SUMMER 2020 MINI PROJECT REPORT

MINI PROJECT GROUP-2

Gautham Venugopal
2018112003

Sahil Manoj Bhatt
2018111002

July 2020

International Institute of Information Technology,
Hyderabad

Understanding in Intuitive Scenarios : A Study of the Wason's Card Selection Task and Simpson's Paradox

Gautham Venugopal Sahil Manoj Bhatt

Abstract

The Wason's card selection task and Simpson's paradox, two very famous logic-based experiments, can be formulated as a verbal/intuitive version and a mathematical/logical version. In both these tasks, the two versions elicit different responses from participants. This study attempts to collect responses from participants and understand the reason behind their differing responses. A questionnaire was prepared with both versions of the two tasks in an attempt to find correlations between the performances by the respondents. No strong correlation was found suggesting the absence of a common factor behind the differing responses that the logical and intuitive versions of both these tasks receive.

Introduction

The Wason selection task (or four-card problem) is a logic puzzle devised by Peter Cathcart Wason in 1966. It is one of the most famous tasks in the study of deductive reasoning. An example of the task goes as follows : *You are shown a set of four cards placed on a table, each of which has a number on one side and a colour on the other side. The visible faces of the cards show 3, 8, Red and Brown. Which card(s) must you turn over in order to test the truth of the proposition that if a card shows an even number on one face, then its opposite face is the color red?*.

Previous experiments (Wason, 1968, 1969a) have established that it is very difficult to decide what information is required to test the truth of an abstract conditional sentence. In fact, hardly any individuals make the correct choice of cards to turn over (Brown and 8) in order to determine the truth of the sentence.

The rule has the logical form, “if p then q’, where p refers to the stimulus mentioned in the antecedent (D); p*, i.e. not p, refers to the stimulus which negates it (K); q refers to the stimulus mentioned in the consequent (3); and q* i.e. not q, refers to the stimulus which negates it (7). In order to solve the problem it is necessary and sufficient to choose p and q*, since if these stimuli were to occur on the same card the rule would be false but otherwise true.

Simpson’s paradox refers to a phenomenon whereby the association between a pair of variables (X, Y) reverses sign upon conditioning of a third variable, Z, regardless of the value taken by Z. If we partition the data into subpopulations, each representing a specific value of the third variable, the phenomenon appears as a sign reversal between the associations measured in the disaggregated subpopulations relative to the aggregated data, which describes the population as a whole.

A practical example of Simpson’s Paradox would be as follows. Suppose that a University is trying to favour women when hiring staff. It advertises positions in the Department of History and in the Department of Geography, and only those departments. Five men apply for the positions in History and one is hired, and eight women apply and two are hired. The success rate for men is twenty percent, and the success rate for women is twenty-five percent. The History Department has favoured women over men. In the Geography Department, eight men apply and six are hired, and five women apply and four are hired. The success rate for men is seventy-five percent and for women it is eighty percent. The Geography Department has favoured women over men. Yet across the University as a

whole 13 men and 13 women applied for jobs, and 7 men and 6 women were hired. The success rate for male applicants is greater than the success rate for female applicants.

	Men	Women
History	1/5	2/8
Geography	6/8	4/5
University	7/13	6/13

In this study, we aim to find correlations between different variants of both of these tasks.

Review of Literature

Significant amount of research has been done on finding out why people get the Wason card selection task wrong and on the paradoxicality of Simpson's paradox. This has led to multiple versions of both the Wason's card selection task and the Simpson's paradox being formulated. It was often found that some of these variants elicited significantly different responses from people when compared to the original.

One such variant is the intuitive version of the Wason's card selection task. An example of such a question is as follows : *Each card has an age on one side, and a drink on the other. Which card(s) must be turned over to test the idea that if you are drinking alcohol then you must be over 18?*.
Perhaps, surprisingly, it turns out that the performance on the Wason card selection task was highly context dependent. Compared with few individuals who were able to answer the original version correctly, a significant majority of participants were able to answer correctly when the question was proposed in certain contexts.

Of the intuitive version of Wason's card selection task itself, there are multiple variants. One such variant, attempts to present the Wason's task in the context of social relations. An example of such a variant is described above in which the recipient is asked to police a social rule, in this case the minimum age after which one can drink alcohol.

According to Cosmides and Tooby(1992), this experimental evidence supports the hypothesis that a Wason task proves to be easier if the rule to be tested is one of social exchange (in order to receive benefit X you need to fulfil condition Y) and the subject is asked to police the rule, but is more difficult otherwise. They argued that such a distinction, if empirically borne out, would support the contention of evolutionary psychologists that human reasoning is governed by context-sensitive mechanisms that have evolved, through natural selection, to solve specific problems of social interaction, rather than context-free, general-purpose mechanisms. In this case, the module is described as a specialized cheater-detection module.

Another popular variant of the intuitive version of the Wason’s card selection task is one that makes use of a story, also called the thematic version of the Wason's card selection task. An example would be where each card represents a journey taken by a person on different days. On one side of the card would be written the destination and on the other side, the mode of travel. On both sides of the cards, the day on which the journey took place would be mentioned. The success of this version of the task has been attributed to the concreteness of the terms used, along with a story emphasizing the relations between the terms used. As a result, it is hypothesized that participants are able to solve the problem in the context of a real world situation rather than in the context of cards and numbers and letters. It is important to note that in the rule(or propositions) used in the above two variants, the cards are not specifically mentioned, as opposed to in the original version.

Meanwhile, research on the paradoxicality of Simpson's paradox led to the collapsibility principle(CP). According to CP, whatever relationship holds between two variables in the sub-groups must hold for the whole group. It is based on CP, that an intuitive version of Simpson's paradox was created. An example of such a question is as follows : *There are only two high schools in a certain school district. Given that the graduation rate for girls in School #1 is higher than the graduation rate for boys, and that the graduation rate for girls in School #2 is higher than the graduation rate for boys, does it follow that the graduation rate for girls in the district is also higher than the graduation rate for boys in the district?*
Through previous experiments, it has been noted that participants are more likely to answer the intuitive version of Simpson's Paradox wrong, as compared to the logical version.

Reason behind hypothesis

The difference in the performance of participants when faced with the intuitive versions of both the Wason's task and the Simpson's paradox as compared to the logical versions of both these tasks forms the basis of our study. In this study, we aim to find if there are any correlations between any of the four tasks discussed above.

It is also worth noting that both these tasks are quite different. While Wason's task is a task on deductive reasoning, the Simpson's Paradox is based on the collapsibility principle.

Methods

We prepared a questionnaire with both the versions of the two tasks.
In addition, some survey questions were irrelevant to the target questions. We did not want students to know what we were testing them on. The order of the questions were decided so that both the versions of the tasks were separated by other questions.

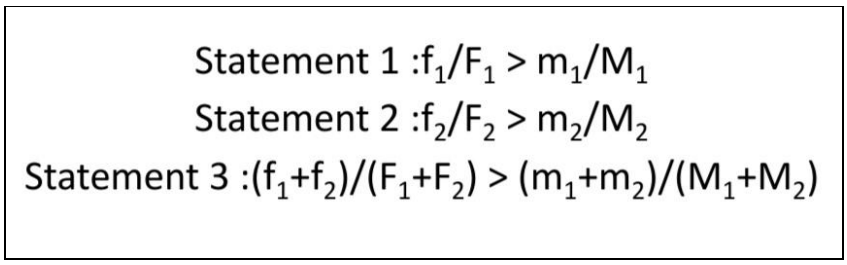


Figure-1

The other questions included one on Raven's progressive matrices, a triad test and one based on the participant's opinion on an unrelated topic.

For the logical version of the Simpson's paradox, we gave 3 statements as shown in Figure-1, and asked participants whether statement 3 would follow provided that statement 1 and 2 are true.

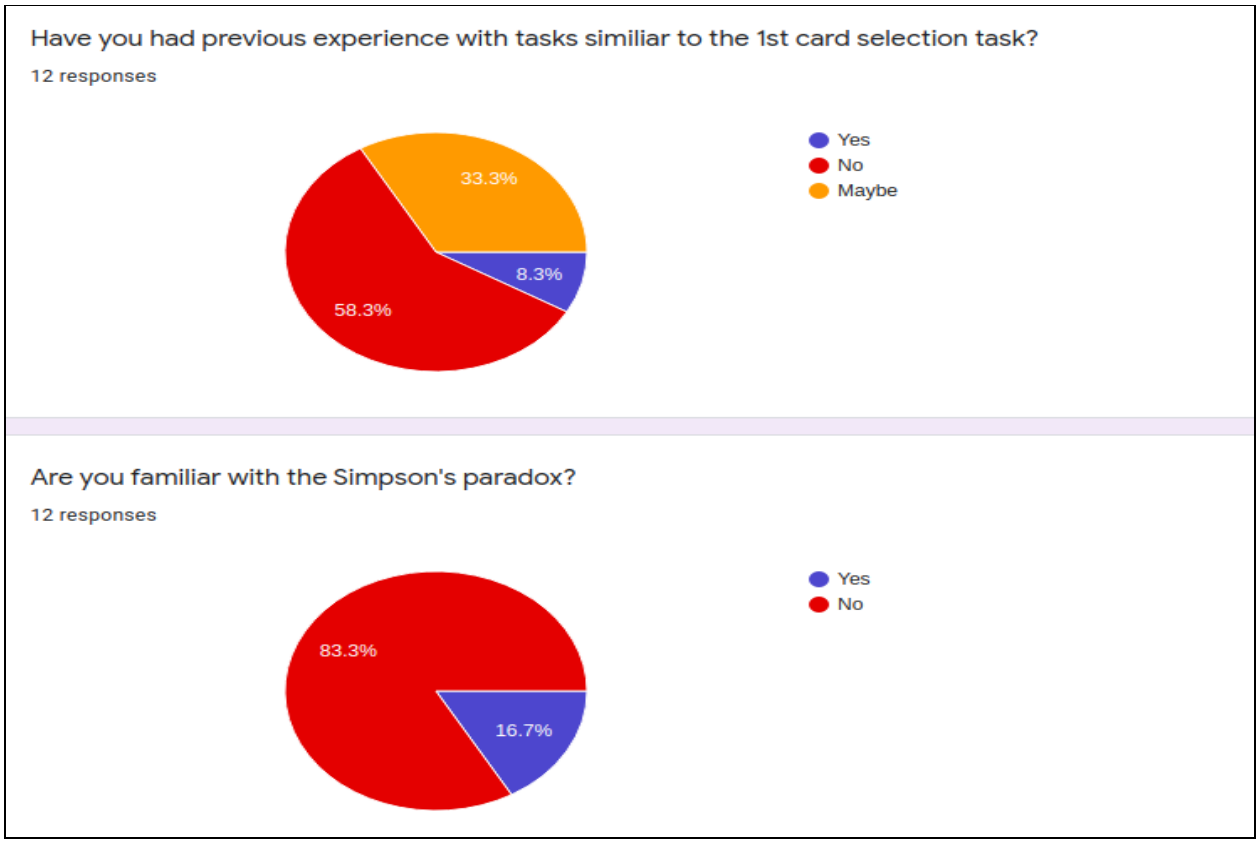
For the logical version of Wason's task, the participants were presented with the standard question involving cards marked '3','8','Red' and 'Brown'. For the intuitive version of Wason's task, we chose the one in the context of social relations.

All our participants were between 19 to 25 years of age. Most of them were students at IIIT Hyderabad, and had previous exposure to logical reasoning based tasks. The participants,however, did not have previous experience with either of the two specific tasks asked in the questionnaire.

Results and Discussion

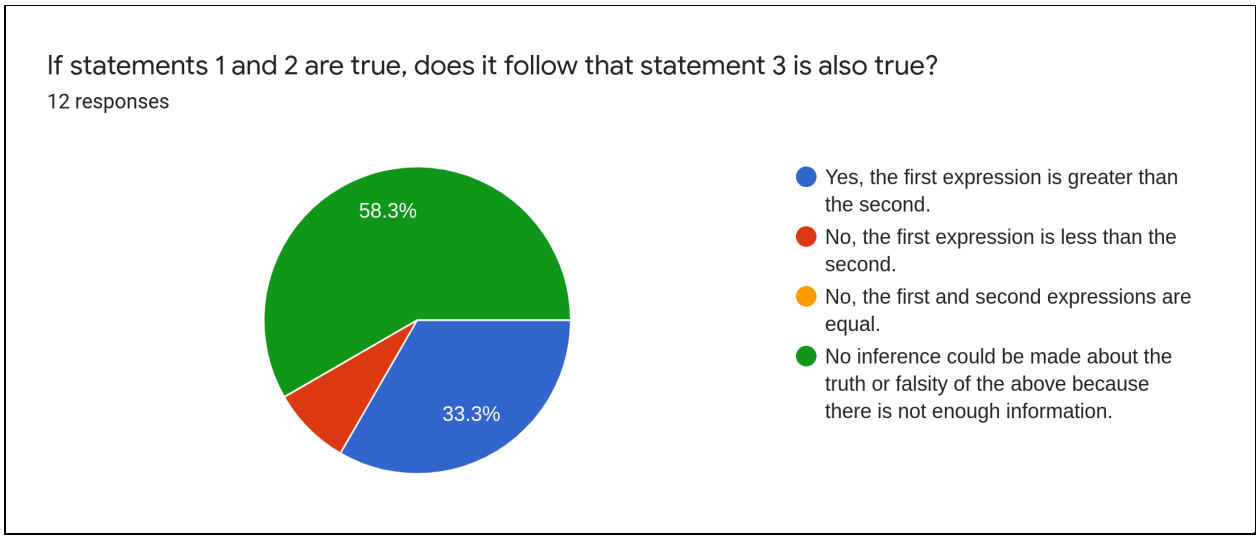
We conducted a survey of twelve people from different fields and asked them questions based on different versions of Simpson's paradox and Wason's card selection task. Apart from these, we also included questions based on IQ, and also asked them their level of confidence in the answers submitted.

Majority of the participants were not familiar with either of the two main tasks presented to them.

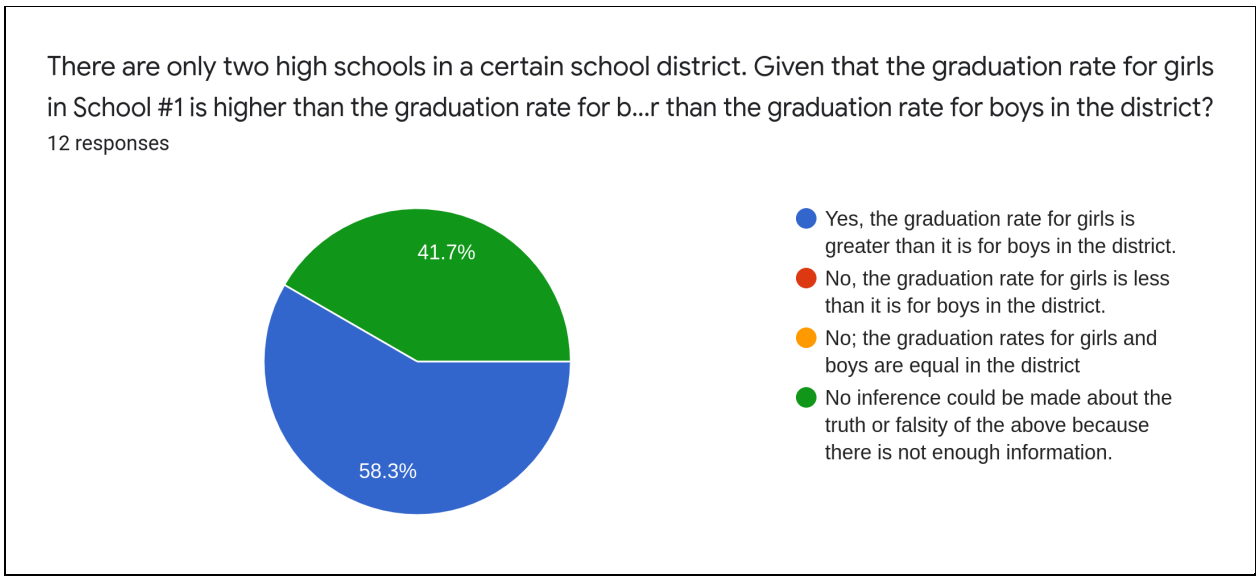


The question above were asked to check if the participants were familiar with the tasks presented to them

The performance of the participants with respect to Simpson's paradox was largely consistent with that seen in previous experiments. A majority of the participants were able to answer the logical version of Simpson's paradox correctly. However only 5 out of 12, were able to answer the intuitive version of the paradox correctly. Also, all the participants who got the intuitive version correctly were able to get the logical version also correct.



Participants' response to the logical form of the Simpson's paradox



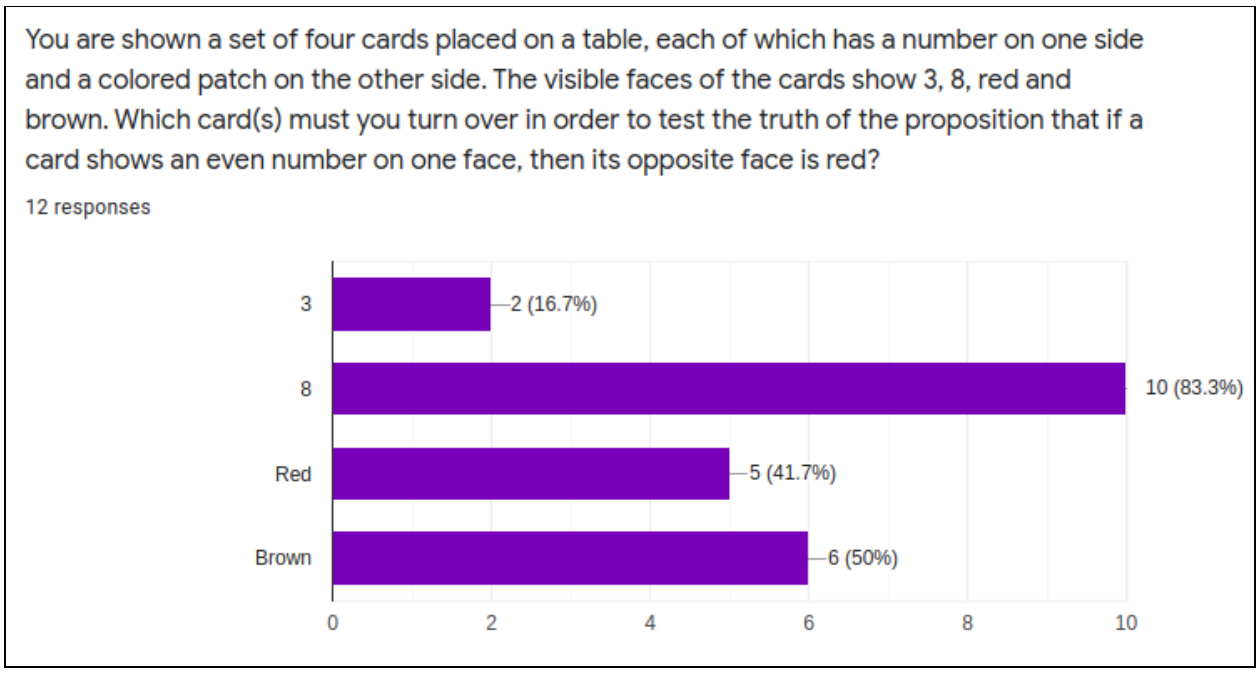
Participants' response to the intuitive form of the Simpson's paradox

	Correct	Incorrect	Total
Logical	7	5	12
Intuitive	4	8	12

Results of Simpson's Paradox

The responses to the confidence scale were also as expected. The mean confidence scale for the intuitive version was 9.08 as compared to the mean confidence scale of 7.67 for the logical version.
For the intuitive version, the mean confidence of the participants who answered incorrectly was 9.57 as compared with 8.4 for those who answered correctly. This result is consistent with Hilary Kornblith's conjecture that our being reflective does no more than reinstate what we have already believed.

The performance of the participants in the Wason's task also was consistent with previous experiments. 4 out of 12 participants were able to get the logical version of Wason's task fully correct, ie, chose both of the correct options. This means around 30% of the participants were able to get the logical version correct, whereas in previous experiments, only 10-15% of the participants were able to get it correct. We believe that a reason for this could be that most of the participants must have had previous exposure to similar logical reasoning-based questions, due to which answering this task would have been relatively easy.



Participants' response to the logical form of the Wason's card selection task

	Correct*	Incorrect*	Total
Logical	4	8	12
Intuitive	5	7	12

Results of Wason's Task

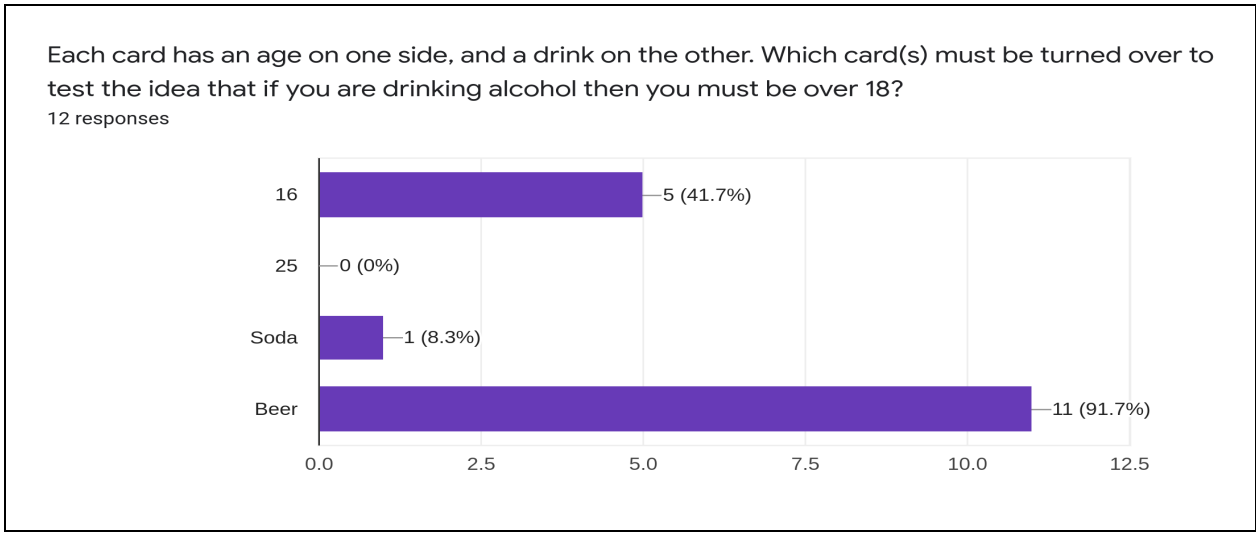
*(*means that the response included <u>both</u> the correct cards and not just one of them.)*

The response though is very different in case of the intuitive version of the card selection task, which involves cards with ‘Age’ marked on one side, and drinks (alcoholic and non alcoholic) marked on the other.
Clearly, almost all participants agreed with turning over the card marked ‘Beer’ in order to test the proposition. Nearly half of them agreed with turning over the card marked ‘16’ as well.

Although not everyone chose both the correct cards as their answer to the intuitive version, we observed that everyone chose at least one of the two correct cards.

However, the reason we believe that both correct cards were not chosen could be due to the nature in which the task was presented. Usually, the task is presented in a physical environment using paper or cardboard cards. Also, participants are often presented with a few sample cards before the actual task is presented. This would've helped the participants understand the question better.

It should be noted however, that the intuitive version had received less incorrect responses than the logical version, in line with what we had expected to observe.



Participants' response to the intuitive version of the Wason's card selection task

Coming to confidence scales, there was not much difference in mean confidence between the two versions. The mean confidence for the logical version was 8.58 compared to a mean confidence of 8.5 for the intuitive version. Interestingly, a high correlation was observed between confidence and performance. Of the 4 who got both the logical and intuitive versions correct, all of them gave confidence scales of 10 for both the versions. The person who got only the intuitive version correct had a confidence of 9 for the intuitive version and a confidence of 10 for the logical version.

Overall, no reliable correlations were found between the two tasks. 3 out of the 8 people who got the intuitive version of Simpson's paradox incorrect and 2 out of the 5 people who got it correct, were able to answer the intuitive version correctly with high confidence. However, it should be noted that only 25% of the participants who got the logical version of Simpson's Paradox incorrect got the intuitive version of the Wason's card task correct, as compared to 50% of the participants who got the logical version of Simpson's Paradox correct. If this result could be replicated in a larger study, then it would lend support to the argument that the use of concrete terms in the intuitive version of the task makes the question more

amenable to the application of the general intelligence used in mathematics.

	Those who got the intuitive version of Simpson's Paradox correct.	Those who got the intuitive version of Simpson's paradox wrong.
Those who got the intuitive version of Wason's correct.	2	3
Those who got the logical version of Wason's correct.	1	3
Total	5	7

Conclusion and Recommendations for further research

The results of the study seem to suggest that the reasoning involved in the Wason’s card selection task and Simpson's paradox are quite different.

For further research, one could attempt to modify the questions associated with the Wason's task so that their results would reflect exactly what was seen in previous studies.

One could also try using different variations of the intuitive version of the Wason's card selection task. If the correlation between the intuitive version of the Wason's card selection task and the logical version of the Simpson's Paradox holds, then, it would lend more credibility to the significance of the concreteness of terms used in the Wason's task.

References

1. PC Wason and Diana Shapiro(1971), Natural and contrived experiences in a reasoning problem
2. Stanford Encyclopaedia of Philosophy entry on Simpson's Paradox
3. Judea Pearl(2014), Understanding Simpson's paradox, The American Statistician
4. Branden Fitelson and James Hawthorne(2010), The Wason task(s) and the paradox of confirmation
5. Simpson's Paradox is suppression, but Lord's Paradox is neither: clarification of and correction to Tu, Gunnell, and Gilthorpe (2008)
6. Cosmides, L.; Tooby, J. (1992). "Cognitive Adaptations for Social Exchange" (PDF). In Barkow, J.; Cosmides, L.; Tooby, J. (eds.). The adapted mind: Evolutionary psychology and the generation of culture. New York: Oxford University Press. pp. 163–228. ISBN 978-0-19-506023-2.
7. Kornblith, Hilary. (2010) 'Why Reflective Endorsement Cannot do'. Philosophy and Phenomenological Research, 80, 1-19.
8. https://medium.com/homeaway-tech-blog/simpsons-paradox-in-a-b-testing-93af7a2f3307
9. https://en.wikipedia.org/wiki/Simpson%27s_paradox
10. https://en.wikipedia.org/wiki/Wason_selection_task