Response strategies of instructed malingerers during forced choice testing

DOI:
10.26481/dis.20190625ro

Summary

Malingered cognitive impairment, such as loss of hearing, or malingered loss of memory for specific events, for example a crime, can be detected with the Forced Choice Test (FCT). This test features a series of binary multiple choice trials and forces the examinee to make a choice on each trial. Genuinely impaired performance is defined as guessing, producing test scores within levels of chance. In contrast, malingered performance is defined as the intentional avoidance of correct answers, producing test scores worse than chance performance. The aim of this thesis was twofold. First, to explore the FCT's optimal detection accuracy, and second, to establish a model that describes the various counterstrategies malingerers use to defeat this test.

Chapter 2 deals with the lack of theoretical conceptualization of malingerers' behaviour during a FCT. To address this, a model was devised based on Cognitive Hierarchy Theory (CHT; Carmerer, Ho, & Chong, 2004) and strategies derived from self-reports reported in previous studies. According to CHT, examinees take the strategies of other players into account when developing their own strategy and the sophistication of this process is limited by the available cognitive resources of the examinee. Another feature is that strategies are hierarchical, indicated with numeric levels, with higher levels being superior to lower. Here we define three strategy levels for malingerers in the FCT ranging from 0 to 2. A level 0 strategy predicts the examinee will endorse the correct information, resulting in overperformance (more items correct than expected by chance). A level 1 strategy predicts that the examinee will avoid selecting correct answers, resulting in underperformance. A level 2 strategy predicts that an examinee will provide a balanced mixture of correct and incorrect answers. We subjected liars and truth tellers to a traditional FCT or to a FCT that included a fake polygraph examination to misdirect examinees' beliefs about the detection mechanism of the FCT. Test performance and response strategy levels

were measured. The main findings were that (i) substantial proportions of liars used level 2 strategies, which suggests correct understanding of the FCT's detection mechanism; (ii) different strategy levels featured different detection accuracies; and (iii) examinees test behaviour could be influenced by misdirecting them from the test detection mechanism. Together, these findings provide the first support for our proposed model. Based on this knowledge, future manipulations of the FCT paradigm can be developed to increase the detection accuracy of the FCT. This experiment has been published in Orthey, Vrij, Leal, and Blank (2017).

Chapter 3 draws on the model introduced in the previous chapter. As stated, the FCT is apt at detecting level 1 strategies, but not level 2 strategies. The underlying theory assumes that higher order strategies require more cognitive resources. Therefore, if cognitive resources are limited examinees may be less likely to select a strategy level with poor detection accuracy (level 2) and instead promotes selection of lower order strategies (level 0 and 1), which are more easily detected. To limit the available cognitive resources we introduced time pressure to the FCT paradigm, by forcing examinees with and without concealed information to select an answer alternative within two seconds. We compared this paradigm with the traditional FCT in terms of strategy selection and detection accuracy. The main findings were that (i) selection of strategy levels was not affected by time pressure; (ii) in both the time pressure and traditional FCT, the number of correct items selected discriminated examinees with and without concealed knowledge better than chance; and (iii) examinees with concealed knowledge, who reported a level 2 strategy, had selected fewer correct items under time pressure than in the traditional FCT paradigm, leading to a considerable higher proportion of cases at underperformance level in the time pressure FCT. These results suggest that time pressure is not suited to affect the strategy selection process, but instead affects execution of the strategy. That is, examinees who report level 2 strategies and intend to randomize between correct and incorrect answers, are expected to avoid detection by the underperformance criterion, as demonstrated in the standard condition limiting the overall detection accuracy of the FCT. The time pressure condition demonstrated that examinees, using the same strategy, were less successful in avoiding detection by the underperformance criterion. Consequently, cognitive load, in terms of time pressure, could be used to limit the effectiveness of a common and effective counterstrategy in the FCT.

Chapter 4 examinees a new criterion to detect level 2 strategies and its value in dealing with cases of coaching. Coaching describes the act of an examinee seeking information on a forensic test prior to administration. This is a concern for the FCT, because once an examinee is aware of the underperformance criterion the examinee is likely to use a level 2 strategy and randomize between correct and incorrect answers. As the detection rate for level 2 strategies is poor, coaching is a threat to the validity of the FCT. The 'runs-test' has been suggested to measure the alternations between correct and incorrect answers. It is based on the idea that examinees who are unaware of the correct answer, have a likelihood of 50% to alternate between trials, thus like the traditional criterion, they are expected to produce a number of alternations within chance levels. In contrast, examinees who are aware of the underperformance criterion, are expected to alternate more frequently between correct and incorrect answers to ensure that the total number of correct items falls within chance levels. Hence, the 'runs-test' detects examinees using a level 2 strategy through elevated alternation rates between correct and incorrect answers. So far, empirical support suggests it is of limited value only (Jelicic et al., 2004; Verschuere et al., 2008). To increase the validity of the runs-test, we attempted to force examinees to choose between a randomizing pattern that 'looks' random and a randomizing pattern that produces a test score within chance levels. To do so we alternate the position of the correct answer alternative (left or right) between trials as well. Consequently, alternating between correct and incorrect answers means only answers on the same side are selected, while alternating between answers presented on the left and right would lead to more extreme scores. We expected that coached examinees would prefer the former, while examinees who are genuinely guessing, would prefer the latter pattern. Detection accuracy of the 'runs-test' would be increased, because both types of randomizing behaviour are anti-correlated, increasing the difference between both groups. The main findings were that (i) coaching was associated with level 2 strategies and underperformance was apt at detecting level 1, but not level 2 strategies; (ii) the runs-test was able to detect coached examinees with concealed knowledge; (iii) the underperformance criterion and runs-test criterion can be utilized as a 2-step classification procedure, with underperformance being sensitive to level 1 strategies and the runs-test to level 2 strategies. Together these findings support the underlying strategy levels and their associated test scores as well as the need to detect level 2 strategies in order to increase detection accuracy of the FCT. This article was published as Orthey, Vrij, Meijer, Leal, and Blank (2018).

Chapter 5 is focused on detecting level 2 strategies in case of malingered red/green blindness. Specifically, the validity of the 'runs-test' and a new criterion was evaluated. The idea behind the new criterion was to elicit a response bias within the generation process of a test score that falls within chance performance. Specifically, we aimed at introducing a manipulation, independent of the actual task (here to discriminate red and green), that elicits a systematic preference. We varied the perceived difficulty of the trials by manipulating the see-throughness of the stimuli, so malingerers could be more likely to select correct answers on trials that are clearly visible and be more likely to select incorrect answers when they are not. If examinees attune their selection preference to the perceived difficulty of the trial, this systematic pattern would deviate from genuine guessing behaviour and could serve as a new criterion. Therefore, we instructed examinees to simulate red/green blindness and subjected them to a FCT of 100 trials embedded into a test battery. The FCT in the standard condition featured a bright red and a bright green rectangle on each trial. Additionally, in the opacity condition, the see-throughness of the rectangles was varied over the trials, creating the illusion that on some trials the correct answer would be more difficult to identify than on others. We re-examined the validity of the 'runs-test', because malingered sensory deficits, as opposed to malingered loss of memory, allows for the construction of FCTs with larger test sizes and the 'runs-test' should be more effective with longer tests. The main findings were that: (i) the runs-test did detect malingerers better than chance in the standard, but not in the opacity condition; (ii) malingerers produced more statistically significant response biases than expected by chance; and (iii) the probability of the individual response biases detected malingerers better than chance. These findings suggest that the 'runs-test' or a systematic association between perceived trial difficulty and endorsement of correct answers could be used as criterion to detect level 2 strategies in malingered sensory deficits.

Finally, Chapter 6 will feature an evaluation of the three strategy levels and their correspondence to malingerers' test behaviour as well as a reassessment of the FCT's detection accuracy. Pathways to increase detection are discussed and two fields of application, malingered sensory dysfunction and malingered loss of memory, are differentiated in terms of test construction and choice of criteria. The chapter closes with a reflection on experimental limitations and future challenges.