

Testing Hypotheses about Mechanical Devices

Aidan Feeney

Department of Psychology
University of Durham
Science Laboratories
South Road
Durham DH1 3LE
United Kingdom
aidan.feeney@durham.ac.uk

Simon J. Handley

Centre for Thinking and Language
Department of Psychology
University of Plymouth
Drake's Circus
Plymouth PL4 8AA
United Kingdom
shandley@plymouth.ac.uk

Abstract

We use the literature on mechanical reasoning to derive predictions about how people will test a mechanical rule. In the presence of a single rule we predict significantly more selections of tests in which the hypothesized cause is manipulated than in the presence of two rules: the original and one casting doubt on the sufficiency of the hypothesized cause for the effect. We describe an experiment using Wason's selection task that confirms our predictions and go on to discuss the implications of our results for recent work on causal cognition.

Introduction

Much of our everyday reasoning concerns the operation of mechanical devices. Many of our interactions with such devices, from figuring out how to program the VCR to mastering the newest piece of software, require us to draw inferences, make predictions and test hypotheses. Some of this reasoning is very important and, accordingly, mistakes can be costly. For example, the Chernobyl explosion (see Medvedev, 1991) has been used by Johnson-Laird (1993) to illustrate how biased interpretation of evidence relevant to an hypothesis about a device can lead to disaster. Fortunately, not all faulty mechanical reasoning leads to such calamitous consequences and sometimes we even have the luxury of experimenting with a device in order to see how it works. For example, the paradigmatic case in the literature on discovery learning (Khlar & Dunbar, 1988) is how people acquire the ability to operate a device by self-guided trial and error learning.

In this paper we will be concerned with the principles that determine the tests that people choose to carry out in order to examine an hypothesis about a mechanical device. Consider, for example, a conditional hypothesis concerning the cooling system in a factory:

(1) If the valve is open then water flows through the pipe.

Our concern is how people will go about testing such a rule rather than how they should test the rule. There are four cases which seem intuitively relevant here: the valve being open; the valve being closed; water flowing through the pipe; and water not flowing through the

pipe. According to a normative theory based on the hypothetico-deductive method (Popper, 1959), in order to test the rule participants should choose to discover information about what happens both when the valve is open and when there is no water running through the pipe. Only these tests can provide evidence that falsifies the rule. Much research suggests that people do not consider these to be the best tests with which to examine a hypothesis and we do not think the case of mechanical reasoning is likely to be any different in this respect.

We suspect, however, that causal hypotheses concerning a mechanical device - such as the one above - will elicit a very specific pattern of testing behaviour amongst participants. The detail of our suspicions and a more comprehensive rationale will be presented in the next section.

Mechanical Reasoning

There is a variety of work on mechanical reasoning in the literature all of which suggests that people use representations that are in some way analog to the device being reasoned about. For example, Hegarty (1992) has proposed an account of how people think about devices claiming that mechanical reasoning processes are isomorphic to the structure of the device that is the subject of those processes. Hegarty's theory was intended to explain results obtained using a mechanical reasoning paradigm where participants are shown static images of pulley systems and are asked to predict the effect of some manipulation of the device. For example, participants might be asked how one or other of the pulleys in the image will move when a rope is pulled. Hegarty's data, both from eye-movements and gross reaction times, suggest that in making inferences about a mechanical device, people animate contiguous elements of the system piecemeal, by inferring a causal chain of events from the input of the system to its output.

Other work supports this conclusion. Schwartz & Black (1996), in showing that people can induce rules to describe a system based on mental depictions of that system, provide evidence suggesting that the initial depictions are constructed in a causal order. That is, in the absence of a rule, people appear to mentally

simulate the effects of a manipulation to the system in the direction of cause and effect.

People's piecemeal animation (Hegarty, 1992) and depiction (Schwartz & Black, 1996) of mechanical devices, suggest that their representations of mechanical devices respect the causal structure of the device. Furthermore, when people perform thought experiments involving those representations, they perform them so as to observe the effects on the system of the manipulation of a causal agent.

Testing Mechanical Hypotheses

We assume that in testing a causal hypothesis concerning a mechanical device people will form a representation of the hypothesized rule that respects its causal direction. So, given hypothesis 1:

If the valve is open then water will flow through the pipe

we expect participants to encode in their representation the information that, under this hypothesis, the valve is of causal significance with respect to water flowing through the pipe.

Given previous work on how people animate and perform mental experiments on models of devices, we expect participants, when imagining the possible consequences of performing a test on the system described in the hypothesis, to represent the consequences of the antecedent condition being in a certain state. That is, we would expect participants to be more interested in tests that respect the causal structure of the hypothesis than in tests that require backwards reasoning from changes in the effect to changes in its putative cause. Specifically, we expect participants to be more interested in cases where the cause is present or absent than in cases where the effect is present or absent.

The prediction that participants will be interested in cases where the effect is absent is a risky one, as participants are not normally interested in the false antecedent case when testing a conditional rule. Indeed, Oaksford and Chater's (1994) probabilistic account of how people test conditional rules claims that the false antecedent case is never informative. The situation for causal conditional rules is very different, however, where interest in the false antecedent case might be interpreted as being due to the use of a counterfactual strategy in testing the causal status of the antecedent. Mackie (1974) claims that we infer causality not only from repeated observations of contiguous events but also from a consideration of what might be observed in the absence of the putative cause. If the effect is also absent under these circumstances then we infer a causal relationship between the two. Harris, German & Mills (1996) have demonstrated such a strategy in the causal reasoning of children aged between 3 and 5 years.

There are, however, conditions under which we would not expect participants to be primarily interested in tests that respect the causal structure of the device. For example, if the hypothesis is presented at the same time as a second rule:

- (1) *If the valve is open then water will flow through the pipe*
- (2) *If the pipe is free from blockages then water will flow through the pipe*

where this second rule specifies an additional antecedent for the consequent, then we would expect participants to select fewer tests where the hypothesized cause is manipulated. This is because the additional antecedent introduces a potential disabling condition (Cummins et al 1991) for the hypothesized cause. This would mean that a failure to find the effect in the presence or absence of the cause might be attributable either to the hypothesized cause being insufficient to produce the effect or to the absence of the enabling condition. In the example above, the valve being open and the pipe being free from blockages might be conjointly necessary for water to flow through the pipe. If this is the case then examining the results of tests involving manipulation of the valve is unlikely to be revealing of the truth or falsity of the rule in the absence of information about the presence or absence of the enabling condition.

A Mechanical Selection Task

To test our intuitions about how causal rules about a mechanical device are represented and hence tested, we constructed a mechanical version of Wason's selection task (Wason, 1968). In our version of the task participants received a scenario (see below) which supplied a context for a conditional rule describing a causal relationship between the state of a component of the device and some output. Underneath were printed four cards representing the true antecedent, the false antecedent, the true consequent and the false consequent states of affairs. To test our hypotheses concerning the conditions under which participants would be primarily interested in tests of the hypothesis that manipulated the putative cause, we constructed a second version of the task. This second version was achieved by adding a second rule to the problem specifying an additional antecedent for the same outcome (see 1 and 2 above).

Our manipulation of number of rules is directly analogous to the presentation of additional antecedents in the conditional arguments task (Byrne, 1989). Participants who receive conditional reasoning problems that specify an additional antecedent are significantly less likely to draw the valid Modus Ponens and Modus Tollens conclusions than are participants who do not receive information about an additional

antecedent. This may be interpreted as the result of the additional antecedent causing participants to doubt the sufficiency of the first antecedent for the rule.

Byrne's work was an extension of experiments reported by Romain, Connell & Braine (1983) who showed that presenting participants with a second conditional rule that specified an alternative antecedent for the consequent suppressed the rate at which the invalid Denial of the Antecedent and Affirmation of the Consequent inferences were drawn. Recently, Feeney and Handley (2000; Handley, Feeney & Harper, 2000) have described the results of selection task experiments where participants received a second rule specifying an alternative antecedent for the consequent in the rule to be tested. Across a series of six experiments large and reliable rates of suppression of Q card selections were found. A meta-analysis of five of the six experiments showed that the rate of not-P card selection was also significantly lower in the presence of an alternative antecedent.

The Q and not-P cards on the selection task are logically equivalent to the DA and AC inferences. Suppressing the rate at which they are selected is analogous to suppressing the rate at which the invalid inferences are made. One prediction about our mechanical selection task, therefore, is that the presence of a second rule specifying an additional antecedent should produce suppression on those cards which are logically equivalent to the MP and MT inferences i.e. the P and not-Q cards. However, we predict that this will not be the case.

Instead we expect rates of antecedent card selection (both true and false antecedent cards) to decrease in the presence of an additional antecedent. This prediction is based on our assumption that in reasoning about an hypothesis concerning a mechanical device people will incorporate into their representation the hypothesized causal status of the antecedent with respect to the consequent. Based on previous work (Hegarty, 1992; Schwartz & Black, 1996) we hypothesize that they will prefer to examine the results for the effect of manipulating the antecedent rather than vice versa. Such a testing strategy should be significantly reduced in the presence of a second rule containing an additional (and perhaps conjointly necessary) antecedent.

Method

Participants: 90 female and 21 male students at the University of Durham participated in this experiment. Participants' mean age was 20.5 years and age ranged from 17 to 42 years.

Design, Materials and Procedure: There were two groups of participants in this experiment. The first received a one-rule selection task containing a scenario and just one rule. The other group of participants received a two-rule selection task comprising of the

same scenario and rule to be tested as well as a second rule specifying an additional antecedent for the consequent in the first rule. For all participants the scenario and the rule to be tested were as follows:

A friend of yours, who works in a factory, takes you on a tour of her place of work. She points to a large pipe and says that the cooling system in the factory obeys the following rule:

If the valve is open then water will flow through the pipe

You are interested in checking whether the cooling system does follow the rule your friend has told you about. Below are four cards which refer to tests that have been carried out on the cooling system. On one side of each of these cards is recorded whether the valve was open when the test was carried out whilst on the other side is recorded whether the water was flowing at the time of the test. Please indicate by circling the appropriate card or cards, which one(s) you need to turn over to decide whether the rule is true or false.

Remember the rule you are testing is:

If the valve is open then water will flow through the pipe

The second rule received by half of the participants was as follows:

If the pipe is free from blockages then water will flow through the pipe

Participants in this latter group were reminded that their task was to test the first rule. Finally, all participants saw four cards labelled 'Valve open', 'Valve closed', 'Water flowing' and 'Water not flowing'. They were asked to indicate those card(s) which were necessary in order to decide whether the rule was true or false.

Results

We performed three analyses on our results. The first examined the effects of experimental condition on individual card selection frequencies whilst the second was of the effect of number of rules on the rate at which antecedent and consequent cards were selected. In addition, we analyzed the frequency of various card combination selections.

Individual Card Selection Frequencies: Our first analysis was of the effects of our number of rules manipulation on the rate at which individual cards were selected in the experiment. These rates are presented in Table 1 below.

Chi-square analyses showed no significant effects of number of rules on the rate at which any of the cards were selected (P card: $\chi^2(1) = 1.73, p > .18$; not-P card: $\chi^2(1) = 1.40, p > .23$; Q card: $\chi^2(1) = 1.96, p > .16$; not-Q card: $\chi^2(1) = .15, p > .70$). As the presence of a

second rule has previously been found to significantly affect the total number of cards selected by participants (Feeney & Handley, 2000) we tested for an effect of our number of

Table 1: Percentage of participants selecting each card as a function of condition.

	P	Not-P	Q	Not-Q
One-Rule	84	40	37	21
Two-Rules	74	30	50	24

rules manipulation on the total selected in this experiment. This effect was not significant ($t(109) = .365, p > .71$). The mean total for the one rule condition was 1.82 cards (S.D. = .60) and 1.78 cards (S.D. = .74) for the two-rules condition.

Cause vs. Effect Selections: To test our predictions about cause and effect card selections we analysed the rate of cause and effect selections. For the purposes of this analysis we computed the number of cause and effect selections made by each participant, where a cause selection was defined as the selection of either of the antecedent cards and an effect selection as choosing either of the consequent cards. The mean numbers of each type of selection are shown in Figure 1 below.

We performed a 2x2 mixed design Anova on the cause and effect data. The between participants variable in this analysis was number of rules whilst the within participants factor was the number of selections concerning the hypothesized cause vs. the number of selections concerning the hypothesized effect specified in the rule. This analysis produced a non-significant main effect of number of rules ($F(1, 109) = .133, MSE = .227$), a highly significant main effect of cause vs. effect ($F(1, 109) = 33.40, MSE = .385, p < .001$) as well as a significant interaction between the within and between participant variables ($F(1, 109) = 4.94, MSE = .385, p < .028$). Tests for simple effects showed that

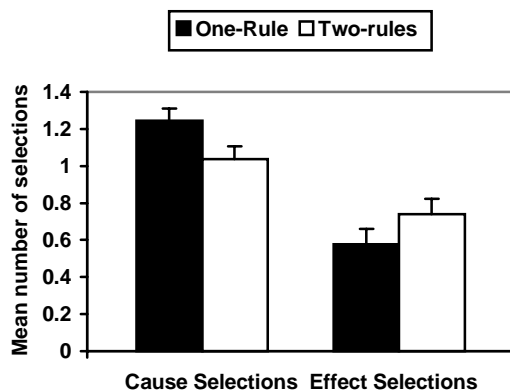


Figure 1: Mean number of cause and effect selections by condition.

there were significantly more antecedent (or cause) selections in the one-rule condition than in the two rule condition ($F(1, 109) = 4.964, MSE = .243, p < .028$). The difference due to number of rules on the rate of consequent (or effect) selections was not significant ($F(1, 109) = 1.965, MSE = .369, p > .16$).

Finally, in order to analyze the effect of our number of rules manipulation on relative rates of cause and effect card selections, we computed an index for each participant of their total of cause selections minus their total of effect selections. The effect of the number of rules manipulation on this index was significant ($t(109) = 2.22, p < .029$). In the one rule condition the mean score on this index was .667 (S.D. = .932) whereas in the two rule condition the mean score was .296 (S.D. = .816).

Card Combination Frequencies: Our final analysis was of the card combination frequency data from the experiment. As may be seen from Table 2, the rate of logically correct responding was not affected by our experimental manipulation with 4 out of 57 participants choosing the logically correct combination in the one-rule condition versus 5 out of 54 in the two-rule condition. The most striking effect of our number of rules manipulation on the combinations of cards that participants choose to select concerned the combined selection of the P and not-P cards only. In the one-rule condition 13 participants (23%) chose this combination whereas in the two-rule condition it was chosen by only 4 participants (7%). This difference is statistically significant ($\chi^2(1) = 5.07, p < .02$) and is in the direction suggested by our hypothesis concerning how people represent and test causal rules concerning mechanical devices.

Table 2: Card combination frequencies as a function of condition.

Combinations	Condition	
	One Rule N = 57	Two Rules N = 54
p	13	13
not-p	2	1
q	0	5
not-q	0	0
p, not-q	4	5
p, not-p	13	4
p, q	14	14
not-p, q	2	2
not-p, not-q	4	5
q, not-q	1	1
p, not-p, not-q	1	1
p, q, not-q	2	0
all four	1	3

Discussion

The results of this experiment support our predictions about how people test causal rules concerning mechanical devices. Across both conditions we found a significantly greater rate of antecedent than consequent selection. This is not surprising and is probably true of most selection task experiments (although for an important exception, see below). Of much more immediate interest is the finding that the difference between the rate of antecedent and consequent card selections was significantly greater in the one-rule condition than in the two-rule condition. In addition, the rate at which people selected antecedent cards was significantly greater in the one-rule condition than in the two-rule condition although the (non-significant) rate of suppression due to the presence of a second rule was the same for each antecedent card. Finally, participants in the one-rule condition were significantly more likely to select the combination of cause present and cause absent cards than they were in the two-rule condition.

We argue that this pattern of results suggests that people's representation of the rule contains information about the putative causal status of the antecedent and when considering the possible consequences of the various tests of the rule people primarily consider tests where the hypothesized cause is manipulated. When compared to the results of selection task experiments where participants are asked to test a standard indicative rule, this experiment may be seen to have produced a very high rate (40%) of not-P card selection in the one rule condition. For example, in the meta-analysis reported by Handley and Feeney (2000) of single-rule conditions from five experiments on the standard indicative selection task, 24% of the 272 participants whose data were included were found to have selected the not-P card. We argue that the elevated rate of not-P selection found in this experiment is due to the use of a counterfactual strategy to test the causal claim made in the experimental rule. In addition to testing for the presence of the effect in the presence of the putative cause, participants were interested in looking to see whether the effect was present or absent when the hypothesized cause was absent.

Strikingly, when an additional antecedent calls the sufficiency of the putative cause into question, people are significantly less likely to select antecedent cards. A reduction in P card selections is expected given that the second rule may lead people to question the sufficiency of the antecedent for the consequent to occur. However, the selection of the not-P card may still be informative regarding the truth of the causal rule. Consider for example the two possible outcomes given the closure of the valve, the absence of water flow or the presence of flow. Assuming the device is working we may expect to observe an absence of water flow when the valve is closed. However, this absence may also be caused if the

antecedent of the second conditional is not satisfied, that is if the pipe is blocked. Hence observing the absence of the effect in the absence of the cause is not informative about the truth of the rule. However, imagine instead that we observed water flowing when the valve was closed. This case would appear more informative regarding the question of whether the mechanical device is operating correctly. What our results suggest is that some participants consider only the first outcome. Hence, they regard the not-P card as uninformative and do not choose it.

Different Types of Causal Hypothesis?

Although our predictions were induced from the literature on mechanical reasoning, our results are of obvious relevance to the literature on causal cognition. For example, it is interesting to compare our results to those of Green & Over (2000) who examined decision theoretic effects in how people test causal conditional hypotheses such as the following:

(3) *If you drink from the well then you will get cholera*

Across all of their conditions, the rate of antecedent selections never exceeded the rate of consequent selections to the same degree as was true of our one rule condition. Collapsed across conditions, their rate of consequent selections was, in fact, marginally greater than the rate of antecedent selections (a total of 171 consequent selections vs. 169 antecedent selections). These results are in stark contrast to our own findings.

Green and Over's experiment was designed to test ideas concerning the relationship between the contingency table and causal hypothesis testing. Their results show that people are sensitive to the probabilities of the cause and the effect when deciding which cards to select. We believe that our results are different to theirs because we asked our participants to test a causal *mechanical* rule whereas their experiment concerned a causal *medical* rule. O'Brien, Costa & Overton (1986) have also found results suggesting that there are domain-specific differences in causal reasoning. In their experiment participants were asked what implications each of the four cases (*cause present + effect present; cause present + effect absent; cause absent + effect present; cause absent + effect absent*) had for a variety of hypotheses concerning medical and mechanical causal relationships. For all cases except the *cause absent + effect absent* case, participants were significantly less certain about the medical hypothesis than about the mechanical hypothesis.

One way to conceive of O'Brien et al's result is that people are unwilling to accept a medical hypothesis in the light of information about just a few exemplars whereas they have more confidence about the status of a mechanical hypothesis given a few confirming or disconfirming cases. In other words, causal medical

hypotheses require reasoning that is likely to be probabilistic in nature whereas causal mechanical hypotheses need not (of course it is possible to design a mechanical task that encourages probabilistic reasoning - see Kirby, 1994).

There are several factors that might be involved in causing one type of reasoning to be essentially probabilistic and the other deterministic. First, it is possible that our knowledge about organisms tells us that even in the presence of a cause, the effect might not occur. In other words, causal rules about organisms admit of many disabling conditions. In addition, illnesses have many possible causes. In Green & Over's example, someone drinking from the well might be immune to cholera or cholera might be present in the well, the local river and the nearby lake. Given all of these possibilities it makes sense that participants in O'Brien et al's study were unwilling to make decisions about the status of a medical rule in the light of information about individual exemplars. Similarly, it is unsurprising that participants in Green and Over's experiment evidenced probabilistic thinking.

Now think about testing a mechanical hypothesis. Such hypotheses are normally tested via intervention. That is, if you think that your car won't start because the plugs are dirty, you will clean the plugs and then try to start the car. If the car starts your hypothesis has been confirmed, if not then it has been disconfirmed. In either case it is unlikely that you will repeat the procedure several times in case of the operation of disabling conditions or alternative causes. Similarly, when interacting with a novel electronic device (see Klahr & Dunbar, 1988) people do not perform the same test several times in order to establish the effect of some manipulation. Their reasoning in such cases tends to be non-probabilistic.

This distinction, between probabilistic causal reasoning and consequential (or deductive) causal reasoning also relates to the literature on mechanical reasoning described in the introduction. The systems that Hegarty (1992) and Schwartz & Black (1996) required their participants to reason about were closed and so did not admit of disabling conditions or alternative causes. Of course, a particular manipulation to the system might not cause the expected effect but given the diagrams that people were shown, disabling conditions and alternative causes were unlikely to be available to reasoners. Accordingly, a strategy based on the piecemeal animation of the device in the causally appropriate direction will be adopted. For analogous reasons, our participants were interested in tests of the hypothesis about the cooling system that involved direct manipulations to the putative cause. Once the possibility of disabling conditions were introduced, they were significantly less interested in such tests.

References

- Byrne, R.M.J. (1989). Suppressing valid inferences with conditionals. *Cognition*, 31, 61-83.
- Cummins, D.D., Lubart, T., Alksnis, O. & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition*, 19, 274-282.
- Feeney, A. & Handley, S. (2000). The suppression of q card selections: Evidence for deduction in Wason's selection task. *Quarterly Journal of Experimental Psychology*, 53A, 1224-1242.
- Green, D.W. & Over, D.E. (2000). Decision theoretic effects in testing a causal conditional. *Current Psychology of Cognition*, 19, 51-68.
- Handley, S.J., Feeney, A. & Harper, C. (2000). Alternative antecedents and the suppression of fallacies in Wason's selection task. Manuscript under review.
- Harris, P.L., German, T. & Mills, P. (1996). Children's use of counterfactual thinking in causal reasoning. *Cognition*, 61, 233-259.
- Hegarty, M. (1992). Mental animation: Inferring motion from static displays of mechanical systems. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 18, 1084-1102.
- Johnson-Laird, P.N. (1993). *Human and machine thinking*. Hove, UK: LEA.
- Kirby, K.N. (1994). Probabilities and utilities of fictional outcomes in Wason's four card selection task. *Cognition*, 51, 1-28.
- Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12, 1-55.
- Mackie, J.L. (1974). *The cement of the universe: A study of causation*. Oxford, UK: Oxford University Press.
- Medvedev, G. (1991). *The truth about Chernobyl*. London: I.B. Tauris & Co. Ltd.
- O'Brien, D.P., Costa, G. & Overton, W.F. (1986). Evaluations of causal and conditional hypotheses. *Quarterly Journal of Experimental Psychology*, 38A, 493-512.
- Popper, K. (1959). *The logic of scientific discovery*. London: Hutchinson.
- Rumain, B., Connell, J., & Braine, M.D.S. (1983). Conversational comprehension processes are responsible for reasoning fallacies in children as well as adults: IF is not the biconditional. *Developmental Psychology*, 19, 471-481.
- Schwartz, D.L. & Black, J.B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, 20, 457-497.
- Wason, P.C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.

Acknowledgements

We would like to thank Lara Webber for her assistance in carrying out the experiment described in this paper.