



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Can dual processing theory explain physics students' performance on the Force Concept Inventory?

**Citation for published version:**

Wood, AK, Galloway, RK & Hardy, J 2016, 'Can dual processing theory explain physics students' performance on the Force Concept Inventory?', *Physical review physics education research*, vol. 12, no. 2, 023101. <https://doi.org/10.1103/PhysRevPhysEducRes.12.023101>

**Digital Object Identifier (DOI):**

[10.1103/PhysRevPhysEducRes.12.023101](https://doi.org/10.1103/PhysRevPhysEducRes.12.023101)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Physical review physics education research

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Can dual processing theory explain physics students' performance on the Force Concept Inventory?

Anna K. Wood,<sup>\*</sup> Ross K. Galloway, and Judy Hardy

*Physics Education Research Group, School of Physics and Astronomy,  
University of Edinburgh, Edinburgh EH9 3FD, United Kingdom*

(Received 29 March 2016; published 28 July 2016)

According to dual processing theory there are two types, or modes, of thinking: system 1, which involves intuitive and nonreflective thinking, and system 2, which is more deliberate and requires conscious effort and thought. The Cognitive Reflection Test (CRT) is a widely used and robust three item instrument that measures the tendency to override system 1 thinking and to engage in reflective, system 2 thinking. Each item on the CRT has an intuitive (but wrong) answer that must be rejected in order to answer the item correctly. We therefore hypothesized that performance on the CRT may give useful insights into the cognitive processes involved in learning physics, where success involves rejecting the common, intuitive ideas about the world (often called misconceptions) and instead carefully applying physical concepts. This paper presents initial results from an ongoing study examining the relationship between students' CRT scores and their performance on the Force Concept Inventory (FCI), which tests students' understanding of Newtonian mechanics. We find that a higher CRT score predicts a higher FCI score for both precourse and postcourse tests. However, we also find that the FCI normalized gain is independent of CRT score. The implications of these results are discussed.

DOI: [10.1103/PhysRevPhysEducRes.12.023101](https://doi.org/10.1103/PhysRevPhysEducRes.12.023101)

### I. INTRODUCTION

Dual processing theory refers to a collection of models in cognitive psychology that describe two systems of thinking: system 1, which involves intuitive and nonreflective thinking, and system 2, which is more deliberate and requires conscious effort and thought (see, e.g., Refs. [1,2]). System 2 enables us to solve complex problems, but tends to be slow and demands high concentration and computational power. System 1, on the other hand, is quick and requires low computational power, but is less suited to many tasks involving higher-order thinking. Humans have a strong bias towards system 1 processing, and are often described as “cognitive misers” [3]. A three item instrument called the Cognitive Reflection Test (CRT) is commonly used to distinguish between peoples' tendency to engage in these two types of thinking. Devised in 2005 by Frederick [4], the CRT has been used particularly for assessing intuitive-analytic cognitive styles [5]. The three questions used in the standard CRT adapted from Ref. [4] are as follows:

- (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? (Intuitive answer 10p, correct answer 5p) [6].

- (2) If it takes 5 machines 5 min to make 5 widgets, how long would it take 100 machines to make 100 widgets? (Intuitive answer 100, correct answer 5).
- (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? (Intuitive answer 24, correct answer 47).

Each question on the CRT is designed to reliably cue an intuitive, but incorrect answer. For example, in question 1, the answer that initially comes to mind for most people is 10p. Although only a very simple calculation is required to check this answer, 10p is consistently the most common wrong answer (even for students at MIT [2]), indicating that people have responded with their initial, intuitive (system 1) response. To get the question correct requires the answerer to consciously reject this initial answer and to correctly calculate the new answer. The CRT therefore measures the tendency to override an initial, intuitive answer (system 1) and then to engage in more reflective thought (system 2) to determine the correct answer.

Performance on the CRT has been shown to correlate with a variety of measures involved with decision making, as well as measures of intelligence, cognitive reflection [4], and logical reasoning ability [7]. Although the CRT correlates with both intelligence measures and numeracy, there is general agreement that the CRT is not simply a measure of numerical ability [3,8,9] and that the CRT explains additional variance in reasoning and decision-making tasks when numeracy and intelligence are taken

<sup>\*</sup> [annakwood@physics.org](mailto:annakwood@physics.org)

into account [3]. The CRT is also a powerful predictor of someone's likelihood of engaging in rational decisions and their ability to make unbiased judgments. However, performance on the CRT can also be affected by context: in one study, students made fewer errors when the test was presented in a lighter, harder to read font compared to the normal version. Alter *et al.* hypothesized that the difficult to read version invoked system 2 thinking [10].

### A. Rationale

Our work was motivated by the possible similarity between the cognitive challenges of the CRT and those faced by students learning physics. Students often demonstrate naive or intuitive ideas when answering physics questions which are based on their experience of the world, but which, from a scientific perspective, are incorrect. These incorrect answers are often referred to as “misconceptions.”

However, there is debate about the role that misconceptions play in learning physics. This is important because how misconceptions are viewed affects approaches to teaching. The two most common theories that discuss cognitive processing in physics are the “knowledge as theory” [11] perspective, which views student knowledge as consistent and resistant to change, and the “knowledge in pieces” perspective [12], in which students' cognition is seen as highly context dependent and therefore changeable. In the knowledge as theory perspective, misconceptions are viewed as ideas that are likely to hinder learning. Teaching is therefore focused on challenging these misconceptions through conceptual change.

In contrast, researchers who view knowledge as being “in pieces” [12,13] see students' ideas as fragmentary, transient, and context dependent. From this perspective, the resources available to students have a much smaller grain size, for instance, knowledge elements [14] or p-prims [15,16], basic units of intuitive resources gained from experience of the world. Learning is seen as a process of developing “patterns of association” [14] which can both help and hinder the individual in solving physics problems. Dual processing theory may provide an alternative approach to understanding the role of misconceptions in learning physics.

### B. Approach

We hypothesized that the cognitive processes involved in answering the CRT are similar to those involved in conceptual understanding of physics, particularly in areas where misconceptions are common, such as Newtonian mechanics.

In order to test this we analyzed how students perform both on the CRT and on a commonly used test of conceptual understanding—the Force Concept Inventory (FCI). The FCI is a 30 item multiple-choice test designed by Hestenes *et al.* [17] and updated in 1995 [18] that

gives a measure of students' conceptual understanding of Newtonian mechanics. The test builds on the work of Halloun and Hestenes [19], which found that common sense beliefs about motion had a large effect on performance in physics. To design the FCI students were interviewed and the incorrect ideas that they had about concepts in Newtonian mechanics were used as the basis for the test distracters. As misconceptions are stable and resistant to change under traditional instruction [11,20], answering FCI questions correctly could be said to involve two steps: first, students must discard options that constitute an intuitive (but incorrect) response, and second, they must use appropriate conceptual reasoning about a physical situation to find the correct answer. We do not know for sure what students are thinking, but as misconceptions are so prevalent, at least for novices, this seems to be plausible and therefore worth investigating further. We therefore hypothesize that students who score higher on the CRT will also score higher on the FCI.

## II. CONTEXT AND METHODOLOGY

The research described here involved students studying an introductory calculus-based Newtonian mechanics course at the University of Edinburgh, UK. Typical class sizes are 200–300 students with a gender ratio of around 80:20 males to females. Students taking this course were both physics majors and nonmajors, although all students must satisfy the entry requirements for the physics degree program. The course has a “flipped” structure in which students are expected to do pre-readings and an online quiz before the lecture. The lecture time then consists of interactive engagement pedagogies, primarily Peer Instruction (see Wood *et al.* [21] for more details). The average normalized FCI gain for the course is  $0.5 \pm 0.1$ .

The FCI was administered online via the web-based course management system in week 1 and again in week 10 of the first semester in order to measure the learning gain of the students. Students completed the test in their own time but were restricted to 90 min. Our results on the FCI are comparable to those from other UK institutions who use a paper-based test under test conditions [22]. The CRT questions were delivered via the Top Hat online response system. This is a web-based system, and students respond via smart phones, tablets, laptops, or any other web-enabled device. The system supports multiple-choice, free text, numerical, ranking, and other question types. For the CRT questions we used the numerical response question type.

The CRT questions were delivered during the second lecture of the course, during the first week of instruction. Since the CRT questions are seemingly trivial, and obviously unrelated to the content of the course, they might seem incongruous to the students. Consequently, the students might seek out the “trick” to the questions, thus possibly overriding system 1 thinking. To attempt to counteract this, a mild subterfuge was used: the Top Hat

TABLE I. Percentage of students with each CRT score. Numbers in brackets give the number of students.

Percentage of students with each CRT score (number of students)	0	1	2	3	Mean
University of Edinburgh (148)	5 (8)	16 (23)	22 (32)	57 (85)	2.31
Meta-analysis [23]	38	23	21	18	Not known

system was being used for the first time in this academic year, and its predecessor system supported only multiple-choice questions. The students were informed that they were going to be given “three quick questions to try out the numerical response capabilities of the new system,” thus obfuscating the intent of the questions. The three CRT questions were delivered in quick succession over a 3 min period, and the correct answers were not revealed until student responses for all questions had been collected, thus not “tipping off” the students that the intuitive answers were not correct.

In total 148 students answered every question on the CRT and also completed both the FCI precourse and FCI postcourse tests. One limitation of this study is that some students only answered one or two questions on the CRT rather than all three. Data from these students were not included in the analysis, as failure to answer a question may have been due to technical problems. However, if students left an answer blank because they did not know the answer to the question, there may be a bias in our results towards those who score more highly on the CRT.

### III. RESULTS

Students scored higher on the CRT than is typical for the general population [23], with 74% answering Q1 correctly, 71% answering Q2 correctly, and 86% answering Q3 correctly. This may be because the students studied in this research are likely to have much better than average mathematical skills and score higher on intelligence measures, both of which are known to aid performance on the CRT [24]. However, the general trend seen between questions is similar to that reported in the literature, with students scoring higher on question 3 compared to questions 1 and 2.

We also found, as expected, that the intuitive heuristic answer accounted for the majority (72%) of all wrong answers. For Q1, 90% of students who answered incorrectly did so with the heuristic answer. For Q2 and Q3, this was 49% and 45%, respectively, but the heuristic answer was still the most common wrong answer.

CRT scores were calculated for each student by counting the number of items answered correctly, giving a CRT score of between zero (all questions answered incorrectly) and three (all questions answered correctly). FCI scores for pre- and postcourse tests (out of a maximum of 30) were calculated by counting the number of questions that the student answered correctly. Table I shows the percentage (and number) of students with each CRT score for the

148 students in this study. For comparison we have also given the results from a meta-analysis of 118 studies across the general population. Our results closely resemble those from MIT students studied by Frederick [4], who compared CRT scores for a range of different populations, including students at a number of different universities. The MIT students were the highest scoring of the populations that he investigated.

Figure 1 shows the median FCI score for students as a function of their CRT score for both precourse and postcourse FCI tests. A clear positive relationship is seen between precourse FCI score and CRT score, with students who score zero on the CRT scoring 12 points on average less on the FCI compared to the students who scored 3 on the CRT. The change in FCI scores is particularly pronounced between those who scored zero and one on the CRT and between those who scored one and two on the CRT. A smaller gap is seen between those scoring two and three on the CRT. The overall pattern is similar, but less pronounced, for the postcourse FCI: the gap in FCI score between students scoring zero and those scoring three on the CRT is 9.5. A Pearson correlation coefficient was calculated to be 0.38 (precourse test) and 0.32 (postcourse test). However, this assumes a linear relationship between the variables. In both pre- and postcourse tests the increase in FCI score for students scoring a CRT of three compared to students scoring two is smaller than would be expected for a linear relationship. It may be that a flooring effect is evident here; the CRT is unable to discriminate effectively between students at the top end of the CRT. This criticism is

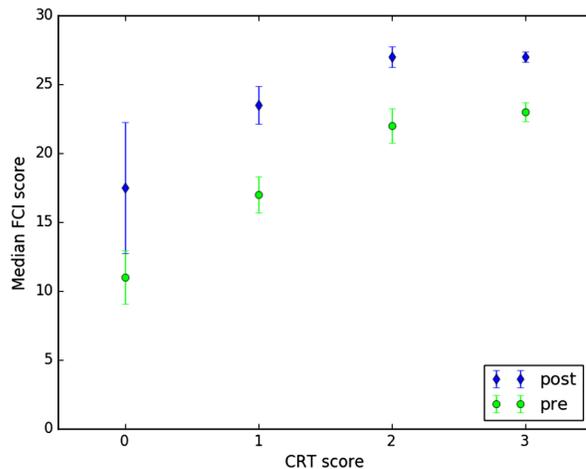


FIG. 1. Median FCI score versus CRT score for precourse and postcourse FCI tests. Uncertainties are standard error on the median calculated as described in Ref. [25].

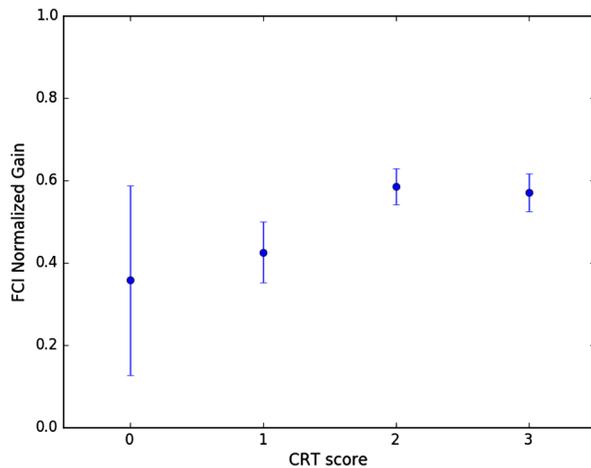


FIG. 2. Median FCI normalized gain versus CRT score. Uncertainties are standard error on the median.

discussed by Toplak *et al.* [9], who have proposed a longer CRT consisting of seven items, which those authors hope will overcome this weakness. A test that is better able to discriminate between students who do well on the CRT may result in a stronger correlation between FCI and CRT scores.

Figure 2 shows the average normalized FCI gain as a function of CRT score. The normalized gain for each student is calculated as

$$\langle G \rangle = \frac{\langle \text{FCI}_{\text{post}} \rangle - \langle \text{FCI}_{\text{pre}} \rangle}{30 - \langle \text{FCI}_{\text{pre}} \rangle}, \quad (1)$$

where 30 corresponds to the maximum possible FCI score and the angle brackets on the right-hand side indicate medians (rather than means). The figure shows that the average normalized gain on the FCI is independent of CRT score (Pearson’s correlation coefficient was calculated to be 0.07). This is a surprising result, implying that improvement in conceptual understanding of Newtonian mechanics is independent of cognitive reflection, as measured by the CRT. In fact, the absolute change in FCI score is slightly larger for the group that scored zero on the CRT compared to the group that scored three (see Fig. 1). This finding is at odds with research reported by Shtulman and McCallum [26], who measured science understanding across a number of different fields (including astronomy, evolution, geology, mechanics, perception, and thermodynamics) and found that cognitive reflection was a significant predictor of science understanding. They concluded that “These results suggest that cognitive reflection may be a prerequisite for changing certain cognitive structures, namely, concepts and theories” [26].

#### IV. DISCUSSION AND CONCLUSIONS

We presented an analysis of the relationship between students’ score on the Cognitive Reflection Test and their

performance on the Force Concept Inventory. We find a clear relationship between the two tests, with students who score higher on the CRT gaining higher scores on the FCI, for both the precourse and postcourse administrations of the test. We also found that the normalized gain on the FCI was independent of CRT score.

These findings indicate that students who are more likely to override the system 1 intuitive response and to engage in the more demanding cognitive reflection needed to answer the CRT question correctly are also more likely to score highly on the FCI, implying that similar cognitive processes account for at least some of the cognitive abilities needed for each test. Although clearly more research is needed, this result has important implications for the way that the nature of student difficulties with physics are thought about. A dual processing theory approach has not received much attention in the physics education literature (however, see Refs. [27,28]), but our findings suggest that this perspective offers a promising approach for understanding physics students’ difficulties.

A dual processing perspective combines aspects of both the “knowledge as theory” and “knowledge in pieces” perspectives, described earlier. Intuitive ideas developed from experiences of the world are likely to be deeply ingrained and could therefore become heuristics that are adopted as system 1 responses. This may explain why some misconceptions seem to be common and universal. Students who tend to be “cognitive misers” will be more likely to answer physics questions with their intuitive ideas. However, students who are able to override these intuitive responses and to engage in system 2 thinking are much more likely to then activate appropriate physics resources to enable them to answer the question correctly. Activation of resources can therefore be seen as something that is dependent on both the context of the problem and the individual characteristics of the students. This view is similar to that described by the “resources model” [29,30] in which students’ incorrect responses are thought to result from students not activating the appropriate resources, rather than because they do not have the physics knowledge necessary to correctly answer the problem.

Our finding that the normalized FCI gain is independent of CRT score raises further questions about the role of misconceptions in learning physics. We find that students who are likely to rely on system 1 thinking (and therefore to choose the intuitive option), as measured by the CRT, see a similar improvement in conceptual understanding as students who override their intuitive ideas and engage system 2 thinking. This result implies that even though misconceptions in the form of common intuitive ideas seem to exist, they do not hinder the development of more scientific views of the world.

The implications of a dual processing theory perspective for physics instruction are twofold. Firstly, the idea

that misconceptions are deeply ingrained intuitive ideas that are activated through system 1 thinking, without conscious thought, implies that students should be helped to develop the cognitive reflection skills necessary to override system 1 and to engage in system 2 thinking. Secondly, rather than challenging misconceptions, teaching should focus on building connections between scientific ideas, so that appropriate resources can be more readily activated.

Although more work is clearly needed, we believe that a dual processing perspective has the potential to further our understanding of learning in physics. These results are the initial findings of a more detailed study which will be presented in due course. We encourage other researchers to replicate, extend, and build on these findings in different contexts and using different measures of physics understanding, so that the role that dual processing theory may have can be further explored.

- 
- [1] J. St. B. T. Evans and K. E. Stanovich, Dual-process theories of higher cognition advancing the debate, *Perspect. Psychol. Sci.* **8**, 223 (2013).
- [2] D. Kahneman, *Thinking, Fast and Slow* (MacMillan, London, 2011).
- [3] M. E. Toplak, R. F. West, and K. E. Stanovich, The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks, *Mem. Cogn.* **39**, 1275 (2011).
- [4] S. Frederick, Cognitive reflection and decision making, *J. Econ. Perspect.* **19**, 4 (2005).
- [5] G. Pennycook, J. A. Cheyne, D. J. Koehler, and J. A. Fugelsang, Is the cognitive reflection test a measure of both reflection and intuition?, *Behav. Res. Meth. Instrum. Comput.* **48**, 341 (2016).
- [6] The original question uses dollars. We asked the question using local currency as this would be more familiar to the students (and would avoid adding unwanted additional cognitive load).
- [7] J. M. Liberali, V. F. Reyna, S. Furlan, L. M. Stein, and S. T. Pardo, Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment, *J. Behav. Decis. Making* **25**, 361 (2012).
- [8] G. Campitelli and P. Gerrans, Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach, *Mem. Cogn.* **42**, 434 (2014).
- [9] M. E. Toplak, R. F. West, and K. E. Stanovich, Assessing miserly information processing: An expansion of the Cognitive Reflection Test, *Think. Reas.* **20**, 147 (2014).
- [10] A. L. Alter, D. M. Oppenheimer, N. Epley, and R. N. Eyal, Overcoming intuition: Metacognitive difficulty activates analytic reasoning., *J. Exp. Psychol. Gen.* **136**, 569 (2007).
- [11] S. Vosniadou, C. Ioannides, A. Dimitrakopoulou, and E. Papademetriou, Designing learning environments to promote conceptual change in science, *Learn. Instr.* **11**, 381 (2001).
- [12] J. P. Smith III, A. A. diSessa, and J. Roschelle, Misconceptions reconceived: A constructivist analysis of knowledge in transition, *J. Learn. Sci.* **3**, 115 (1994).
- [13] A. A. DiSessa, N. M. Gillespie, and J. B. Esterly, Coherence versus fragmentation in the development of the concept of force, *Cogn. Sci.* **28**, 843 (2004).
- [14] J. Tuminaro and E. F. Redish, Elements of a cognitive model of physics problem solving: Epistemic games, *Phys. Rev. ST Phys. Educ. Res.* **3**, 020101 (2007).
- [15] A. A. DiSessa, *Changing Minds: Computers, Learning, and Literacy* (MIT Press, Cambridge, MA, 2001).
- [16] A. A. DiSessa, Toward an epistemology of physics, *Cognit. Instr.* **10**, 105 (1993).
- [17] D. Hestenes, M. Wells, and G. Swackhamer, Force Concept Inventory, *Phys. Teach.* **30**, 141 (1992).
- [18] The 1995 version of FCI is available at <http://modeling.asu.edu/R&E/Research.html>.
- [19] I. A. Halloun and D. Hestenes, The initial knowledge state of college physics students, *Am. J. Phys.* **53**, 1043 (1985).
- [20] S. Vosniadou, Capturing and modeling the process of conceptual change, *Learn. Instr.* **4**, 45 (1994).
- [21] A. K. Wood, R. Galloway, R. Donnelly, and J. Hardy, Characterizing interactive engagement activities in a flipped introductory physics class, *Phys. Rev. Phys. Educ. Res.* **12**, 010140 (2016).
- [22] R. Donnelly, Ph.D. thesis, University of Edinburgh, 2014.
- [23] P. Branas-Garza, P. Kujal, and B. Lenkei, “15–25 Cognitive Reflection Test: Whom, how, when” (to be published).
- [24] A. Sinayev and E. Peters, Cognitive reflection vs. calculation in decision making, *Front. Psychol.* **6**, 532 (2015).
- [25] R. McGill, J. W. Tukey, and W. A. Larsen, Variations of box plots, *Am. Stat.* **32**, 12 (1978).
- [26] A. Shtulman and K. McCallum, Cognitive reflection predicts science understanding, in *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (Quebec City, 2014), pp. 2937–2942, <https://mindmodeling.org/cogsci2014/>.
- [27] D. Sands, Evidence for dual processing theory, in *Proceedings of the International Conference on Physics Education* (Prague, 2013), [http://www.icpe2013.org/uploads/ICPE-EPEC\\_2013\\_ConferenceProceedings.pdf](http://www.icpe2013.org/uploads/ICPE-EPEC_2013_ConferenceProceedings.pdf).
- [28] M. Kryjevskaja, M. R. Stetzer, and N. Grosz, Answer first: Applying the heuristic-analytic theory of reasoning to examine student intuitive thinking in the context of physics, *Phys. Rev. ST Phys. Educ. Res.* **10**, 020109 (2014).
- [29] D. Hammer and A. Elby, Epistemological resources, in *Proceedings of the Fourth International Conference of the Learning Sciences*, University of Michigan (Erlbaum, Mahwah, NJ, 2000), pp. 4–5.
- [30] E. F. Redish, A Theoretical Framework for Physics Education Research: Modeling Student Thinking, in *Proceedings of the International School of Physics “Enrico Fermi”* (IOS Press, Varenna, 2003), Vol. 156, pp. 1–63.