardiovascular correlates. In G. Gendolla, M. Tops, & S. Koole (Eds.), *Handbook of Biobehavioral Approaches to Self-Regulation* (pp.287-299). Springer.

Wright, R. A., and Mlynski, C. (2019). Fatigue determination of inhibitory strength and control: A babe in a bath. *Motivation Science 5*, 66-78. doi:10.1037/mot0000114

Wright, R. A., Patrick, B. M., Thomas, C., and Barreto, P. (2013). When fatigue promotes striving: Confirmation that success importance moderates resource depletion influence on effort-related cardiovascular response. *Biological Psychology 93*, 316–324. doi:10.1016/j.biopsycho.2013.02.016

Wright, R. A., Shim, J. J., Hogan, B. K., Duncan, J., and Thomas, C. (2012). Interactional influence of fatigue and task difficulty on cardiovascular response: Demonstrations involving an aerobic exercise challenge. *Psychophysiology 49*, 1049–1058.

12 Data Availability Statement

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.