LUNDS UNIVERSITET

DEPARTMENT of PSYCHOLOGY

Reasoning Through the Application of Crystallized and Fluid Intelligence in a Causal Inference Task

Martynas Matuzevicius

Master's Thesis (30hp)
Spring 2023

Supervisor: Jean-Christophe Rohner

Abstract

The aim of the study was to find out how well individuals perform when solving hypothetical causally related problems, and further understand what intelligence factors interact with this. This was done via a survey that consisted of hypothetical, linear relation, causal scenarios. Two experiments were conducted to ensure that no validity issues confounded the results. Experiment 1 consisted of 77 participants who performed unexpectedly well across all conditions. This may have been due to a later identified heuristic simplifying the level of reasoning required, hence Experiment 2 was conducted. Only 19 participants were used for the follow up experiment, thus may need to be further replicated with a larger sample size. The results however have shown that negative-negative-negative condition was the most challenging among the others presented, resulting in the lowest overall mean (M = 0.8). Causal reasoning in the context of this study shows that most individuals are able to correctly infer the relationship A to C when B is an intermediate cause, and A to D when B and C are intermediate causes. Future studies should focus on validity concerns and explore more controlled conditions for the current paradigm.

Keywords: causal reasoning, inference, crystallized intelligence, fluid intelligence, heuristic

Reasoning Through the Application of Crystallized and Fluid Intelligence in a Causal Inference Task

Planning requires us to create step by step instructions that can be followed, preferably with ease, to achieve a desired result. The current study made use of causal reasoning procedures which can be applicable when trying to understand how a specific set of events follow one another in logical form. This form of reasoning has been defined as an influence of one variable, acting as an effect, on another variable which represents a cause, with a consequent change in the latter (Pearl & Mackenzie, 2018). It represents more, through observable evidence, than what a simple correlation may be able to show (Sloman & Hagmayer, 2006). An example to illustrate a similar relationship that shares the structure to the current study can be used, where A, B, and C represent different intermediate locations. For example, departure from Paris leads to an arrival in Florence, in turn a departure from Florence leads to Prague. This may be a simplistic example, however knowing that Florence provides the necessary transport from Paris to Prague, we can causally infer that Florence is a stop that will have to be considered in the plan, if no other information is given or no alternative route is available. Understanding how we can measure intelligence and how to apply it in abstract as well as pragmatic reasoning tasks is important, as it would allow us to influence education systems, thus improving how different establishments relay information to various individuals.

The current study is concerned with binary relationships where an implied causal effect could mean either an increase or decrease based on a causally related premise. Positive relations (increases) may be more intuitively comprehensive compared to negative relations (decreases). Negation may be a concept where a more elaborate understanding is necessary, thus explaining how it functions is fundamental when causal meaning is of concern (Barbey & Wolff, 2007). An implied linear relation from an example, which asks to infer the relation between A and C, describing that an increase in A causes an increase in B, which in turn causes an increase in C is not cognitively demanding. This is due the idea that we can simply infer a positive relationship between A and C. On the other hand given a scenario where an increase in A causes a decrease in B, which in turn causes a decrease in C, is less intuitive. The reason for this lies at the interaction between all set points when we infer the relationship between A and C. This is due to the two negative relations between A and B as well as B and C to form an inferred positive relation between A and C. Similarly mixed relations where increases and decreases occur may have similar consequences, and prove to be more difficult to reason through, especially when more relations are involved in the causal scenario.

The Role of Causal Reasoning and Intelligence in Problem-Solving

Causal scenarios structure the world around us and through inference we gain an insight to how different events interact with one another. Oftentimes we can anticipate how these interactions intertwine through the application of general intelligence factors, such as crystallized and fluid intelligence. Crystallized intelligence is known as previously acquired knowledge available for recollection and applied for problem solving. Fluid intelligence refers to a form of thought flexibility and the ability to adapt as well as develop new problem-solving techniques for novel scenarios (Sligh et al. 2005). The way we apply our knowledge in the world can be understood through different forms of causal reasoning. We can define this form of reasoning as the process that connects a cause to a subsequent effect, through an information organising process that helps us understand how these causes and effects interact. Through the process of inferring, estimating, or judging observed information, we can reorganise the way that we perceive it (Anshakov & Gergely, 2010). The ability to reason causally allows us to problem-solve and make accurate decisions. Additionally, reasoning to solve a problem may be completely reliant on how easily we can access the information needed (Sloman, 2005). Therefore, crystallized intelligence plays a very important role for our general understanding of any given situation and will be thoroughly discussed throughout this thesis.

Alternatively, Sligh et al. (2005) argue that a higher general knowledge base of an individual increases their likelihood of novel idea generation. In support of this Buades-Sitjar and Duñabeitia (2022) identified a common influence on our approaches to dealing with and overcoming new situations that require an application of fluid intelligence. They explain that knowledge such as spoken language, engrained cultural norms, and skills which an individual acquires contribute to how we reason about upcoming novel events. As such, it can be assumed that creativity is enabled by crystallized intelligence and our knowledge capacity, sequentially, fluid intelligence by creativity as we are able to recollect more ideas in new unique patterns (Sligh et al. 2005). Further claims and discussions situated around fluid intelligence will be introduced in the later sections.

A study by Deary et al. (2010) was carried out to provide evidence for a sole general intelligence factor. It is presented as a compound ability of crystallized, fluid intelligence and other underlying factors, also found to be variable based on heritability and sex. This neuroimaging study shows us conclusive evidence that provides a strong basis for a need to deepen the understanding of what we know about intelligence. Previous research has tested crystallized and fluid factors of intelligence through an application of different terms, such as cause, enable or allow and prevent (Barbey & Wolff, 2007; Sloman 2005; Sloman et al. 2009). Researchers have found that terms such as cause, enable or allow are active words, so that the likelihood of a follow up effect is increased, contrastingly the prevent term implies a decrease in likelihood. These studies investigate causal attribution and the perception of the most likely causal event for tested scenarios, aiming to understand how causal relationships and any plausible derived implications can affect our reasoning abilities, and thus decision making. Similarly, in the current study the explanation of directionality in a causal relationship is important, on the contrary, theories of probability are not so much. Sloman (2005) states that human thought is driven by causal models that guide the representations of how causal relationships in various situations intertwine. Hence, through reasoning it is possible to infer what the likely outcomes of those relation interactions may be. Inferring a direct relationship from previously known causes and effects is exemplary of crystallized intelligence. Through this process we can apply obtained knowledge and use it further in novel scenarios, thus leading to an integration of fluid intelligence. Numerous different fields such as science, entertainment, or planning make use of inference and reasoning abilities, where specific follow up events are affected by previous actions. To be able to grasp where the limits of inference abilities lay, we must explore simple yet pragmatic solutions. By applying findings of the mentioned studies, I aim to enhance the understanding of how individuals express their crystallized and fluid intelligence ability, as well as their application of causal reasoning in a set of hypothetical problems.

Dynamics of Memory, Causation, and Intelligence

When encountering a problem, we can simply rely on our natural understanding of causation, given that the problem requires us to make inferences to help us arrive at a conclusive response. Previous research has established that logical errors frequently occur in scenarios such as: “Some A are B. No B are C. Therefore some A are not C”, where simple deductions must be made (Evans, 2002). It is assumed that this may be due to novelty in the way that the content is presented, as not many people would naturally come across such a problem on a daily basis. Since the problem is phrased in simple terms, where the relations are also easily identifiable, it can only be assumed that errors occur due to a lack of further reasoning beyond what previously acquired knowledge can offer. In this case, knowledge and memory can be used as interchangeable terms, meaning that everything that has been memorised forms the base of knowledge. Once specific memories are acquired, they can be applied to problem-solve, however, within a limited scope of what those memories entail and given the context of application. As such, an individual with specific memories in biochemistry would be able to solve biochemistry problems. Furthermore, those specific memories would fail to be applied correctly in a case of a social science problem for example. In respect, some would argue that memory and intelligence are one and the same factor (Ackerman et al., 2005). And so, it was proposed that intelligence is no more than what we can recollect in the moment. Evans (2002) outlined a question of interest, how competent are individuals without any formal training in causal logic? To which I now pose a subsequent question, would the main difference for those with training in causal logic rely solely on memory when encountering similar yet unfamiliar scenarios? In the current study I will use a set of novel hypothetical problems which similarly to the example illustrated earlier ("Some A are B. No B are C. Therefore some A are not C") will have no real relation to any factual events. Thus, seeing how different individuals reason through the problems will provide evidence for another approach to intelligence testing.

Memory is a fallible aspect of cognitive function, and consequently so must be intelligence, in respect to how we measure it. Although, crystallized intelligence is arguably less prone to changes across a lifespan, being more resistant to cognitive function fluctuations caused by an infliction of stress or mood induced changes (Xu et al., 2023). Therefore, under controlled conditions we should to an extent expect some consistencies in how individuals perform in specific reasoning tasks where previous knowledge can be put into effect. Evans (2002) explains that participants make logical errors systematically rather than at random. Such errors can be attributed to misinterpretation, cognitive bias, or the complexity of information in a given task. It may not be completely accurate to claim that a problem, which is not understood in the first place, measures some form of intelligent reasoning ability. It is important to note that, albeit reasoning in distinct ways, different people might come to similar conclusions. The existence of such systematic errors, however, does not suggest the evidence of irrationality, but that previous knowledge and belief are some of the factors influencing reasoning, irrespective of the structure of logical problems (Evans, 2002). The current study is set out to capture fluid reasoning, although we must reflect on crystallized ability too, and understand how they differ if we want to discuss them separately and effectively.

A study has found that the neuroanatomical structures in the brain have very different reactivity when crystallized intelligence is enabled, compared to fluid intelligence (Xu et al., 2023). Additionally, Deary et al. (2010) outline that intelligence uses a multitude of brain regions, where an extensively interconnected network in the parieto-frontal region has a notable basis for intelligence. From this we can infer that the same cortical reactivity does not support both types of intelligence, as well as crystallized intelligence being distinctly different to fluid intelligence on cortical level. Geake and Hansen (2010) were measuring whether a difference in activations occur when participants are presented with analogical stimuli of fluid or crystallized intelligence through a neuroimaging study. Examples of both have been provided, where an exact analogy 'Hot is to cold as summer is to winter' represents crystallized form of reasoning. If we are presented with this example, it would not be difficult to also solve for the equivalent of words such as summer. This has little novelty therefore previous knowledge can easily be applied. Fluid form on the other hand was analogised through examples such as 'What is the Mozart of painting?'. This scenario could lead to various conclusions that only partially rely on crystallized intelligence, but further require one to construct a model where an equivalently perceived artist, not a musician, paints on par to someone like Mozart and his composed music. Geake and Hansen (2010) therefore concluded that to arrive at any given response, some form of fluid reasoning must occur, as a correct response does not necessarily have to be completely accurate. The response may be based on any given number of specific criteria, such as skill of Mozart, his ability to find harmony, his technical proficiency, or ability to evoke emotion through music. Through the mental process of abstraction this could be compared to composition, texture, or colour balance in paintings by Picasso. Geake and Hansen (2010) also state that through use of analogies we can better our understanding of the fundamental process underlying intelligence. This may be a broad argument which does not concern itself with a plethora of factors underlying intelligence, although their point of focus may lead to well distinguished arguments which are more simplistically tested against. Their analyses show that the fronto-parietal cortex reacted differently to fluid and crystallized analogies sustaining the results provided by Xu et al. (2023). Analogies are therefore an effective method to test how intelligence differs between creative and knowledge-based conditions. Implications of these results support my current study by providing a neuroscience perspective of the neural mechanisms of intelligence and guides the understanding of the cognitive functions correlated with their respective neural substrates.

Despite the findings on how intelligence differs on the cortical level, researchers have reflected on how these neural connections highly depend on one another. Higher levels of fluid intelligence may be the key to higher crystallized intelligence (Nisbett et al., 2012). It has been further proposed that an increase in crystallized intelligence depends on an individual ability to reason fluidly, however given that fluid intelligence begins to fall, the knowledge base of an individual would not be affected. Thus, the individual could continue to accumulate knowledge further irrespective of a decrease in fluid intelligence. Hence, explaining an important aspect of the relationship between these two neuroanatomically distinct cognitive processes. The extent to which crystallized, and fluid intelligence depend on one another is not clear. Sligh et al. (2005) discusses that practice in creative problem-solving leads to a higher increase in fluid intelligence, and not so much crystallized intelligence, which may be more naturally assumed. This is not an intuitive argument and so Nisbett et al. (2012) promotes the idea of a direct link where any newly obtained knowledge leads to a higher general knowledge base, thus increasing primarily crystallized intelligence. Since growth of crystallized abilities are not assumed to be independent of fluid intelligence, there may be other factors influencing how newly obtained information from solving novel situations is transferred to crystallized intelligence (Schipolowski et al. 2014). Although the original conclusion by Sligh et al. (2005) still stands firmly, because applying fluid intelligence to formulate novel problem solutions could lead to an enhancement of itself, which in turn suggests an improvement in reasoning and mental flexibility. This is a very plausible point to consider because certain tasks may not be very informative in terms of providing further knowledge. However, the capacity necessary for solving in a creative manner may still be extensively used, and therefore improved more considerably in contrast.

The current study has a focus on fluid intelligence as it enables participants to reason outside of any known causal relationships that exist, thus having a low impact on what is previously known. I make use of novel words such as alomerance or isolance and form novel relations between these words through an application of increase and decrease verbs to define these relations. Since both factors of intelligence are arguably overlapping, it may not be completely accurate to state the extent to which either of those are predominant (Kyllonen &

Kell, 2017). Although low impact on crystallized intelligence does not necessarily mean that previous knowledge plays a minor role in deciphering a given problem. Specific knowledge can sometimes be sufficient to solve a problem. Professionals in various fields can construct domain specific predictions based on causal influences that they have trained on (Sloman, 2005). An athlete for example can anticipate specific ball motion based on predefined causal models, without having to rely on any advanced physics equations. Sloman (2005) expands on his argument suggesting that despite a likely misrepresentation of true events that follow, individuals construct a causal frame based on preconceived knowledge. What this means in practice is that people still attribute causality even without practice, which in turn could fail to solve a dynamic problem such as anticipating a ball trajectory.

The way that the current study has been designed requires some form of understanding about what kind of causal influence an event A has on an event C, when B is an intermediate causal event. This can be completed in various methods, such as drawing out linear graphs or drawing arrows to visually see what increases and where a decrease may occur. However, the problem at hand can be solved using mathematics. An example where an increase in A causes a decrease in B, and a decrease in B causes a decrease in C can be used illustrate this. We can express A to B relationship as $B = 1 * A$, where one denotes a positive relation. For B to C relationship we use the following $C = -1 * B$, in which case the minus one denotes a negative relationship. Next, we can substitute B into $1 * A$, leaving us with $C = -1 * 1 * A$. Solving this results in $C = -1 * A$, therefore indicating a negative relation between A and C. Memory is a capital from which we can retrieve information, specifically, the desired equation for the problem (Sloman, 2005). This is solely based on the idea that there has been some form of training with mathematical expressions which are applicable for the correct solution to the problem. This method fails when no equation matches the problem being solved. Thus, more creative problem-solving solutions discussed by Sligh et al. (2005) may be necessary to further assess the given problem. This process outlines that if causal reasoning cannot take place through use of memorable information, an application of more abstract reasoning is needed.

Interactive Implications of Working Memory and Fluid Intelligence

In the previous section I have discussed that memory is relevant when considering various arguments of intelligence. In recent literature, it has been noted numerous times that fluid intelligence and working memory specifically are concepts that share a lot of common factors (Nisbett et al., 2012). It was said that working memory acts as a storage and a medium for filtering out irrelevant information which prevents incorrect responses from surfacing. In the current study I make use of various terms that hold no meaning in relation to one another or have been completely made up. Thus, the terms themselves may become redundant given that this function of working memory is applicable to focusing attention on relational terms irrespective of the domain specific word variables. It is not quite clear whether the proposed relationship between intelligence and working memory is evident given that intelligence is highly dynamic ranging from learning how to use an instrument to creating social bonds Roth (2015). Additionally, Stadler et al. (2015) argue that a simple use of an everyday object such as a phone or computer require a certain degree of knowledge, especially to understand how these objects interact with everything else around us.

The distinction between working memory and intelligence is relevant to the current study as working memory may be a construct that encompasses fluid intelligence. It is important to clarify this as the focus of the research should be made clear through either involvement of the additional concept or otherwise its dismissal. Evidence from Fry and Hale (1996) states that the process of aging into adulthood improves a major influence on fluid intelligence, which was found to be processing speed and an improved functioning ability of working memory. Furthermore, when age-related differences were accounted for, results still suggest that processing speed of information has a direct effect on short-term recollection, thus a direct influence on problem-solving ability can be inferred. As such it can be argued that working memory has a strong similarity to intelligence or may even be viewed as the mechanism that fuels fluid intelligence. Kyllonen and Kell (2017) provides evidence for working memory training as a means of improving the fluid intelligence quotient. However, opposing evidence suggests that even though improvements in short-term working memory can be observed after some training, no directly transferable abilities can be assumed to influence fluid intelligence (Hulme & Melby-Lervåg, 2012).

Claims have been made that the neuroanatomy of working-memory and fluid intelligence have a lot of resemblance; however, these have been rejected by Burgess et al. (2006). They carried out a study which identified that neuroscience results hold much more complexity than what was originally believed. Therefore, continued progression on exploring via neuroscientific approaches is important as it can help us understand the neural basis of this thoroughly discussed topic. I further assert that abstract thinking is reliant on working memory capacity despite of age, which may have developmental implications for individual differences. Correlational analyses have shown that there is a substantial overlap between reasoning in different individuals and the information capacity they are able to retain in their working memory (Evans, 2002). This has been found to be the case while information is in the process of some kind of transformation within the scope of working memory. Thus, abstract thinking may simply rely on the most appropriate combination of temporarily recalled items, irrespective of their meaning in face value alone. Furthermore, having a higher capacity in working memory can potentially result in a higher number of abstract combinations that can be temporarily held. If we apply a simple mathematical factorial expression to the ‘magical number’ seven by Miller (1956) which is still a highly discussed number and relevant to the limited capacity value of working memory (Ma et al., 2014), it becomes clear as to how much more information a single item can introduce. An alternative recall limit has also been proposed by Cowan (2010), suggesting that the value is more likely to be around four. We can infer that recollecting six items would be equal to the ability of formulating 720 representations based on the function *6!*, whereas seven items would result in a substantially higher figure of 5,040 (*7!*). Salthouse and Pink (2008) suggests that the relation between working memory and intelligence is not reliant on the quantity of information maintained. Regarding this however, we must be able to distinguish between contextual similes and actual phenomenon that are empirically supported. The capacity of working memory, and thus fluid intelligence may be dependent on factors such as genetics, although a major influence could also be said to occur from childhood into adulthood. Aging may play an important role in sustaining memory improvements into adulthood, serving as an underlying mechanism that helps explain the relationship between working memory and intelligence. To establish a more thorough theoretical framework further hypotheses have been pushed forward. First, processes concerning memory, such as search and retrieval, second, focused attention on relevant information regardless of distractions (Mogle et al., 2008). When tested, these theories may help us explain the underlying mechanisms of intelligence. In relation to the current study, individuals with a greater working memory capacity would have more possible means to reason through every problem provided. This would essentially mean that more cognitive resources are invested towards problem-solving, however with an enhanced likelihood of reasoning through the problems correctly.

Cognitive Flexibility and Abstract Reasoning

When a causal problem is encountered, some form of reasoning occurs to help us understand how an event A is associated with an event C, given that B acts as an intermediate cause and effect. The association between these terms is not always evident, thus the following argument applies: novel situations require solutions that make use of abstract thinking. As previously mentioned, it has been speculated that negative causality is much more difficult to comprehend compared to positive, as it is far easier to reason about events that take place and follow one another, rather than events that do not happen because of their predicates (Sloman et al. 2009). Language may therefore be an issue when testing negative relations, as articulation is not intuitively simplistic. Ability to reason may be more challenging in negative terms (Barbey & Wolff, 2007), however not impossible as reasoning with these terms occurs on a daily basis, such as not going to work decreases the chance of one getting paid, or alternatively prevents the employer from simply paying. This scenario depicts a factual event which may occur, however reasoning in new, possibly hypothetical scenarios may not so much depend on the same external factors. Consequently, the ability to adapt previous knowledge becomes the key to the extent of reasoning that one can demonstrate in these types of contexts (Evans, 2002). This may hold some truth when it comes to vague or speculative scenarios, where factual relation understanding of A, B, C is impractical to use.

New knowledge does not always mean that a new stimulus must be observed, instead previous knowledge can act as a mode of learning, where new ideas are developed (Barbey & Wolff, 2007). If we observe the following hypothetical example: “An increase in integrity causes an increase in interest. An increase in interest causes a decrease in amnesia. What effect does an increase in integrity have on amnesia?”, we can infer a novel relationship between previously known words despite their factual causal relations, due to the way that problem is presented. This process can be explained by rearrangement of simple representations that one holds, where a new arrangement of representations leads to new conclusions. Additionally, this can be expressed as a relearning of how a concept becomes structurally coherent through a different, possibly hypothetical, rearrangement of its building blocks. Thus, Barbey and Wolff (2007) view this as a form of reasoning, which eventually becomes newly acquired knowledge without any or much external stimulus. Through this process we may be able to gain a more detailed insight into different domains of logical reasoning, where we can explain and possibly measure how the ability to grasp a concept improves with each idea generation.

As previously mentioned, fluid intelligence has a high reliance on crystallized intelligence thus a likely scenario of abstraction is that some sort of symbol manipulation takes place, and in turn the correct answer is eventually formulated (Barbey & Wolff, 2007). Symbols in this context refer to smaller parts of a given problem, such as memorable events or ideas that could bring about a change to the way that the problem is perceived. It is therefore a mental ability providing an easier layout to solve a problem. Additionally, the way that the intelligence factors interact here is by aiding one another. To illustrate this, we can view crystallized intelligence as a box full of Legos, where fluid intelligence acts as the instruction manual suggesting numerous ways in which they can be attached together. If for example we face a problem such as: “An increase in stelimity causes an increase in isoliritance. An increase in isoliritance causes an increase in alomerance. What effect does an increase in stelimity have on alomerance?”, then inferences based on observed items does not work. However, provided that we have knowledge of dealing with linear relationships, the answer is not difficult to deduce. Alternatively, without such knowledge more creative solutions, such as visualising information in various ways, are needed, although the range of ideas that may be applicable in this case are vast. This idea works when all necessary knowledge is acquired prior to solving a given problem, or at least the necessary components required to formulate an accurate response are available for the so-called symbol manipulation. Alternatively, the solution to the problem would turn out incorrect, unless achieved by chance. Furthermore, some form of pragmatic thought must still be evident for the solutions to turn out correct. Evans (2002) argues that reasoning abilities are variable, where experience plays a vital role in its development. Additionally, they state that exercising these abilities lead to more efficient reasoning. This supports the idea that the intelligence factors are not independent of each other, but instead systematically provide input based on the approach needed to problem-solve. Schooling, working-memory training, certain pharmaceuticals (wakefulness inducing) and attentional factors have been identified as potential influences of fluid abilities (Kyllonen & Kell, 2017), including the ability to reason in abstract thought. These factors are either reliant on practice, which in turn builds experience, or levels of awareness related to efficiency when reasoning.

Sloman (2005) proposes that not all situations demand the same type of logic, and some forms of logical reasoning are not inherently intuitive to human beings. Causal reasoning on the contrary is supposedly more naturalistic. He also states that it may not necessarily be the reasoning itself that is unnatural, but the situational applicability of specific type of logic which fails to be applied when necessary. If for example a causal inference must be made, but an individual relies on deontic logic, which deals with principles of ethics or lawfulness, due to a supposed ethical concern of a relevant issue, then the accuracy of the logical solution will be biased by the type of logic used to reason in the situation. Furthermore, to what extent then is logical reasoning transferable across conditions? Or is causal reasoning clearcut where one form exists without a direct association to another? And would some of those logical systems be more hardwired to implement previous knowledge whereas other to abstract thought?

In the current study I have set out to answer four research questions situated around causal reasoning ability. First, will there be a difference in accuracy scores between positive, negative, and mixed relation scenarios in two relation conditions? Additionally, will such a difference be evident for three relation conditions? Second, will the longer scenarios consisting of three positive relations affect accuracy scores more than two positive relation scenarios? Similarly, will three negative relation conditions result in more errors compared to two negative relation conditions?

Experiment 1

Method

Participants

A total of 77 participants took part in this study, 34 male, 39 female and 4 who identified as other. The age ranged from 16 to 64 (M = 29.60, SD = 10.47). Participants were recruited through an opportunity sampling method via an anonymous link distributed on social media platforms (Facebook and Reddit). No common occupation categories could be identified from the responses provided. Similarly, there was a broad range of study backgrounds, with no predominant identifiable categories. However, majority of participants had obtained a bachelor’s degree (50%). Fewer had the highest level of education as master’s degree (25%), and only several doctoral degrees (7%) were reported. Remaining participants had a high school diploma or equivalent (18%). Participants that did not completely respond to all the questions were excluded from data analysis.

Materials

This study was designed using a software programme Qualtrics. It required participants the use of a computer or a smartphone and access to internet to complete.

A total of 24 separate causal inference problems were formulated for this study . Each problem consisted of two or three sentences that described the relationship between different word variables. Descriptions were binary in nature indicating either a positive or a negative relationship in terms of increase and decrease (see Table 1). Twelve of these causal problems contained three separate variable words with two relations: Positive-Positive (PP), Negative-Negative (NN), Positive-Negative (PN), and Negative-Positive (NP). The remaining twelve problems contained four separate variable words with three relations: Positive-Positive-Positive (PPP), Negative-Negative-Negative (NNN), Positive-Positive-Negative (PPN), Negative-Negative-Positive (NNP). Table 1 shows all experimental conditions and example problems for each condition.

Word variables were chosen specifically based on their definitions, which could then allow us to categorise them in the three following domains: psychological, natural science, and a made-up domain, where each encompassed all eight conditions once. Every domain was presented randomly for counterbalancing purposes to reduce any order effects. Similarly, the conditions within the domains were randomised in their respective blocks to avoid order effects. Furthermore, before the causal inference sentences could be formulated, the word variables were listed in their respective domains. Following steps involved pairing of words, once again within their respective domains, to ensure that non-causally related combinations of words in psychology and natural science domains were possible. Once it was clear that plenty of possible combinations existed which involved no real causal relations, I used SWISH Prolog application to randomly generate three- and four-word sequences for their respective domains. These sequences were also checked at face value, for whether they accidentally represented an actual causal relation (e.g. "A decrease in volume causes a decrease in pressure"). In that case a sequence was removed so that all problems represented novel causal relations. Every causal problem was later presented alongside a question during the survey, asking how the first word variable was causally related to the last in the sequence (A-C/ A-D). The following example has been extracted directly from the survey and it illustrates a singular problem as well as the corresponding response alternatives presented: An increase in mindfulness causes an increase in masculinity. An increase in masculinity causes an increase in conformity. An increase in conformity causes a decrease in conscientiousness. What effect does an increase in mindfulness have on conscientiousness? (a) An increase in conscientiousness, (b) A decrease in conscientiousness, (c) No effect on conscientiousness. The dependent variable in this study was accuracy score, more precisely the number of responses that participants were able to infer correctly.

Table 1

The Conditions Outlined in the Materials Section Along with Examples

Condition	Example problem	Correct response
PP	"An increase in obedience causes an increase in assertiveness. An increase in assertiveness causes an increase in mindfulness.". Prompt: "What effect does an increase in obedience have on mindfulness?"	"An increase in mindfulness"
NN	"An increase in memory causes a decrease in calmness. A decrease in calmness causes an increase in extraversion.". Prompt: "What effect does an increase in memory have on extraversion?"	"An increase in extraversion"
PN	"An increase in momentum causes an increase in expansion. An increase in expansion causes a decrease in buoyancy.". Prompt: "What effect does an increase in momentum have on buoyancy?"	"A decrease in buoyancy"
NP	"An increase in limirance causes a decrease in berence. A decrease in berence causes a decrease in marance.". Prompt: "What does an increase in limirance cause?"	"A decrease in marance"
PPP	"An increase in height causes an increase in decay. An increase in decay causes an increase in tension. An increase in tension causes an increase in velocity.". Prompt: "What effect does an increase in height have on velocity?"	"An increase in velocity"
NNN	"An increase in control causes a decrease in femininity. A decrease in femininity causes an increase in performance. An increase in performance causes a decrease in stability.". Prompt: "What effect does an increase in control have on stability?"	"A decrease in stability"

Table 1 Continued

The Conditions Outlined in the Materials Section Along with Examples

PPN	"An increase in luferity causes an increase in idealority. An increase in idealority causes an increase in readicity. An increase in readicity causes a decrease in borence.". Prompt: "What effect does an increase in luferity have on borence?"	"A decrease in borence"
NNP	"An increase in intensity causes a decrease in height. A decrease in height causes an increase in acceleration. An increase in acceleration causes an increase in conservation.". Prompt: "What effect does an increase in intensity have on conservation?"	"An increase in conservation"

Procedure

The link to the survey was accessible to public via the social media platforms (Facebook and Reddit). Once accessed, participants were taken to the Qualtrics survey page where they were broadly provided information concerning experimental aim, contact information, as well as assurance of anonymity. After consenting to the terms, participants moved onto a more detailed instruction page where the survey was thoroughly explained. Here, participants were instructed to deduce how two phenomena are causally related based on the information provided. The following example was also given stating a real and factual scenario: "An increase in outdoor temperature causes a decrease in heating costs." And "A decrease in heating costs causes an increase in the amount of money saved.". This was followed up by a correct response as an example: "An increase in outdoor temperature causes an increase in the amount of money saved". Additionally, two graphs were displayed for familiarising with the experiment and as a prompt for the type of reasoning that was expected. First was a line graph that provided a visual representation of a positive linear relationship, whereas the second was a line graph that showed a negative linear relationship. Both had two descriptors, where the positive relationship suggested that if X increases, Y increases and if X decreases, Y decreases. Similarly, applicable descriptors were provided for the negative relationship, such as if X increases, Y decreases, and if X decreases, Y increases. This ensured that participants had a clear idea of what was expected of them.

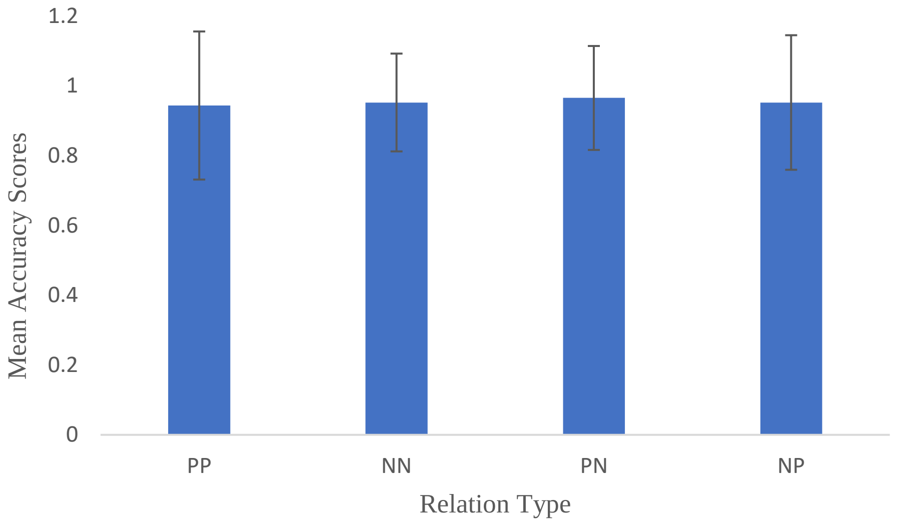
In the following pages 24 causal reasoning problems were presented in three blocks, where a single block consisted of eight different hypothetical relation problems from a single domain. Participants were instructed that this should take them approximately 25 minutes to complete. Format of the questions was uniform throughout so that “An increase in A causes an increase in B. An increase in B causes a decrease in C. What effect does an increase in A have on C?” was applicable in all cases, except from three relation problems where an additional clause was added. Additionally, the relation type was changed based on the conditions described in the Materials section. To answer these, three multiple choices were provided stating that an increase, decrease or no effect will occur. Only one response could be made per problem. Finally, basic demographic, education and primary occupation details were collected.

Results

The mean accuracy scores across two relation conditions can be seen in Figure 1. For the mean accuracy scores of three relation conditions please see Figure 2. Four, one-way repeated measures ANOVAs were conducted to analyse the data. The first ANOVA was conducted to assert whether there was a difference between relationship types in respect to accuracy as the dependent variable. The conditions were PP, NN, PN, and NP. From this analysis it could be inferred that there was no significant effect $F(3, 234) = 0.45$, $p = .719$. The second ANOVA was carried out for three-relation conditions (PPP, NNN, PPN, NNP) and it established that there was no significant effect of relation type under these circumstances either $F(3, 234) = 1.44$, $p = .231$. The third ANOVA compared whether there was a difference in accuracy between relation length for PP and PPP conditions, however no significant effect was found $F(1, 78) = 1.29$, $p = .259$. The last ANOVA was done to check if the relation length influenced accuracy in NN and NNN conditions, although once again no significant effect was found $F(1, 78) = 2.03$, $p = .159$.

Figure 1

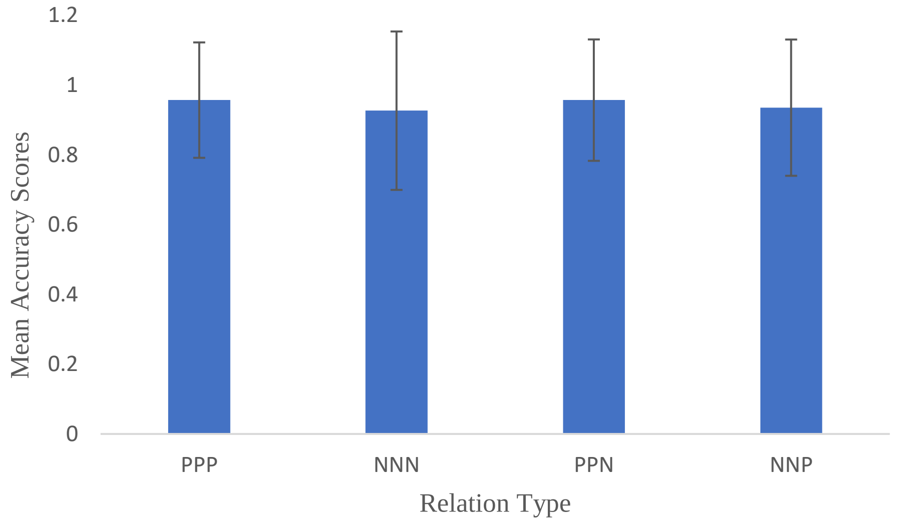
The Mean Accuracy Scores Across Two Relation Conditions



Note. This figure displays the mean accuracy scores and mean + 1 *SD* and mean – 1 *SD* for positive-positive, negative-negative, positive-negative, and negative-positive conditions.

Figure 2

The Mean Accuracy Scores Across Three Relation Conditions



Note. This figure displays the mean accuracy scores and mean + 1 *SD* and mean – 1 *SD* for positive-positive-positive, negative-negative-negative, positive-positive-negative, and negative-negative-positive conditions.

Discussion

The aim of the study was to find out how well individuals perform when solving hypothetical causally related problems, and further understand what intelligence factors interact with this. This was done via a survey that consisted of hypothetical, linear relation, causal scenarios. When relation types were tested to answer the first research question asking whether there will be a difference in accuracy between positive, negative, and mixed relation scenarios, I found evidence suggesting that irrespective of relation type (PP, NN, PN, NP) participants performer remarkably well. Three relation conditions (PPP, NNN, PPN, NNP) also showed no evidence of any major differences in accuracy scores when relation types were tested. Next, I wanted to gain an empirical understanding of relation length to accuracy scores and analyse whether people performed better or worse across conditions PP and PPP, as well as NN and NNN, however the results show that no such difference is evident. Therefore, three relation conditions did not lead to more errors in the responses compared to two relation conditions irrespective of relation type.

Based on the results provided it can be assumed that the difficulty of the causal problems was not sufficiently adjusted for the group that participated, hence the high accuracy scores. Simple linear relationships, in terms of increases and decreases may be too familiar to those with a background in further mathematics, or else some form of statistical reasoning. A small percentage of participants had a high school diploma or equivalent (18%), therefore a higher proportion of errors should have been expected at least on the lower spectrum of education within the given sample. These results are not consistent with Barbey and Wolff (2007) who claimed that effects of negation require a more elaborate understanding, thus creating a higher level of difficulty when solving. The conclusion that one could arrive at based on these results would be that causal reasoning specific to these types of problems is far more commonly understood than anticipated. However, there is a possibility that participants found an alternative method to conclude the answers throughout the survey as the causal problems were uniform in structure.

Formatting of the causal problems were kept same throughout the survey as this was thought to prevent any random effects from being introduced that may influence the results. Instead, this presented a potential heuristic which participants could apply. The term has been defined as a mental shortcut that enables efficient processing of information through a simplified strategic approach, which dismisses elements of information deemed unnecessary

(Gigerenzer & Gaissmaier, 2011). The last relation in all the causal problems could have been used to deduce the correct responses, and bypass most of the task, as it would become irrelevant information. Heuristic approaches are known to be an adaptive method to problem-solve, and in this case it would enable a participant to make the necessary decisions quicker than solving the problems in any other way. Thus, if such a heuristic was noticed and applied, the outcome would result in a high number of correct responses. Heuristics do however have certain flaws, where an overapplication could lead to errors, given that one of the conditions was formulated in a way that would mislead the participants from making the correct choice. An example condition NNP could be rearranged so that the positive relation appears first (PNN), in which case the last term would lead to a negative and incorrect conclusion.

Although, some form of reasoning must still have been applied to find such a heuristic as it is not clearly presented. Given that the heuristic was identified within the first few problems, this would enable a participant to ignore majority of the information in the survey and simply attend to the last relation term and get a perfect score. Additionally, eight problems were presented per block, meaning even if some of the responses were provided incorrectly there was an opportunity to amend the answers. This was not possible if a participant moved onto the next block, as no possibility to return to a previous block was provided. Depending on whether the heuristic approach has been used in this experiment, it may not be reasonable to claim that causal reasoning was accurately captured by this experiment. Due to validity concerns which will be further expanded on, I have conducted a second experiment.

Experiment 2

Method

Participants

A total of 20 participants took part in this study, 11 male, 6 female and 3 who identified as other. The age ranged from 18 to 55 (*M* = 31,37, *SD* = 11.20). Participants were informed to not take part in this experiment if they had already been recruited for experiment one, therefore I decided to go with the assumption that none of the participant in Experiment 2 participated in Experiment 1. Data from a single participant had to be excluded due to incomplete responses. The sampling method used to recruit participants remained the same as in the first experiment, through an opportunity sampling method via an anonymous link distributed on social media platforms (Facebook and Reddit). Once again, no common occupation categories could be identified from the responses provided. Similarly, there was a broad range of study backgrounds, with no predominant identifiable categories. In the current sample five of the participants had achieved a high school diploma or equivalent (25%), 11 had a bachelor's degree (55%), three had achieved a master's degree (15%), and only one participant had a doctoral degree (5%).

Materials

Much of this section remains the same for Experiment 2 as presented in Experiment 1, however specific changes will be described in detail. The condition from Experiment 1 with three relations (PPN) has been altered for Experiment 2 (PNN) for all three domains. This ensured that the heuristic approach could not be used by simply referring to the last relation term in the problem. A further change has been implemented that prevented the same issue from occurring, where in the following example: "An increase in X causes a decrease in Y. An increase in Y causes a decrease in Z", the relation term of Y did not remain consistent when defining the X to Y relation and the following Y to Z relation, which was not the case in Experiment 1. This was relevant for all of the conditions, however only PN and NN conditions were additionally affected in terms of preventing the heuristic from occurring. In total, nine out of the 24 problems did not lead to a correct solution based on the last relation term in the sequence. Each causal problem was no longer followed up by a question, instead prior to answering the causal problems an instruction has been given stating what the necessary steps are throughout the whole survey. Two answers per causal problem were applicable, where fewer or more responses resulted in an error message asking to adjust number of responses chosen. Multiple choice answers were still used, however the number of responses for this survey has been changed to four to reduce the likelihood of answering correctly by chance from 33% to 17%. Response descriptions were also different, where each described the relation between A and C or D, so that both combinations of an increase or decrease in A were covered for both an increase or decrease in C or D in their respective conditions. Additionally, response order was randomized to reduce answering correctly by chance. To illustrate a singular problem has been extracted from the original survey: An increase in mindfulness causes an increase in masculinity. An increase in masculinity causes an increase in conformity. An increase in conformity causes a decrease in conscientiousness. The response alternatives included the following options: (a) An increase in mindfulness causes an increase in conscientiousness, (b) A decrease in mindfulness causes a decrease in conscientiousness, (c) An increase in mindfulness causes a decrease in conscientiousness, (d) A decrease in mindfulness causes an increase in conscientiousness. In every problem two correct responses were available. For the example above, indicating a negative relationship, participants were required to select answers (c) and (d). In contrast, for a problem indicating a positive relationship, the correct answers would resemble the structure of answers (a) and (b). Both responses had to be chosen correctly to mark a problem as accurately answered. The dependent variable in this study was also accuracy score.

Procedure

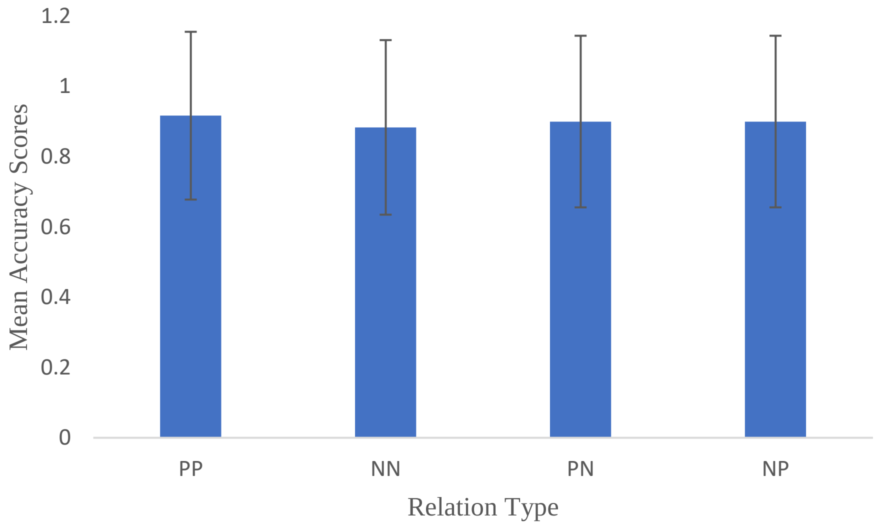
Access to the survey and a brief with necessary information along with assurance of anonymity were replicated from Experiment 1. The instruction page still provided all the information stated in Experiment 1; however, participants were informed that Experiment 2 would take them approximately 20 minutes to complete. Since two responses were required in this experiment, an example stating both relations to the factual problem "An increase in outdoor temperature causes a decrease in heating costs. A decrease in heating costs causes an increase in the amount of money saved." in Experiment 1 were provided, such that both "An increase in outdoor temperature causes an increase in the amount of money saved" and "A decrease in outdoor temperature causes a decrease in the amount of money saved" were correct and should be selected. Individual questions that followed each causal problem were removed, instead participants were instructed to identify two correct responses based on the information provided in every scenario.

Results

The mean accuracy scores across two relation conditions can be seen in Figure 3. For the mean accuracy scores of three relation conditions please see Figure 4. For Experiment 2, data analysis followed a similar pattern where four one-way repeated measures ANOVAs were conducted. I ran the first ANOVA to once again assert whether there is a difference between relation types and accuracy scores (the dependent variable) for the two relation conditions (PP, NN, PN, NP). The analysis resulted in a non-significant main effect $F(3, 57) = 0.28$, $p = .843$. Similarly, I tested whether there was a difference between relation types for three relation conditions (PPP, NNN, NPP, PNN) and accuracy scores. This ANOVA showed that there is no significant effect $F(3, 57) = 2.62$, $p = .060$. The remaining analyses were carried out to see if there was a difference between relation length and accuracy scores in positive and negative relation conditions. When conditions PP and PPP were tested no significant effect was found $F(1, 19) = 0.00$, $p = 1.000$[1]. Lastly, NN and NNN conditions were tested where the ANOVA showed a non-significant effect $F(1, 19) = 2.45$, $p = 0.135$.

Figure 3

The Mean Accuracy Scores Across Two Relation Conditions

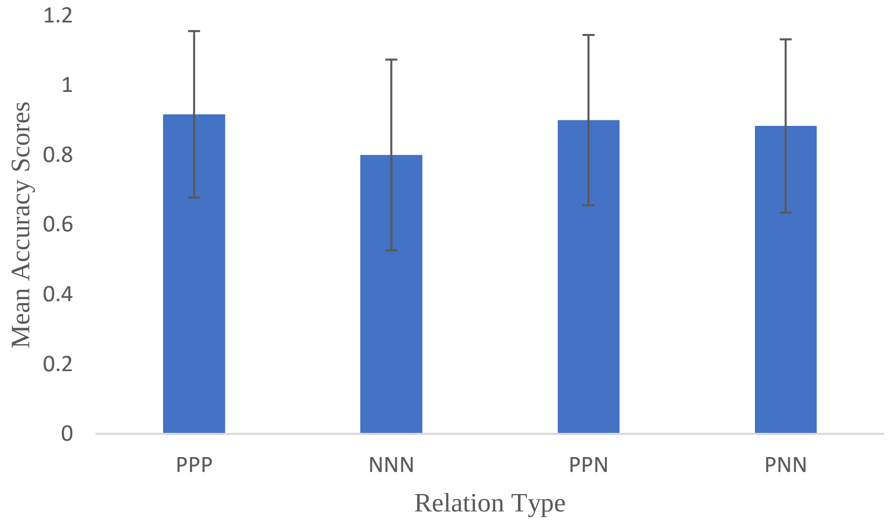


Note. This figure displays the mean accuracy scores and mean + 1 *SD* and mean – 1 *SD* for positive-positive, negative-negative, positive-negative, and negative-positive conditions.

[1] The *F* value of 0.00 occurred due to a small range of possible means in the DV (0, 0.33, 0.66, 1.00).

Figure 4

The Mean Accuracy Scores Across Three Relation Conditions



Note. This figure displays the mean accuracy scores and mean + 1 *SD* and mean – 1 *SD* for positive-positive-positive, negative-negative-negative, positive-positive-negative, and negative-negative-positive conditions.

Discussion

To answer the first research question I analysed the data which shows us that relation type does not influence accuracy scores when two relation conditions were tested. Therefore, negative, and mixed conditions were answered correctly with a very small difference in number of errors on average. In the case of three relation conditions, the differences between means were close to reaching levels where an effect would suggest a trend where relation type affects accuracy scores, however this is not definitive. Relation length analyses have shown that positive relation conditions are solved just as accurately whether two or three relations are present. On the other hand, data suggests that there is a visible decline in accuracy (see Figure 4) when three negative relation conditions are solved in comparison to two relations, however this was not significantly different between tested means.

General Discussion

Results from Experiment 1 have shown that relation type or length does not influence accuracy scores irrespective of the conditions tested, where performance was really high across all conditions. It is however important to note that these results were obtained from a sample where most of the participants had achieved at least a bachelor's degree (82%). Despite that, the remaining participants still performed near to perfect accuracy, meaning variability in levels of education do not necessarily reflect on performance in the current study. The findings from Experiment 2 have shown that participants did not perform significantly worse in negative or mixed relation conditions either. The statistical significance is more applicable to two relation conditions however further observations can be made. In the three relation conditions, results have a significance value above the 5% probability threshold, however a general trend may be inferred, where a larger sample size could potentially lead to the desired significance value, as the current study only consisted of 19 participants. Thus a significant main effect of relation type under those circumstances may be a reasonable assumption to make. Furthermore, a direct comparison was made between relation length in the following conditions, PP and PPP, as well as NN and NNN respectively. Evidence obtained has shown that no main effects of relation length were present in the data set.

Taken together my findings indicate that participants were able to comfortably reason across all the conditions without any additional difficulties posed by the negative relations or added length to the causal problems. Contrastingly, further observations may be made concerning the resulting significance values. The means reflect an observable difference in NNN condition, where the accuracy scores were comparably lower to other conditions (see Figures 3 and 4). Once again, majority of the participants were found to be highly educated where 75% of the sample had achieved at least a bachelor's degree, and the issue that the sample may have the necessary knowledge to solve the problems provided remains unchanged from Experiment 1. Similarly, it is important to note that primary work roles and previous education may play a significant part in acquisition of correct responses in the carried-out study. This is due to the background knowledge situated around mathematics and statistics which may support the role of crystallized intelligence, thus having the ability to infer correct responses through an application of a simple mathematical formula. Further implications involve fluid intelligence, which may have been applicable to those without previous knowledge in causal reasoning in terms of increases and decreases. Since there is a possibility that Experiment 1 was simplified through an application of a heuristic, which nevertheless involves logical reasoning, Experiment 2 may be a better measure of fluid intelligence.

Theoretical and Practical Implications

Despite some limitations that will later be outlined, the results suggest several theoretical and practical implications. Causal inference is an everyday task that is a major factor of planning and decision-making. Before the day begins, we can decide for example when to start preparing for work, and how we would like to approach everyday tasks such as ones concerning preparations before a workday. The structure of reasoning may change in a dynamic setting, however reasoning in terms of increases and decreases is not unusual. Such that increasing the length of time dedicated to preparing may in a sense decrease the stress levels, and similarly decreasing that time would result in an increase in stress levels. Thus, the argument that people reason by weighing the consequences and outcomes of different scenarios appears to be moderately reasonable as well as pragmatic. Establishing further discourse to gain a better insight into these problems can help us understand how we can converge theory into practice.

Intelligence is influenced by numerous factors and the abilities surrounding it are just as relevant when reflecting on a particular task. How different individuals reason about the causal problems in this study may be guided by their previous knowledge such as spoken language, ingrained cultural norms, and skills which are acquired throughout the lifetime (Buades-Sitjar & Duñabeitia, 2022). Thus, how participants reason through the causal problems in the current study is based purely on individual approaches, which are once again influenced by their crystallized intelligence. Although, it is more likely that the problems tackle some type of fluid intelligence, given that they consist of hypothetical scenarios. A claim that the inference task in the current study measures purely fluid intelligence may be too specific, also theoretically unlikely. Both intelligence factors are assumed to be overlapping (Kyllonen & Kell, 2017), therefore an activation in one would naturally lead to the other. Since participants had to perform some form of abstract reasoning to make sense of the hypothetical scenarios and the made-up terms, it is difficult to argue that participants could simply perform such transformations based on previous knowledge. The results however show that in both experimental settings participants performed above expected norm. These findings suggest that participants are well accustomed to the task demands, and as proposed by Sloman (2005), this is due to causal models that drive human thought. Hence, the ability to solve these problems may be more innate than anticipated. The rationale is that understanding causality helps us function and reason through numerous daily events without constantly making errors or mistakes in how we attribute causality. Although, this attribution should not come as naturally in the context of the current study.

On the other hand, since this task does capture problem-solving and decision-making, I would assume that some form of inference requiring fluid ability can be deduced. To solve the given problems, a simple transformation from made up words such as luferity, idealority, readicity, and borence could be made, where each word is converted to a mental representation A, B, C, and D. Based on the definition by Anshakov and Gergely (2010) on fluid intelligence, this is the processing that would likely be applied when reasoning through such hypothetical scenarios. It is difficult to say whether participants have previous knowledge that would allow them to make such a transformation due to its specific role in the context of hypothetical causal reasoning. Evans (2002) argues that content and context-based differences are often a major influence of many commonly identified logical errors. Given more significant results it may be possible to say whether the errors that have occurred in certain conditions are based on novelty effects, or a lack of inference capabilities and successful application of intelligence factors. Instead, the results now suggest that content and context-based differences extend the issue of causal reasoning, as they can be seen as additional influencing factors. This study has shown that participants can perform exceptionally well in simple hypothetical situations. In a scenario that relies more on contextual information, which Evans (2002) refers to, participants are possibly more likely to produce errors.

Another aspect may be involved in the outcome of the current study. Since working memory and intelligence are argued to be one and the same factor (Ackerman et al., 2005), it could be said that simply having a vast knowledge-base and strong recollection skills could result in high accuracy in responses. Furthermore, if this knowledge can be easily retrieved (Sloman, 2005) it can be used to solve given problems purely through recollected knowledge implying crystallized intelligence as being the primary source of reasoning. More specifically, working memory and fluid intelligence have also shown to be highly correlated aspects of cognitive function. Furthermore, working memory may even be the leading factor of fluid intelligence. Fry and Hale (1996) state that by simply aging into adulthood, which implies natural development of processing speed as well as functioning of working memory, the ability to reason fluidly improves. This claim supports the high accuracy scores of the tested sample, as every participant was an adult with a reasonable level of education. These may be the factors necessary to answer all or majority of the given causal problems successfully. How intelligence interacts with the directly presented stimuli, however, remains unclear.

Fluid intelligence may be distinguishable from its crystallized aspect in application; however it is not viewed as an independent factor. Nisbett et al. (2012) suggests that an interaction is always in effect, thus, to say a task measures fluid intelligence may need to be replaced with a question asking to what extent does this task measure reasoning abilities outside of previously acquired knowledge. We can infer that unseen stimuli and novel problems refer to some form of fluid reasoning, however in this case causality is a very natural everyday function. Consequently, this creates difficulties when formulating problems for the current study, situated around increase and decrease verbs, that could accurately measure causal reasoning and fluid intelligence together. Since intelligence factors are arguably intertwined in many ways, fluid intelligence may simply be a tool that extends from crystallized abilities allowing us to acquire further knowledge. This explanation works alongside the suggested relationship by Nisbett et al. (2012) who state that for crystallized intelligence to increase fluid ability is essential. Furthermore, a decrease in fluid intelligence would not impact the crystallized factor of intelligence, thus the previous knowledge of an individual could continue to increase. In general, specific aspects of intelligence can be referred to broadly, such that crystallized ability has a distinctly different function in contrast to fluid, for grouping of the respective abilities. Classifying this type of terminology, however based on task specifics may not be the best way to discuss the underlying inferential abilities of intelligence. It is difficult to be certain of the factors that may be contributing the most, however an interactive influence from both fluid and crystallized intelligence is likely. Possible limitations of the current study, as well as recommendations for future studies to further develop the theoretical framework will be discussed in the following sections.

Limitations

The ability that allows us to filter out irrelevant information and prevent incorrect responses from emerging is sustained by working memory (Nisbett et al., 2012). An application of this ability in a uniformly designed survey could lead to attentive fixation over time leading to specific pattern recognition. This refers to the idea that only the last relation word can be read to come to the correct conclusion, without concerning with the rest of the question. Thus, the idea that only the last relation term has been attended to throughout the entire survey could be a likely explanation as to why most of the respondents got near to perfect accuracy score on the causal problems in Experiment 1. This becomes an even more plausible explanation alongside the heuristic definition provided by Gigerenzer and Gaissmaier (2011) which suggests that mental shortcuts occur when information is processed efficiently, and irrelevant information is dismissed. Despite this information not being overtly present, some form of systematic reasoning could have taken place, although with a deviation from the original expectation of how the causal problems should have been solved. The heuristic method still implies some form of causal inference; however it is not reasonable to assume that participants simply observed the last causal term in the scenario from the beginning of the survey and were able to commit to this method without ensuring that the application is correct for all available conditions. Martín and Valiña (2023) investigated the issue of heuristics in decision making and argue that mental shortcuts are a frequent occurrence when in need to make some kind of inference to a given problem. They further discuss that these strategies do not imply any form of logic yet could lead participants to systematically deviate from the normative, in this case expected, decision making. It is therefore not unreasonable to assume that this is the process of how participants came to the right conclusions so frequently. To reiterate, the negative conditions in the current study were supposedly more challenging, thus should have resulted in more errors, however this was not the case. It is unclear whether such a limitation is also applicable to Experiment 2 due to the lack of control over numerous extraneous variables, and causal problems which could be solved in numerous different ways, in terms of what deduction steps took place to infer the answers. Despite the prompt graphs displaying positive and negative relationships in the instructions it does necessarily mean that participants used these to understand the task, as the inference is open to interpretation based on individual skills in abstract reasoning. Furthermore, the example provided in the instructions is based on a realistic causal relationship which does not hold the same relational ambiguity as the related hypothetical terms in the survey.

The major downside to the design of Experiment 1 comes from its ambiguous construct validity, as the intended measure may have been confounded by later identified design flaws. As previously discussed, the specified heuristic may have been applied, consequently altering the way that the originally tested causal inference should be interpreted. Essentially, this does not mean that causal reasoning was not captured by the experiment, however in terms of how this heuristic may have affected accuracy results is unclear, as it was not accounted for when the survey was designed. Due to these concerns Experiment 2 was introduced to establish whether the findings may have been misrepresented. Since an alternative method to solve every causal problem in the survey was available, the second survey was adjusted to ensure that the last relation term in the causal problems cannot always be applied. More specifically, this was the case for nine out of the 24 causal problems (38%). This alteration means that the heuristic may still be applied, though we can imply that causal reasoning must still occur to achieve correct conclusions. However, given that this is applied without further discretion of what the problems state, numerous errors would occur due to not attending to all the causal terms provided. Since this was not the case, I assume that the small sample size in Experiment 2 could have prevented that from becoming evident as not enough variability was captured from the 19 participants.

It is possible that the experimental paradigm has been either too simplistic, or people are generally far more capable at solving linear relationship problems than anticipated, whether those relationships are positive, negative, or mixed. Also, the sample may not be a good representation of how most individuals normally reason. This is due to a high proportion of the sample having at least a bachelor’s degree or higher. Implications of this would suggest that the results which are far more accurate than expected, are not due to a flawed paradigm, but rather a specific sample that would perform better than average. Furthermore, background research used to support the current study is topic relevant in terms of explaining the applications of causal reasoning, however the specific paradigm from the current study has not been previously replicated. As such, designing a logical survey in relation to previously established paradigms may provide more continuity for subsequent studies. This is a concern due to the potential application of the relational heuristic present throughout the survey of Experiment 1.

Other limitations to the current study include the environment that the study may have been completed in. It is unclear whether all participants had similar conditions when taking part in the study. Despite lack of significant results, it may be an influencing factor when considering intelligence, as it is reliant on working memory and attention. Further limiting factors may be attributed to cultural differences as it is unclear how intelligence is affected by cultural norms. The way fluid and crystallized intelligence are presented may be a variable component to reasoning, where our everyday lives influence how we perceive causal interactions.

Future recommendations

Application of crystallized and fluid intelligence in everyday causal reasoning tasks is a natural occurrence. A lack of time constraints may be an issue when considering logical reasoning problems. This is because participants can dwell on any given problem for an indefinite amount of time until arriving at a correct solution. With a prespecified time limit, the heuristic may not have been identified as easily, or at all, as the added temporal factor would encourage participants to progress through the questions at a specified rate. Furthermore, it is important to consider language used throughout the survey, as it may be more engaging to formulate problems that have an alternative structure, due to the design being uniform throughout. This would likely introduce between-subjects effects, however further comparisons are necessary to gain a more thorough understanding of implications underlying the verbs increase and decrease. Similarly, this may lead to a reduction in opportunities to identify a heuristic related to the current causal problems.

To correctly derive an A-C or A-D interaction among other intermediate elements in a causal problem, some form of reasoning must take place, however there are multiple ways to achieve this. In addition, the method of completion is even less evident when too many naturally occurring variables are not accounted for, thus bringing in ambiguity the experiment results. Therefore, given specific control conditions such as the environment that the task has been completed in or the time given to complete it, we can more precisely measure the limitations of reasoning. This would allow us to account for the different methods of problem-solving in the current study task. Furthermore, Fry and Hale (1996) state that the process of aging into adulthood improves a major influence on fluid intelligence, which was found to be processing speed and an improved functioning ability of working memory. Future studies should therefore focus on either younger or an elderly population, which would allow clearer observations of what changes occur across different age groups. Additionally, a larger sample size would contribute to power of Experiment 2, and with additional data significance values may be pointing more towards a main effect. This may be especially relevant to the three relation conditions.

Sloman et al. (2009) have researched the implications of verbs *cause*, *allow*, and *prevent*, which similarly were identified to provide a more thorough understanding of how causal language is comprehended in various conversations or scenarios. Their study also provides a strong basis for further empirical development and different interpretations of these causal verbs. If we can attribute specific causal meaning of a larger variety of verbs such as influence, affect, or create, that imply a broad causal interaction, similarly, to cause, increase or decrease, it would benefit our understanding of how reasoning is influenced by these terms. The implications of this would help us develop our knowledge of the underlying factors of intelligence in application of language.

Lastly, as outlined in the limitations controlling the environment that the survey is completed in may result in more generalisable data, as well as provide easier replicability of potentially significant results. Additionally, cultural norms may be a needed aspect to consider when designing future studies as it could show us how intelligence factor correspond to the positive, negative, and mixed relational conditions tested.

Conclusion

This study is relevant in the modern research of causal reasoning and two major factors of intelligence. It provides empirical evidence for how well individuals are able to infer correct responses across various hypothetical causally related conditions. The key words for causation were increase and decrease which dictated the direction of the inferred relationship. The related words on the other hand were chosen from three distinct domains (natural science, psychology, and made-up) to account for the hypothetical events. The results showed that participants were able to reason across all conditions surprisingly well, however, Experiment 2 has shown more variability across means where the NNN condition was responded to incorrectly the most, thus proved to be the most challenging. Due to a possible heuristic in Experiment 1, a follow up experiment was necessary. This work demonstrates that it is not clear what type of intelligence causal inference processes entail, however it is likely that the combination of both take place when reasoning. Future studies should focus on validity of intelligence testing when implementing similar paradigms. Additionally, including timed and controlled experimental conditions may be able to create more representative results.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working Memory and Intelligence: The Same or Different Constructs? Psychological Bulletin, 131(1), 30–60. https://doi.org/10.1037/0033-2909.131.1.30

Anshakov, O. M., & Gergely, T. (2009). Cognitive Reasoning Framework. *Cognitive Technologies*, 51–68. https://doi.org/10.1007/978-3-540-68875-4_5

Anshakov, O. M., & Gergely, T. (2009). Cognitive Reasoning Framework. *Cognitive Technologies*, 51–68. https://doi.org/10.1007/978-3-540-68875-4_5

Barbey, A. K., & Wolff, P. (2007). Learning causal structure from reasoning. *In Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 29, No. 29).

Buades-Sitjar, F., & Duñabeitia, J. A. (2022). Intelligence subcomponents and their relationship to general knowledge. *Journal of Cultural Cognitive Science, 6(*3), 343–355. https://doi.org/10.1007/s41809-022-00113-z

Burgess, G. C., Braver, T. S., & Gray, J. R. (2006). Exactly how are fluid intelligence, working memory, and executive function related? Cognitive neuroscience approaches to investigating the mechanisms of fluid cognition. *Behavioral and Brain Sciences, 29*(2), 128-129. https://doi.org/10.1017/s0140525x0625903x

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why?. *Current directions in psychological science*, *19*(1), 51-57. https://doi.org/10.1177/0963721409359277

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature reviews neuroscience*, *11*(3), 201-211. https://doi.org/10.1038/nrn2793

Evans, J. S. B. (2002). Logic and human reasoning: an assessment of the deduction paradigm. *Psychological bulletin*, *128*(6), 978. https://doi.org/10.1037/0033-2909.128.6.978

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological science*, *7*(4), 237-241. https://doi.org/10.1111/j.1467-9280.1996.tb00366.x

Geake, J. G., & Hansen, P. C. (2010). Functional neural correlates of fluid and crystallized analogizing. *NeuroImage*, *49*(4), 3489-3497. https://doi.org/10.1016/j.neuroimage.2009.09.008

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, *62*, 451-482. https://doi.org/10.1146/annurev-psych-120709-145346

Hulme, C., & Melby-Lervåg, M. (2012). Current evidence does not support the claims made for CogMed working memory training. *Journal of Applied Research in Memory and Cognition*, *1*(3), 197–200. https://doi.org/10.1016/j.jarmac.2012.06.006

Kyllonen, P., & Kell, H. (2017). What is fluid intelligence? Can it be improved?. *Cognitive abilities and educational outcomes: A Festschrift in honour of Jan-Eric Gustafsson*, 15-37. https://doi.org/10.1007/978-3-319-43473-5_2

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature neuroscience*, *17*(3), 347-356. 347–356. https://doi.org/10.1038/nn.3655

Martín, M., & Valiña, M. D. (2023). Heuristics, Biases and the Psychology of Reasoning: State of the Art. *Psychology*, *14*(2), 264-294. https://doi.org/10.4236/psych.2023.142016

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81. https://doi.org/10.1037/h0043158

Mogle, J. A., Lovett, B. J., Stawski, R. S., & Sliwinski, M. J. (2008). What's so special about working memory? An examination of the relationships among working memory, secondary memory, and fluid intelligence. *Psychological science*, *19*(11), 1071-1077. https://doi.org/10.1111/j.1467-9280.2008.02202.x

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: new findings and theoretical developments. *American psychologist*, *67*(2), 130. https://doi.org/10.1037/a0026699

Pearl, J., & Mackenzie, D. (2018). *The book of why: the new science of cause and effect*. Basic books.

Roth, G. (2015). Convergent evolution of complex brains and high intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1684), 20150049. https://doi.org/10.1098/rstb.2015.0049

Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence?. *Psychonomic bulletin & review*, *15*, 364-371. https://doi.org/10.3758/pbr.15.2.364

Schipolowski, S., Wilhelm, O., & Schroeders, U. (2014). On the nature of crystallized intelligence: The relationship between verbal ability and factual knowledge. *Intelligence*, *46*, 156-168. https://doi.org/10.1016/j.intell.2014.05.014

Sligh, A. C., Conners, F. A., & Roskos-Ewoldsen, B. E. V. E. R. L. Y. (2005). Relation of creativity to fluid and crystallized intelligence. *The Journal of Creative Behavior*, *39*(2), 123-136. https://doi.org/10.1002/j.2162-6057.2005.tb01254.x

Sloman, S. (2005). *Causal models: How people think about the world and its alternatives*. Oxford University Press.

Sloman, S. A., & Hagmayer, Y. (2006). The causal psycho-logic of choice. *Trends in cognitive sciences*, *10*(9), 407-412. https://doi.org/10.1016/j.tics.2006.07.001

Sloman, S., Barbey, A. K., & Hotaling, J. M. (2009). A causal model theory of the meaning of cause, enable, and prevent. *Cognitive Science*, *33*(1), 21-50. https://doi.org/10.1111/j.1551-6709.2008.01002.x

Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*, *53*, 92-101. https://doi.org/10.1016/j.intell.2015.09.005

Sternberg, R. J. (Ed.). (2020). *The cambridge handbook of intelligence*. Cambridge University Press.

Xu, H., Xu, C., Yang, Z., Bai, G., & Yin, B. (2023). Two sides of the same coin: distinct neuroanatomical patterns predict crystallized and fluid intelligence in adults. *Frontiers in Neuroscience*, *17*, 1199106. https://doi.org/10.3389/fnins.2023.1199106