Cognitive Reflection and Moral Reasoning

Ana Proroković and Ljiljana Gregov

University of Zadar, Department of Psychology, Croatia

ABSTRACT

The goal of this study was to examine the relationship between reflectivity/impulsivity and aspects of moral reasoning (general level and individual stages) while considering assessment times and relevance of moral arguments. The study involved 442 participants (163 female and 279 male) aged between 19 and 76, with different levels of education. The study was conducted online and two measuring instruments were applied: the cognitive reflection test and the test of moral reasoning The obtained results showed that problem solving time was significantly shorter for intuitive answers as opposed to correct answers. Predominantly reflective and predominantly impulsive individuals differed in various aspects concerning problem solving and the assessment of moral arguments. Predominantly impulsive individuals demonstrated: significantly longer problem solving time for correct answers (there were no differences for intuitive answers), lower general level of moral reasoning, longer assessment time, and higher assessment of the relevance of moral arguments (sensitivity to argument strength) in almost all stages of moral development. The results suggest that there are different ways in which dominant cognitive styles determine the effects in tasks of different types.

KEYWORDS

cognitive reflection test
moral reasoning
reflectivity/impulsivity
dual process model

INTRODUCTION

In some earlier studies (Sternberg & Grigorenko, 1977; Witkin et al., 1977), cognitive styles are described as relatively stable, general perceptual and cognitive characteristics that determine an individual's approach to thinking, learning, and problem solving in different contexts. That is, they represent individual preferences with regards to information processing. Sternberg and Grigorenko (1997) recognized three different approaches to studying cognitive styles that have been previously applied in research: the cognition-centered approach, the personality-centered approach, and the activity-centered approach. The first emphasizes the characteristics of the cognitive processes as determinants of cognitive styles (e.g., reflective and impulsive styles). The second is oriented towards the interrelationships between cognitive styles and personality traits (e.g., perceptive, emotional, reflective, or intuitive). The third approach focuses on the practical application, where cognitive styles are observed as mediator variables. The applied research has been typically conducted within the educational domain and it has confirmed the existence of different learning and teaching styles (e.g., intuitive, analytic, and integrated style). The current study adopted the framework of the cognition-centered approach, emphasizing the role of information processing and decision making. In this context, most authors start from the so-called dual process theories/models, according to which decision making is the result of the interaction between two cognitive systems (Gawronski & Creighton, 2013). These two systems are most often described as intuitive and analytical (Hammond, 1996), heuristic and analytical (Evans, 2006), affective and intentional (Stanovich & West, 2000) or simply as Systems 1 and 2 (Greene, 2009; Kahneman & Frederick, 2002). The former is characterised by fast decision making based on habits and experiences. Such decisions are made automatically, unconsciously, and are mostly driven by emotion. The latter is characterised by slow decision making and need for cognition, it is focused on goals, controlled, and conscious, that is, reflective.

Although studies in the field of cognitive psychology and neuroscience agree on the existence of two separate systems (Baron et al., 2015; Gawronski & Creighton, 2013), there is still a lack of empirically vali-

Corresponding author: Ana Proroković, University of Zadar, Department of Psychology, Krešimirova obala 2, Zadar 23000, Croatia.
E-mail: aprorok@unizd.hr

dated models that can explain the dynamic aspect of their interaction. Typically, there are two basic viewpoints. According to the first, which has been often empirically tested, the two systems work sequentially. System 1 always starts first (led by heuristics) and as a result, gives an intuitive answer. System 2, which always intervenes after System 1, can modify the intuitive answer with another, rational answer (Evans, 2006; Kahneman & Frederick, 2002). From the standpoint of sequential dual process theory, many cognitive tasks include a fast, intuitive process. The intuitive judgement is instantaneous and automatic, and does not include cognitive reflection (effort) other than the one needed for understanding the given task. Occasionally, that process is followed by a slow and reflective reasoning, that is, System 2 activates, and corrects the errors which come from the System 1 (Kahneman, 2011). Therefore, answering correctly implies a longer answering time, because it also includes the time of correcting the initial, intuitive answer. An alternative standpoint assumes that these two systems can operate simultaneously and be mutually competitive (e.g., Sloman, 1996). That is, System 1 competes with System 2 and the "winner" determines the answer. Seeing as System 1 is naturally faster, it should always win over System 2. However, if the reflective (correct) and the intuitive (wrong) answers are two competitive responses, and individuals are biased while selecting one of them (by their cognitive style), the frequency of the typical responses and the answering time should depend on the preferred cognitive style (shorter time relates to more frequent responses). It should be noted that the probability of individual answers will depend on some characteristics of the question/task (e.g., their difficulty and familiarity). Therefore, according to this approach, these are two simultaneous processes that lead to different answers in which the faster process usually wins. However, when it loses, it is unusually slow, which makes the decision making time longer (when the predominantly intuitive individual gives a reflective response and vice versa).

The cognitive reflection test (CRT) by Frederick (2005) has proven to be a useful measure in researching individual differences in thinking, reasoning, and decision making. The test consists of three questions that provoke wrong, intuitive answers (lures) by activating System 1, while the correct answer requires the activation of System 2. Numerous studies have shown that the CRT scores moderately correlate with cognitive abilities (Stanovich & West, 2008), as well as with many other constructs such as biases in reasoning and decision making (Campitelli & Labollita, 2010), risk preferences (Andersson et al., 2016), religiosity and belief in the supernatural (Gervis & Norenzayan, 2012), utilitarian moral judgments (Paxton et al., 2011) and others. Regarding the stated CRT correlates, Cokely and Kelley (2009) argue that it can serve as a valid measure of cognitive styles, that is, reflectivity/impulsivity as a trait. Individuals that are more reflective solve the CRT correctly but more slowly, while the more impulsive individuals solve it faster, but not as correctly, so it can be assumed that a similar pattern of cognitive activity will be present in other areas of deliberation and decision making. A crucial hypothesis of a dual-process reasoning model implies the individuals' willingness to check or reconsider their initial answer and change it (Baron et al., 2015). Many psychologists (Baron et al., 2015; Gawronski & Creighton, 2013) think that this disposition is the most important characteristic of rational thinking and even intelligence.

Despite some of the controversies regarding previous studies of these two systems, it can be concluded that the distinction between fast, intuitive answers and reflective answers is rather clear and relevant for many interpretations of reasoning and assessing in general. Therefore, many approaches to moral reasoning and decision making utilize different versions of dual-process models. For example, results of numerous studies (Campitelli & Labollita, 2010; Paxton et al., 2011). confirm that CRT scores correlate with utilitarian moral judgement in cases where a moral dilemma presupposes a conflict between the prohibition of certain rule-based actions (conditioned by social norms) and the utilitarian assessment of the overall consequences (e.g., killing one person in order to save five people). The main explanations for this relationship imply that the utilitarian answer prevails over the prohibition-based intuitive answer (Paxton et al., 2011), which is in accordance with the sequential approach. This approach cannot adequately explain some of the results related to the effect of difficulty and type of moral dilemmas on response time (Baron et al., 2015). For example, in the case of more difficult moral dilemmas, the time of deontological responses is prolonged, and limiting one's time for responding reduces the number of utilitarian judgments. Furthermore, cognitive interference (when the individual makes qualitative and quantitative assessments of consequences) prolongs the time of utilitarian answers. Finally, the correlation between utilitarian answers and CRT scores can be the result of individual differences in reflectivity/impulsivity (as a personality trait), that is, the individual's tendency to predominantly focus on accuracy or speed. Such results could be interpreted within the so-called parallel dual-process model of decision making.

Nonetheless, previous studies on CRT and moral reasoning have mainly used utilitarian moral dilemmas involving a certain type of intuitive response (prohibition). However, there is a lack of studies that place CRT in relation to other aspects of moral reasoning (e.g., general level of moral reasoning and phases of moral development) that are consistent with the predominant cognitive developmental moral theories (Kohlberg's and neo-Kohlbergian approaches). In order to assess the stages of moral development, various moral arguments, which do not have clear characteristics of intuitive answers, are most commonly derived to justify (or not) certain behaviours. According to Kohlberg (1973), people progress through three phases of moral thinking (preconventional, conventional, and postconventional phase) that build on cognitive development, where each phase is characterized by two developmental stages (there is a total of six developmental stages, the first of which represents the lowest level, while the sixth stage represents the highest level of moral reasoning). The preconventional or egocentric phase includes two stages, and in it, moral behaviour is focused on obedience, avoiding the punishment and the inability to take into account other peoples' perspectives (first stage). In the second stage, the earliest form of moral reciprocity (personal interest and exchange) appears. The rules are followed as long as they are in line with personal interest, and morality is based on equal exchange (Hren, 2008). In the conventional phase, reasoning is related to social perspective and leg-

islation. The third stage of moral reasoning is within the framework of social conformity. Morality is in line with the expectations of the social environment and is determined in terms of stereotypical roles, while on the fourth stage (the so-called normative stage), morality is defined by laws and norms. Finally, the third general phase is the postconventional phase, and it is based on the concept of universal justice. At stage five (the so-called stage of fundamental justice and social contract), law and morality cannot be equalized because the law itself is imperfect, and the sixth and highest stage of moral reasoning implies the general principles of justice which are above society and law, where morality is considered to be the goal, not a means.

The goal of the current study was to examine whether CRT is predictive of some moral reasoning aspects when there are no clear intuitive answers, and whether reflectivity/impulsivity is related to the general level and individual stages of moral reasoning.

The first general hypothesis concerned the fact that information processing within System 1 and System 2 at the manifest level differs in the type and time of responses, which would be reflected in CRT performance. Accordingly, individuals could be differentiated with regards to the frequency of their use of these two systems as predominantly impulsive and predominantly reflective ones.

The second general hypothesis was that the dominant cognitive style would also reflect on some aspects of moral judgments, or more specifically, on the assessment time and sensitivity to the strength of moral arguments. Although arguments at certain stages of moral reasoning do not have clear intuitive answers (lure), some arguments are experientially more recognizable (especially at lower levels), so it could be expected that they will be processed predominantly within System 1. Contrary to this, the complex and less experiential arguments at higher levels would more probably be processed within System 2. Therefore, it was assumed that predominantly impulsive individuals would assess the arguments corresponding to the lower stages of moral development as more important than predominantly reflective ones. On the other hand, predominantly reflective individuals would judge arguments as more important at higher stages of moral development (analytic approach, need for cognition). These differences could have an indirect effect on the general measure of moral reasoning in favor of predominantly reflective individuals. Due to the more frequent use of System 1, it can generally be hypothesized that predominantly impulsive individuals would be faster in assessments of the argument strength than predominantly reflective ones, at all stages of moral development.

METHODS

Participants

Out of a total of 1009 subjects that initially participated in the study, data of 442 (279 male, 163 female) were retained for the final analysis. Participants who did not finish all the tests, whose solving time was longer than an hour, and those who did not show variability in their answers have been excluded from the final sample. The reasons for this rather large dropout rate most likely related to the fact that the test of moral reasoning was demanding in terms of cognitive engagement, and the participants were not motivated enough to give careful and consistent answers (approximately 40% of the sample did not complete the test or they often took breaks longer than 15 minutes). Such a dropout rate is not uncommon in similar studies in which participants solve moral dilemmas. For example, Doyle and O'Flaherty (2013) reported an overall response rate of 48% for the defining issues test (DIT) applied together with others measures. The additional dropout on the basis of consistency checks was 16%, which is in line with other studies (Rest, 1990; William, 2004). The sample was culturally and nationally heterogeneous (EU citizens), with an age range between 19 and 76 (M = 34.52; SD = 13.12) and with different levels of education. The distribution according to the levels of education within the sample corresponded to the expectations within the population.

COGNITIVE REFLECTION TEST

The CRT originally contained three tasks (Friederick, 2005). In addition to the usual correct and incorrect answers, there are also the so-called intuitive responses present in the tasks. However, due to the frequent use of this test, some studies showed that nearly half of the participants have already seen or solved the test. The familiarity of the test frequently affected the results, and with it, the test's validity (Haigh, 2016). In order to reduce this test familiarity effect, new alternative tests of cognitive reflection were developed, containing larger numbers of tasks of the same type (Primi et al., 2016: Toplak et al., 2014). Their correlations with the results of the original test are moderate to high.

Therefore, an extended version of the CRT was also used in the current study, which included the six following tasks (the first three tasks refer to Frederick's original version):

1. "A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost?" (correct answer is $0.05; intuitive answer is $0.1)
2. "If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?" (correct answer is 5 minutes; intuitive answer is 100 minutes)
3. "In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?" (correct answer is 47 days; intuitive answer is 24 days)
4. "The merchant reduced the price of a pair of shoes that cost 100$ by 10%. The next week he reduced the price by another 10%. What is the price of the shoes now?" (correct answer is $81; intuitive answer is $80)
5. "If you divide 30 by ½ and add 10, what will you get?" (correct answer is 70; intuitive answer is 25)
6. "The doctor gave you 3 tablets and told you to take one every half hour. How long before you take all three tablets if you take the first one right away?" (correct answer is 1 hour; intuitive answer is 1 hour and 30 minutes)

A confirmatory factor analysis confirmed the expected unidimensional structure of this version of the CRT (see Table 1, Figure 1), satisfying internal consistency (Cronbach's α = 0.768), and a relatively high item correlations with the total test result (0.44-0.57).

TEST OF MORAL REASONING

The test of moral reasoning (TMR, Proroković, 2016) consists of two moral dilemmas in which the character from the story makes a certain decision. For every decision made, there are six pro and six contra arguments that would either justify or not the decision the character made. The arguments are, by their content, adapted to Kohlberg's stages of moral development (two stages of each level: the preconventional, conventional, and postconventional level). The participants assessed the extent to which the proposed arguments are acceptable/unacceptable on a six-point scale (without the possibility of a neutral answer). The test uses an index of moral reasoning (IMR) as a general measure of the level of moral reasoning that is based on the deviation from the "optimal profile." The premise for defining the optimal profile starts with the assumption that the person with the highest level of moral reasoning will assess the argument on the lowest stage to be the least acceptable, and the arguments on each subsequent stage to be more acceptable by a point. The IMR represents a parameter that ranges from 0 to 1, where a lower result indicates a lower level of moral reasoning, and a higher result indicates a higher level. In previous studies, this measure showed very good criterion validity and construct validity (Proroković, 2016).

Procedure

The study was conducted online using PsyToolkit (Stoet, 2017). It was voluntary and anonymous in nature. The invitation to participate in the study came with the link posted on different social media platforms, and a condition that participants should be older than 17 years of age. Before the tests were presented, the participants were informed of the primary goals of the study, and permission was requested for the use of their data for research purposes. It should be noted that in the online version of the TMR, certain pro and contra arguments were presented successively (one by one); the same applies to the CRT test.

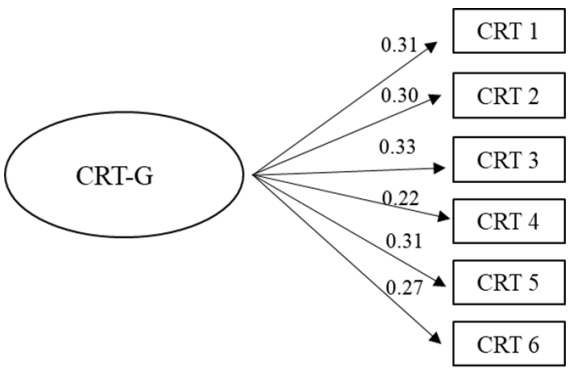

FIGURE 1.
Path diagram for the confirmatory factor analysis of CRT with standardized regression weights.

There were no possibilities of skipping a task/question or returning to previous tasks. Questions in both tests were randomly rotated, and the participants filled out the sociodemographic questionnaire after finishing both tests. Each answer was recorded, as well as the time the participants took to provide it. Before the statistical analysis of the collected data, the time of each answer on the CRT and the TMR was relativized and reduced with respect to the average time needed for reading the tasks and arguments (due to a different number of words in each of the arguments), and the extreme results were excluded. First, the average reading time per character/letter for each participant was calculated with respect to the reading time needed for the first instruction. Afterwards, the revised time was calculated for each of the tasks/arguments by subtracting the average reading time multiplied by the number of corresponding characters /letters for each argument/task from the total time (from the beginning of the presentation of each task/argument to the participants's reaction).

RESULTS

Problem Solving Times and Cognitive Styles

Previous studies in this area have regularly classified the participants with respect to their scores on the CRT, that is, either as dominantly reflective or dominantly intuitive. Thus, in the initial analysis, we sought to answer the question of whether such a classification is justified, not only when it comes to the total number of correct answers, but also the corresponding problem solving time. In the first step of the analysis, the differences in problem solving time between correct and intuitive answers in the CRT were examined using an analysis of variance (ANOVA). As expected, the participants demonstrated a significantly shorter solving time for intuitive answers as opposed to correct answers, $F(2, 206) = 16.83$, $p < .001$, $n_p^2 = .140$, with a relatively large effect size. Furthermore, the average time needed for responding was the longest for nonintuitive wrong answers (see Figure 2). A post-hoc analysis (Bonferroni test, pairwise comparisons) showed that all differences between the categories were statistically significant ($p < .01$). It can be assumed that correct and incorrect answers are two different outcomes of the same analytical, that is, reflective thinking, which does not necessarily involve overcoming the first, intuitive answer.

TABLE 1.
Single Sample Goodness of Fit Indices (unidimensional model of the Cognitive Reflection Test)

Index	Value
ML Chi-Square	24.887 (*df* = 9)
RMS standardized residual	0,034
Steiger Lind RMSEA	0.038
McDonald non centrality index	0.957
Bollen's Rho	0,921

Note. ML = maximum likelihood; RMS = root mean square; RMSEA = root mean square error of approximation.

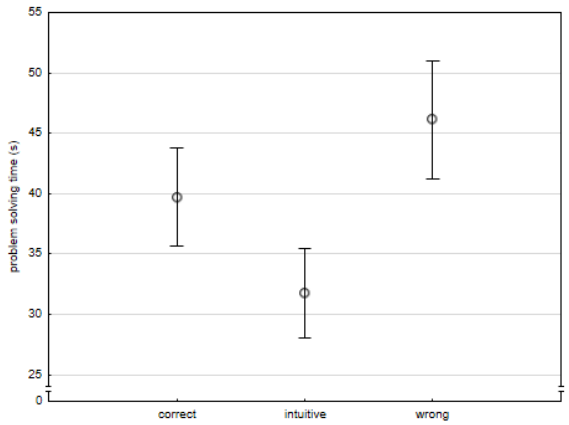

FIGURE 2.
Differences in problem solving times for correct, wrong and intuitive answers (means, 95% confidence intervals).

In subsequent analyses, emphasis was placed on the possible differences in different parameters of dependent variables considering the primary cognitive systems determined by the score on the CRT. The participants were divided into two groups: predominantly reflective and predominantly impulsive individuals. Those participants who achieved a minimum of 4 correct answers were classified as predominantly reflective ones ($N = 211$), while those who achieved a minimum of 3 intuitive answers ($N = 171$) were classified as predominantly impulsive ones (wrong answers were not taken into account). Basic demographic information for these two groups is given in the Table 2. The results indicated a predominance of male participants in the group of predominantly reflective individuals, $\chi^2 = 12.72$, $df = 1$, $p < .001$. Furthermore, the group contained significantly more participants with a higher level of education, $\chi^2 = 15.99$, $df = 4$, $p = .003$, while no significant differences were found with respect to the age of participants, $F = 3.33$, $df = 1/373$, $p = .068$.

It should be noted that a possible reason for the higher number of reflective individuals is partly methodological in nature, that is, due to online test application (our inability to control whether problem solving was done independently) as well as the difficulty of tasks (which were relatively easy). In this categorization, we intentionally disregarded the wrong answers (those that were not intuitive answers), which were sometimes treated as equivalents of intuitive answers by other authors (Jimenez et al., 2018). This is fully supported by the results shown in Figure 2, where it is evident that different cognitive processes can be identified behind the intuitive and wrong answers

The next step in the analysis was to compare the potential differences in problem solving time between predominantly reflective and predominantly impulsive individuals. They were analysed separately (ANOVA) in regard to the total problem solving time and in regard to the problem solving time related to correct and intuitive answers. The reason we opted for a separate ANOVA and not a factorial ANOVA was the fact that there was a considerable number of participants who provided no intuitive answers (the predominantly reflective ones), just as there was a number of participants who did not provide any correct answers (predominantly intuitive ones). There were no significant differences in the total and average problem solving time between predominantly reflective and predominantly impulsive participants, $F(1, 380) = .001$, $p = .984$, $n_p^2 < .001$. However, when analyzing correct and intuitive answers separately (see Figure 3), predominantly impulsive participants were significantly slower with providing correct answers compared to reflective participants, $F(1, 334) = 5.57$, $p = .019$, $n_p^2 = .017$), with a relatively small effect size. All other differences (intuitive answers) were not significant, $F(1, 275) = .20$, $p = .658$, $n_p^2 < .001$.

Moral Reasoning and Cognitive Styles

The main study goals were answered in the following analysis, where the participants with different cognitive styles were tested for differences regarding certain aspects of moral reasoning. Potential differences were confirmed in the general level of moral reasoning (IMR) between participants who can be categorized as predominantly reflective and those who can be categorized as predominantly intuitive ones (ANOVA). There was a significant difference in the general level of moral reasoning, $F(1, 309) = 4.94$, $p = .027$, $n_p^2 < .016$, in favour of reflective participants (MR = .44, SDR = .08; MI = .42, SDI = .08). To obtain more precise insight into these differences, the effect size calculated, and it showed that this difference was not statistically strong, that is, the predominantly cognitive type of reasoning (reflective vs. impulsive) explained only 1,57% of total moral reasoning variance.

In the next step, we conducted the analysis of potential differences in the assessments of moral arguments strengths, using a repeated-meas-

TABLE 2.
Basic Demographic Information for the Predominantly Reflective and Predominantly Impulsive Individuals

		Reflective	Impulsive	Total sample
Gender (%)	Male	72.51	54.97	64.66
	Female	27.49	45.03	35.34
Level of education (%)	Primary school	5.21	14.62	9.42
	Secondary school	38.86	43.86	41.10
	Bachelor	21.80	19.88	20.94
	Master	29.86	20.47	25.65
	Phd	4.27	1.17	2.88
Age (M, SD)		36.13 (12.83)	33.65 (13.35)	34.56 (13.13)

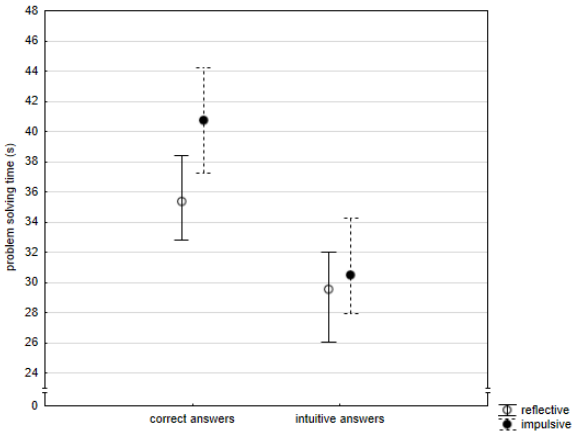

FIGURE 3.

Differences in problem solving times (s) between predominantly reflective and predominantly impulsive individuals considering the correct and intuitive answers (means, 95% confidence intervals).

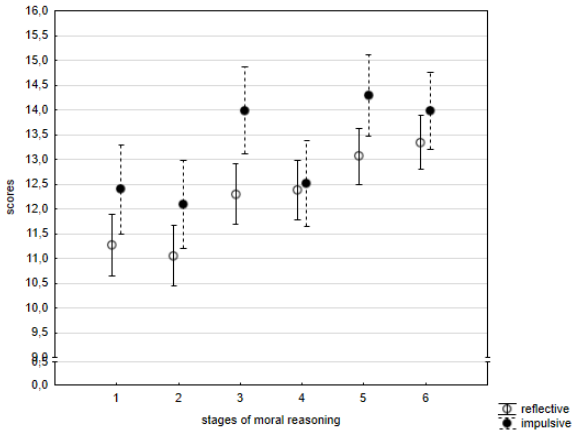

FIGURE 4.

Differences in the assessments of moral arguments between predominantly reflective and predominantly impulsive individuals (means, 95% confidence intervals).

ures multivariate analysis of variance (MANOVA). The first factor in the MANOVA was reflectivity/impulsivity (two levels), and the second factor was moral reasoning stages (6 levels). As expected, there was a significant difference in estimating moral arguments considering the predominant reflectivity/impulsivity, $F(1, 1900) = 9.77$, $p = .002$, $n_p^2 = .025$, and considering the stages of moral development, $F(5, 1900) = 28.01$, $p < .001$, $n_p^2 = .069$). Arguments that were on higher stages of moral development were estimated to be more important/stronger (medium effect size) with no significant group × stage interaction, $F(5, 1900) = 1.047$, $p = .388$, $n_p^2 < .001$. Predominantly impulsive individuals estimated arguments as more important at all stages of moral development except at the fourth, the normative stage, with relatively small effect size (see Figure 4).

Further data analysis focused on the time needed to assess the strength of moral arguments using a 2 × 6 MANOVA (reflectivity/impulsivity and moral reasoning stages as independent variables). The results showed a statistically significant difference in the average time needed to assess TMR arguments between predominantly reflective and predominantly impulsive participants, $F(1, 1900) = 15.62$, $p < .001$, $n_p^2 = .039$. The time needed to assess the arguments concerning every stage of moral development in impulsive participants was statistically significantly longer, with a small effect size. There were also statistically significant differences in time needed to assess moral arguments regarding the stages of moral development, $F(5, 1900) = 142.32$, $p < .001$, $n_p^2 = .272$, and the group × stage interaction was statistically significant, $F(5, 1900) = 4.71$, $p < .001$, $n_p^2 = .012$) with a very small effect size. As expected, longer assessment time was connected to later stages of moral development (complex arguments), and shorter time was generally connected to lower stages of moral development with very large effect size (see Figure 5).

Finally, the correlation between problem solving time, total CRT scores, the time needed to assess moral arguments, and the general level of moral reasoning was calculated for the entire sample because the variability of CRT scores is very limited in categorized groups (see Table 3).

DISCUSSION

The hypothesis of the current study was that the difference in the use of the System 1 and System 2 would reflect on task solving time and the type of response. Namely, response time should be longer for correct answers and shorter for intuitive ones. The obtained results are in accordance with the dual-process theory of decision making. Intuitive answers were processed within System 1 (quick decisions and heuristics), and correct and incorrect answers - within System 2 (slower, analytical answers). Also, given the frequency of intuitive or correct answers, that is, the frequency of use of System 1 or System 2, the differentiation into predominately impulsive and predominately reflective individuals was justifiable. Approximately 88% of participants were able to be distinguished within the defined categories, which is in accordance to results of some previous studies (Cintamulya, 2019). On the other hand, there was no certainty in assuming either successive or parallel processing, or that there was some interaction between these two

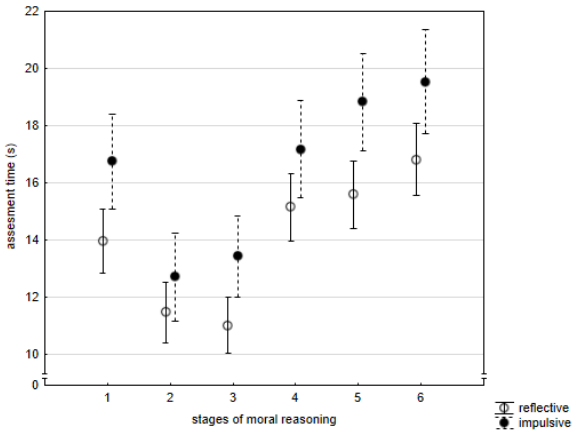

FIGURE 5.

Differences in assessment times of moral arguments between predominantly reflective and predominantly impulsive individuals (means, 95% confidence intervals).

TABLE 3.
Correlation Coefficients Between Problem Solving Time, Total Assessment Time (TMR), IMR, and CRT score

	Assessment time	IMR	CRT score
Problem solving time	0,473**	0,0145	0,070
Total assessment time		-0,022	-0,187*
IMR			0,125*

Note. IMR = index of moral reasoning; CRT = cognitive reflection test.
*$p < .05$ **$p < .01$

systems. For time needed for answering the questions, the sequential theory implies that the intuitive answers will be provided more quickly than correct answers because of the added time that System 2 requires for the correction of the System 1 response. Nonetheless, the probability of specific answers and problem solving times should also depend on some characteristics of each task (e.g., lure strength, task difficulty). Some tasks will mostly provoke an intuitive, biased answer, making the answering time shorter and the number of intuitive answers higher, regardless of the dominant cognitive style (parallel approach). This conflicting hypothesis is consistent with the idea of competition between simultaneous processes leading to different responses (Baron et al., 2015). The fastest process usually wins, but when it loses, the slower processes take over. It seems that the "intuitive answers prevailing" hypothesis (prolonged time for correct answers), is valid only with predominantly impulsive individuals, which is rather in favour of the sequential processing model. Such a conclusion arises from the fact that predominately impulsive individuals showed a longer time for correct responses compared to predominately reflective individuals, but there were no statistically significant differences for intuitive responses. In predominantly reflective individuals, the analytical approach is probably present from the beginning of problem solving, which is rather in favour of the parallel processing model. While the results of Cokely and Kelley (2009) stated that reflective individuals solve the CRT more slowly and more correctly in comparison to impulsive individuals, this was not confirmed in our study. Predominantly reflective individuals showed shorter time of correct responses compared to predominantly impulsive ones. Nevertheless, it should be noted that this analysis had a significantly smaller number of participants (ones that did not have a single correct or intuitive answer were omitted due to the impossibility of comparison), and that we encountered a relatively small number of intuitive answers with reflective participants and vice versa (there were statistically significant differences in number of participants in individual categories). Considering that fact, the obtained results have a limited possibility of generalization, which is why it is more appropriate to talk about tendencies in observed differences in both groups of participants, with the need of additional empirical verification.

The next hypothesis was that dominant cognitive styles would affect some aspects of moral reasoning, that is, predominantly reflective individuals should have a higher level of moral reasoning than predominantly impulsive individuals, most probably as an indirect effect of differences in moral argument strength assessments in lower and higher stages of moral development. It is also possible that the general measure of moral reasoning is influenced by some other characteristic of dominant cognitive styles like investing greater cognitive effort or differences in some cognitive abilities. The obtained results are in accordance with these expectations, considering that Frederick (2005) proved the existence of a moderate positive correlation between the CRT scores and some aspects of intelligence, as well as a high correlation with some mental heuristic measures. It has also been shown that suppressing the first intuitive answer is not the only cause for being successful on the CRT, but that mathematical abilities and orientation towards reflectivity also play a big role (Szaszi et al., 2017). Furthermore, IMR most often shows moderate and positive correlation with different aspects of intellectual abilities (Proroković et al., 2017).

We also hypothesized that individuals with distinct cognitive styles would differ in their argument assessments related to different stages of moral development, and that this would most likely reflect in their assessment of argument strength as well as the time needed for this assessment. This hypothesis was partially confirmed. Predominantly impulsive individuals assessed the arguments corresponding to the lower stages of moral development as more important than predominantly reflective ones. However, predominantly impulsive individuals also judged arguments at higher stages of moral development as more important, which was not expected. Generally, predominantly impulsive individuals had a smaller range of assessments between the importance of lower and higher levels of moral judgment, which confirmed the hypothesized indirect effect on IMR. It is possible that when assessing the importance of arguments, predominantly impulsive individuals use less information and are less focused on relevant information, as Cintamulya (2019) showed in her qualitative study. The highest difference between the predominantly reflective and predominantly impulsive participants was in their estimation of the arguments of the third stage (social contracts), which impulsive individuals considered highly relevant. Among other things, the third stage is within the so-called conventional level of moral reasoning, and it is characterized by primarily mutual interpersonal expectations, stereotypic roles, and social conformity (when the expectations of people who are closest to us are considered morally correct). Overall, it could be argued that the predominantly impulsive individuals were (on average) "stopped" at the third stage of moral development, while the almost linear growth in the assessment of argument importance in the function of successive moral development stages was noticed in predominantly reflective individuals.

Finally, we expected that predominantly impulsive individuals would be faster in assessments of argument strength at all stages of moral development than predominantly reflective ones. The results did not support this hypothesis. In fact, they showed that predominantly impulsive individuals take longer to assess the importance of all arguments regardless of the stage of moral development. One possible explanation could be that the predominantly impulsive participants were not sure in their assessments in situations when there were no evident desirable/intuitive answers, which needs to be confirmed in future studies. Although the predominantly impulsive individuals usually prefer situations in which they can use System 1 (fast/intuitive/experienced answers), they tend to be slower than predominantly re-

flective individuals in situations where it is necessary to make a greater cognitive effort (System 2). With predominantly impulsive individuals, it is possible that there is a sort of cognitive "laziness" or "inactivity" when it comes to analytical and time-consuming reflectivity, as well as their inability to adjust to metacognitive monitoring and control of reasoning (Ackerman & Thompson, 2017). Metacognitive processes frequently do not have a direct insight into the accuracy of generated answers, and an indirect sign of metacognition is answer fluency (clear, correct, familiar, sure answers), that is, slower answers are connected to less certainty in their adequacy (Ackerman & Zalmanov, 2012).

As could be expected, the results indicated that the time needed for argument assessment grows as a function of argument complexity (represented by stages of moral development). The exception was the first stage that refers to preconventional level, in which a moral behaviour is directed towards obedience and avoiding punishment, and it is characterized primarily by egocentrism and the inability of taking into account other people's perspectives. It is possible that assessment time is longer at this stage due to the potential conflict between the "naturally egocentric" answer and the perception that the same answer is probably socially undesirable.

Finally, it is important to point out some of the observed correlations between different aspects of efficiency in the CRT and TMR. In general, the correlations showed that there was a statistically significant relationship between the CRT scores and the corresponding assessment time. It is evident that there was a certain tendency towards faster/slower decision making regardless of the response type (assessment/answer). Additionally, there was a statistically significant negative correlation between the average time needed for assessing the strength of moral arguments and the CRT scores. This is derived from the fact that predominantly impulsive individuals in general take longer to assess every moral argument and have lower CRT scores (they were classified according to this criteria) and vice versa. As could be expected, the correlation between the CRT scores and the IMR was also statistically significant, most likely as a result of the part of the common variance that could be attributed to general intellectual abilities. However, there was no significant correlation between the CRT scores, IMR, and corresponding solving times. As already stated, when it comes to the time needed to assess the strength of each moral argument, predominantly impulsive individuals take significantly longer and provide higher assessments in all stages (overestimation), while predominantly reflective individuals demonstrate opposite results (underestimation). Therefore, it could be assumed that the insignificant correlation was the effect of the interaction between the longer times in lower stages of moral development (lower IMR) and the longer times in higher stages (higher IMR). Moreover, one could have expected that the CRT scores would positively correlate with the total/average problem solving time (longer time presupposes analytical/reflective thinking and is expected for the correct answer). However, the absence of this correlation was probably the result of the fact that generally wrong answers are related to the longest solving time, and that predominantly impulsive individuals take longer to provide the correct answer as opposed to predominantly reflective ones.

CONCLUSIONS AND LIMITATIONS

The results of this study partially confirmed the hypotheses. Theoretically, it could be expected that predominately impulsive individuals are faster and less accurate on the CRT and the assessment of the moral arguments strength than predominantly reflective ones. As expected, there was a clear tendency to give accurate or intuitive CRT responses in most participants, which justifies the categorization of the individuals with respect to dominant cognitive styles. When assessing moral argument importance, predominately impulsive participants assessed arguments as more important at the first stages of moral development than predominantly reflective ones. However, the assumption that predominantly reflective individuals will assess arguments at higher stages of moral development as more important than predominantly impulsive ones has not been confirmed. Moreover, predominantly impulsive individuals evaluated moral arguments as more important at almost all stages of moral development.

Although there was a difference in the CRT response times of accurate and intuitive answers in the expected direction, contrary to expectations, predominantly impulsive individuals were significantly slower in making decisions and solving tasks in general, regardless of the task type. This result contradicts the previous assumptions that the dominant use of System 1 presupposes a faster response. Namely, it seems that predominantly impulsive individuals used System 1 more often, most probably because of some kind of cognitive laziness, and not due to experientially recognized responses.

When it comes to the parallel approach hypothesis of information processing, we can expect that predominantly reflective individuals tend to solve each task analytically and are faster in providing correct answers in comparison to impulsive ones, as opposed to when they answer intuitively (which is atypical for them). The opposite would apply to predominantly impulsive individuals. The results partially fit this theory, that is, only with predominantly impulsive individuals. Atypical answers in this group (result of reflective thinking) were related to longer solving time, especially for complex arguments, and longer time for the correct answers on the CRT. The results of predominantly impulsive individuals only partially fit the sequential model (longer CRT solving times for correct answers). With regard to the assessments of moral arguments that do not have typical intuitive responses, predominantly impulsive individuals were significantly slower than reflective ones, which challenges the basic premise of the sequential approach, whereby the correct answer takes more time because it includes the time needed to correct the initial (wrong) answer.

It appears that empirical studies of sequential and parallel dual information processing, whose tests of predominantly cognitive styles are based on the overall time needed to solve individual tasks, are not entirely adequate, that is, the same result can often be interpreted in the context of both approaches. Evidently, future studies should incorporate the assessment of different processes included in metareasoning. Furthermore, it seems that cognitive styles determine performance differently in tasks of various types (tasks with lure answers and those without them). Therefore, it can be assumed that cognitive styles

(reflectivity/impulsivity) indeed represent a relatively stable personality trait that is reflected in decision making in contextually different problem situations. However, the principal limitation of this study (and other similar studies) is that the division into dominant cognitive styles was based solely on the CRT scores, and this measure is highly related to different cognitive abilities. It would be more beneficial to find individuals of different cognitive styles, but with similar cognitive abilities. This would enable us to capture the specific variance of cognitive style not conditioned primarily by cognitive abilities. Such a research model would likely provide a better understanding of the various underlying processes of decision making, as well as contribute to further evaluation of dual process theories.

DATA AVAILABILITY

Data will be made available upon reasonable request to the corresponding author.

REFERENCES

Ackerman, R., & Zalmanov, H. (2012). The persistence of the fluency-confidence association in problem solving. *Psychonomic Bulletin & Review, 19*, 1187–1192. https://doi.org/10.3758/s13423-012-0305-z

Ackerman, R., & Thompson, V. (2017). Meta-reasoning: Shedding meta-cognitive light on reasoning research. In L. Ball & V. Thompson (Eds.), *The Routledge international handbook of thinking & reasoning* (pp. 1–15). Psychology Press.

Andersson, O., Holm, H. J., Tyran, J-R., & Wengström, E. (2016). Risk aversion relates to cognitive ability: Preferences or noise? *Journal of the European Economic Association, 14*, 1129–1154. https://doi.org/10.1111/jeea.12179

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition, 4*, 265–284. https://doi.org/10.1016/j.jarmac.2014.09.003

Campitelli, G., & Labollita, M. (2010). Correlations of cognitive reflection with judgments and choices. *Judgment and Decision Making, 5*, 182–191.

Cintamulya, I. (2019). Analysis of students' critical thinking skills with reflective and impulsive cognitive styles on conservation and environmental knowledge learning. *Asia-Pacific Forum on Science Learning and Teaching, 20*, 2.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making, 4*, 20–33.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron, 80*(2), 312–325. https://doi.org/ 10.1016/j.neuron.2013.09.007

Doyle, E., & O'Flaherty, J. (2013). The impact of education level and type on moral reasoning. *Irish Educational Studies, 32*, 377–393. https://doi.org/10.1080/03323315.2013.823273

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review, 13*, 378–395. https://doi.org/10.3758/bf03193858

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*, 25–42. https://doi.org/10.1257/089533005775196732

Gawronski, B., & Creighton, L. A. (2013). Dual process theories. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 282–312). Oxford University Press.

Gervis, W. M., & Norenzayan, A. (2012). Analytic thinking promotes religious disbelief. *Science, 336*, 493–496. https://doi.org/10.1126/science.1215647

Greene, J. D. (2009). Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology, 45*, 581–584. https://doi.org/10.1016/j.jesp.2009.01.003

Haigh, M. (2016). Has the standard cognitive reflection test become a victim of its own success? *Advances in Cognitive Psychology, 12*, 145–149. https://doi.org/10.5709/acp-0193-5

Hammond, K. R. (1996). *Human judgement and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice.* Oxford University Press.

Hren, D. (2008). *Utjecaj visokoškolskog obrazovanja na razvoj moralnog rasuđivanja osoba mlađe odrasle dobi* [The impact of higher education on the development of moral reasoning of people of younger adulthood; Doctoral thesis, University of Zagreb].

Jimenez, N., Rodriguez-Lara, I., Tyran, J-R., & Wengström, E. (2018). Thinking fast, thinking badly. *Economics Letters, 162*, 41–44. https://doi.org/10.1016/j.econlet.2017.10.018

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgement. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgement* (pp. 49–81). Cambridge University Press.

Kahneman, D. (2011). *Thinking fast and slow.* Farrar, Straus and Giroux.

Kohlberg, L. (1973). Stages and aging in moral development – Some speculations. *The Gerontologist, 13*, 497–502. https://doi.org/10.1093/geront/13.4.497

Paxton, J. M., Ungar, L., & Greene, J. D. (2011). Reflection and reasoning in moral judgment. *Cognitive Science, 36*, 163–177. https://doi.org/10.1111/j.1551-6709.2011.01210.x

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making, 2*, 453–469. https://doi.org/10.1002/bdm.1883

Proroković, A. (2016). Test moralnog rasuđivanja (TMR) [Test of moral reasoning]. *Zbirka Psihologijskih Skala i Upitnika, 8*, 63–74.

Proroković, A., Nikolić, M., & Šimić, N. (2017). Moral reasoning and its correlates in job applicants. *Arhiv za Higijenu Rada i Toksikologiju, 68*, 59–65. https://doi.org/10.1515/aiht-2017-68-2893

Rest, J. (1990). *Guide for the defining issues test.* Center for the Study of Ethical Development Minneapolis, University of Minnesota.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22. https://doi.org/10.1037/0033-2909.119.1.3

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral & Brain Sciences, 23*, 645–726. https://doi.org/10.1017/S0140525X00003435

Stanovich, K. E.,, & West, R. F. (2008). On the relative independence of thinking biases and cognitive ability. *Personality Processes and Individual Differences, 94*, 672–695. https://doi.org/doi:10.1037/0022-3514.94.4.672

Sternberg, R. J. & Grigorenko, E. L. (1997). Are cognitive styles still in style? *American Psychologist, 52*, 700–712. https://doi.org/10.1037/0003-066X.52.7.700

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology, 44*(1), 24–31.

Szaszi, B., Szollosi, A., Palfi, B., & Aczél B. (2017). The cognitive reflection test revisited: exploring the ways individuals solve the test. *Thinking and Reasoning, 23*, 207–234. https://doi.org/10.1080/13546783.2017.1292954

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking and Reasoning, 20*, 147–168. https://doi.org/10.1080/13546783.2013.844729

William W. J. (2004). Determinants of moral reasoning: academic factors, gender, richness-of-life experiences, and religious preferences. *Delta Pi Epsilon Journal, 46*, 105–123.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field dependent and field independent cognitive styles and their educational implications. *Review of Educational Research, 47*, 1–64. https://doi.org/10.2307/1169967

RECEIVED 01.08.2021 | ACCEPTED 20.04.2022