How Do People Come Up with Right Answers?

Analysis of Participants' Reasoning Patterns in Two Televison Game Shows

Ewelina Gajewska

January 13, 2022

Index number:
[redacted]
ewegaj@st.amu.edu.pl

Abstract

Most of the people have watched at least once one of the television game shows. Some of them maybe even imagined themselves taking part in it or played along with participants in front of their TVs. It is a piece of cake and fun when you know the answer but what do you do when you are unsure about your decision? One can simply guess or, contrary, employ her reasoning skills in order to infer the correct answer based on her general knowledge and available information at the time. As there are many types of reasoning, there are many ways of reaching a conclusion as well. In present paper I analyze those patterns of reasoning of contestants in two popular TV game shows — *Who Wants to Be a Millionaire?* and *The Money Drop*. The idea is not new, as many authors studied people's decision-making under uncertainty in experiments based on another popular TV show — *Let's Make a Deal*. With the use of real case scenarios, I answer the question, which types of reasoning are employed by players and whether it is influenced by the degree of question difficulty.

Keywords: decision making, heuristics, rationality, reasoning

1 Introduction

Aristotle stated that human beings are rational animals, making *rationality* a key characteristic of human nature (Ross, 1956). However, the concept of rationality is not that straightforward and can be divided into the theoretical and the practical. The former asks what is rational to believe while the latter is concerned with what is rational to do. Thus, the former definition is more suitable for the subject of the present work. Then, if one believes that p and that if p then q, it is rational for one to believe that q — i.e. believing that these premises are true, it is rational to believe that the conclusion is true as well. This is an example of the rationality of the transitions between the beliefs (Joyce et al., 2004).

The concept of rationality was central to many research agendas about decision-making processes. However, in the last two decades of twentieth century some researchers shed a new light on human rationality — they discovered that many human decision deviates "from the performance deemed normative according to various models of decision making and rational judgment (e.g., the basic axioms of utility theory)" (Stanovich and West, 2000, p. 645). Some even started to view humans as irrational beings — researchers associated with the heuristics and biases tradition (Joyce et al., 2004). However, as Stanovich and West (2000) point out this phenomena can be explained by the following four factors: 1) performance errors, 2) computational limitations, 3) application of wrong norms by a researcher and 4) a different understanding of the task by the subject. First of them refers to temporary failures of some cognitive processes such as shifts of attention and memory deactivation. Second reflects inherent to human cognition information processing limitations such as restricted attention span and working memory capacity. Third one assumes that the researcher applied the incorrect model to interpret human behavior in a studied case. And the fourth assumes that subjects might interpret the problem differently and act rational according to this problem statement (different than the one framed by the experimenter).

Closely related to rationality is the concept of reasoning, i.e., "a transition in thought, where some beliefs (or thoughts) provide the ground or reason for coming to another" (Adler and Rips, 2008, p. 1). Here, similarly as in the case of a definition of rationality, we can distinguish theoretical and practical reasoning — the former states that "there are propositions to serve as objects of belief, and these propositions can be reasons for further beliefs — beliefs that can be acquired by reasoning" (Joyce et al., 2004, p. 6) and the latter being concerned with actions (outcomes). Again, the concept of theoretical reasoning is central to the subject of this work, nonetheless making decisions, for example choosing option a and not b or c, comprises to some degree both types of reasoning.

Worth mentioning is also the fact that much of everyday reasoning is carried out implicitly (inferences are drawn unconsciously), nonetheless one can reason explicitly as well. The latter type is central to the subject of the present work as participants in chosen television game shows are encouraged to verbally report their reasoning processes.

"A deductive inference is the one which draws out a conclusion which is latent but implicit in the information given (...) In contrast, an inductive inference is one which adds information and is therefore not a logically necessary conclusion" (Evans et al., 2019, p. 2). Thus, in deduction conclusions necessarily follow from their premises. On the other hand, when inferences are not sure — given true premises — but are only likely to be true (can be assigned with some likelihood) they are defined as inductive inferences. Thus, the main attribute of deductive reasoning (deductive arguments) is validity, that is high logical strength of inferences given true premises. On the other hand, the concept of inductive strength is associated with inductive reasoning (arguments) — it reflects their logical

strength. What makes reasoning (arguments) inductively strong is the fact that when all premises are true, there is high probability that conclusions are true as well (in other words, it is highly unlikely that conclusions are false in this case). This strength increases also when one cannot think of any example that stands in contrast to the evidence provided in the premise. Also, the more examples with relevant conditions one can provide, the stronger the argument becomes. Therefore, inductive strength is a matter of degree, ranging from weak to strong arguments (Johnson-Laird, 2013; IEP Staff, 1995). Important to note is also the fact that inductive reasoning — in contrast to deduction — yields a conclusion with an increase of semantic information. That is because one makes inference from facts that she/he has observed (information one possesses) to those (new ones) that she/he has not observed.

Which type of reasoning is employed in a particular case is known only by the arguer. So, if one (from an "outside perspective") wants to assess a given argument, then she/he needs to check both scenarios — whether it is deductively valid or inductively strong (IEP Staff, 1995).

As stated in (Johnson-Laird, 2013) deductive reasoning comprises three stages: 1) understanding information presented in a task — generating mental representation of the situation based on given premises, 2) generating a putative conclusion, and 3) evaluating the conclusion's validity. Induction is composed of similar stages, as firstly one needs to grasp available information, secondly generate a hypothesis based on this information and general knowledge one possesses and finally evaluate reached conclusions.

When measuring deductive reasoning competence subjects are usually faced with different types of syllogistic inference (*modus ponens*, *modus tollens*, negation of the antecedent, affirmation of the consequent) — given premise P and conclusion C the task is to evaluate the conclusion. Two of those scenarios — negation of the antecedent and affirmation of the consequent — are common logical fallacies that participants tend to make (Spiel and Glück, 2008). Three main findings follow from research on deductive reasoning: 1) most people are good at solving logical problems, 2) nonetheless systematic errors and biases are part and parcel of human reasoning, 3) responses to reasoning tasks depend on the content of problems stated in those tasks. Additionally, it is interesting to note that although people understand the rule of *modus tollens*, they often fail to apply it — what is in contrast to the rule of *modus ponens* which people use correctly at much higher rates (Evans et al., 2019; Oaksford and Hahn, 2007).

Induction can take a few forms — there are three main types of this form of reasoning:

1. causal reasoning
2. reasoning from analogy (also called inductive analogy)
3. enumerative generalizations

Crucial to note here are also two other types of inductive reasoning, i.e. John Stuart Mill's methods (canons) of induction (which follow Francis Bacon's views on inductive reasoning) and the statistical syllogism (Overholser, 1993; Tsang and Williams, 2012; Osbeck, 2014).

If one wants to have valid and/or strong argument, it would be better to make use of deduction. However, when a decision context becomes complicated rationality is bounded and therefore using deductive reasoning in these cases may not be the best choice. Instead, making use of induction would be more helpful (Arthur, 1994).

With the dawn of behavioural economics field in the late 1970s some assumptions about human rationality were changed considerably. From now decision-making processes were viewed as relying

heavily on so-called "mental shortcuts" or heuristics. For example, availability heuristic describes our tendency to use information that comes to mind quickly and easily when making decisions. Thus, this heuristic can considerably bias human reasoning being a critical factor in the construction of a mental model of a given situation. Therefore, this approach is based on the concept on bounded rationality and takes into account many cognitive limitations that human minds possess.

Closely related to the subject of my work is The Monty Hall problem, analyzed by the research community with the use of examples from another popular television game show — *Let's Make a Deal*. Here, researchers focus on investigating fundamentals of reasoning about uncertainty. Here, a decision problem is stated as follows:

> You are given the choice of three doors: behind one door there is a grand prize and behind the others, there are small gifts. Firstly, you pick only one out of three doors, for example number 1, and the host, who knows what is behind every door, opens another door, for example number 3, which has some small gift. Then, he asks you whether you want to pick door number 2, or stay with number 1. Now the question is the following: is it to your advantage to switch your choice?

Most claim that the best answer in this scenario is to switch, however most people prefer to hold to the option they originally chose. People's decision making processes in this problem were analyzed by many scientists from various fields — mathematicians, psychologists as well as logicians (Krauss and Wang, 2003; Halpern, 2017; Manfredotti and Viappiani, 2021).

Lastly, interesting to note is also the fact that there have been attempts at making a computer (machine) play one of those game — *Who Wants to Be a Millionaire?* — with quite successful results so far (Lam et al., 2012).

2 Methodology and Materials

I base the present study on the verbal protocol analysis — a method that analyzes people's verbal reports while they attempt to solve the decision problems. In both cases I gathered data from 3 different episodes. And that is because most participants answer only a few questions (win relatively small prizes) and usually the ones they are (almost) certain about. Therefore, I skipped those questions in which participants are sure they know the answer and considered the rest, however only the ones in which contestants provide some verbal reports — in order to be able to reconstruct their reasoning (thinking) patterns.

2.1 Who Wants to Be a Millionaire?

This television game show was first released in 1998 in the United Kingdom and today there are many international versions of the show. In the British version — which was used in the present study — money prizes range from £100 to £1 million — and of course £0 if a person gives wrong answer to one of the first two questions. There are 15 questions which difficulty increases at each step.

In this game show the player has three so-called lifelines each of which he/she may use once to help him/herself in answering a question. These are the following: "50:50" — reducing the number of

possible answers from four to two, with remaining one right answer and one wrong answer, "phone-a-friend" — the opportunity of consulting a friend in answering a question, and "ask the audience" — each person in the audience votes for one answer he/she believes is the correct one and summarized results in a form of percentages of votes for each answer is displayed for a participant.

In addition, there are two "safety nets", i.e., certain thresholds (guaranteed levels), which once reached provide a guaranteed minimum of prize money. It means, that if one gives a wrong answer to a question, but earlier had reached one of safety nets, he/she will not leave the game empty-handed. There are two guaranteed thresholds: £1,000 and £32,000.

In the case of *Who Wants to Be a Millionaire?* I additionally disregarded examples when players used a help in a form of one of the following two lifelines: "phone-a-friend" and "ask the audience". When it comes to the third lifeline ("50:50") participants still have to make a decision (have to reason) by themselves, so in my analysis I included those examples when participants chose this option of help, but not the other two mentioned earlier.

2.1.1 Examples

1. Link to the episode on YouTube platform: `https://youtu.be/bB6LAcVcdtQ`
   Level of difficulty: £ 125k
   Question: *What does an ichthyologist study?*

   At the beginning, the participant said his fire intuition was that the correct answer is *d. reptiles*. However, after a while he took the "50/50" lifeline, and as a result was left with two possible answers: *c. fish* and *d. reptiles*. Then, he reports that something is telling him that the correct answer is *c. fish*. He recalls a situation when he read a book and there was a picture of a fossil of a fish and its description contained something like "ichth" prefix. Participant's verbal reports let me classify this line of reasoning as an example of reasoning by analogy. It may be reconstructed as follows: the participant thinks that once he saw a fossil of a fish being called something like "ichth"; now he has a question that asks about ichthyologist; it contains an "ichth" prefix, so (by analogy) it may be also linked with the topic of fish; therefore, the correct answer is that ichthyologist deals with fish.

2. Link to the episode on YouTube platform: `https://youtu.be/_b0M5_gyiuc`
   Level of difficulty: £ 4k
   Question: *In 2001, Ben was born to which Coronation Street couple?*

   At the beginning the participant eliminated the answer *d.* saying that they are an older couple. Here, he relies on inductive statistical reasoning (also called statistical syllogism) as usually younger couples (not older) give birth to children, so answer *d.* is less likely to be true than the other ones. Then he takes "50/50" and is left with two possible answers *b. Emma and Norman* and *c. Janice and Les*. Now he hesitates for a second and says he probably has to guess which one is the right answer. But then again recalls the age of couples mentioned in answers *b.* and *c.* and says that "Les Battersby is quite old [so the option c. is not correct]". Here, his reasoning goes as follows (matching the logical rule of *modus tollendo ponens*):
   scheme: $((p \vee q) \wedge \neg p) \rightarrow q$

It is b. or c.
It is not c.
Therefore, it is b.

3. Link to the episode on YouTube platform: the same as above (in number 2).
   Level of difficulty: £ 16k
   Question: *Albert Park is the setting for which Formula One Grand Prix?*

   At the beginning the participant disregards two answers straight away. Therefore, he is left with the other two — *b. Australian* and *d. Canadian*. Thus, the correct option is *b.* or *d.*. Then, the participant says: "My wife actually lived for a year in Adelaide [in Australia] and I don't remember her talking about Albert Park. That gets me back to [answer d] Canadian". Now his reasoning matches the logical rule of *modus tollens* and can be reconstructed as follows:
   scheme: $((p \rightarrow q) \wedge \neg q) \rightarrow \neg p$
   If Albert Park was in Australia, my wife would mention it.
   My wife did not mention it.
   Therefore, Albert Park is not in Australia.

   Then, his reasoning follows the rule of *modus tollendo ponens*:
   scheme: $((p \vee q) \wedge \neg p) \rightarrow q$
   It is b. or d.
   It is not b.
   Therefore, it is d.

4. Link to the episode on YouTube platform: the same as above (in number 2).
   Level of difficulty: £ 8k
   Question: *Which flower takes its name from a Latin word for wolf?*

   Here, the participant knows that "wolf" in Latin is called "lupus", and among the answers there is similar word — "lupin", so relying on inductive reasoning from analogy he chooses the option *c. lupin.*

5. Link to the episode on YouTube platform: `https://www.youtube.com/watch?v=riP84I_CkaA`
   Level of difficulty: £ 16k
   Question: *Which is the coldest planet in our solar system?*

   Some form of inductive reasoning is involved in this example and most probably this is reasoning from analogy. That is because the participant seems to know something about the farthest planet from the Sun (i.e., about Pluto) and now he is trying to figure out what are the properties of the coldest planet in our solar system and whether (and how) they are related. Thus, he compares attributes of "the farthest planet" and "the coldest planet" and concludes that these two have some things in common. And that is enough for him to assume that they refer to the same thing — the same planet Pluto. This makes Pluto the coldest in our solar system and therefore the correct answer to the question.

2.2 The Money Drop

The Money Drop was first introduced in the United Kingdom in 2010. Since then over 15 international versions of the show were produced, including the Polish one known as *Postaw na milion* hosted by Łukasz Nowicki and produced by TVP2 television station. Episodes from this Polish version of the show were used in the present study.

In each game a team of two friends or relatives have to answer a total of 8 questions in order to win some amount of money. Participants start with 1 million PLN in a form of 40 bundles comprised of 25k PLN each. In this show the lowest prize is equal to 25k PLN and the highest is 1 million PLN. However, at each step the game can end, if players bet all their money on the wrong answer. So, in order to win any amount of money they must answer all questions.

The first four questions have 4 possible answers. The next three questions have three options and the final eighth question has only two possible answers. Additionally, each question comes from a specific category, which can be chosen (to some degree) by participants — participants are presented with the pair of possible categories from which they choose one of them.

In this show each answer corresponds to one trapdoor. In each question participants allocate money on all but one trapdoor — i.e., one of them must remain empty. Additionally, there is a fixed time range — one minute — to make decision and allocate the money at each step. Trapdoors corresponding to wrong options collapse at each round (after participants made decisions and time to answer the question has expired) and participants lose money they bet on these incorrect answers. Therefore, at each round players are left with the amount of money they had put only on the correct option. Contestants win the money they carried to the final question — but only if they choose the correct answer in this final question. Otherwise they leave the show empty-handed.

Here, question difficulty increases with a decrease in a number of possible answers. That is because in the first four questions participants can allocate money on three answers, in the next stage they can bet the money on two answers and in the final question they can choose only one option.

Last note is that because questions were taken from the Polish version of the show I did not translated their content to English so as to best present to the reader the meaning of questions (as some degree of information is often lost with translation and I wanted to avoid that).

2.2.1 Examples

1. Link to the episode on Vod platform: `https://vod.tvp.pl/video/postaw-na-milion,30102021,56300876`
   Level of difficulty: 4 answers
   Question: *"O czym ty do mnie rozmawiasz. Głowa Cię nie boli?" To cytat z filmu:*
   a. Rejs, b. Chłopaki Nie Płaczą, c. Kiler, d. Nie lubię poniedziałku

   Participants in their attempt to infer the correct answer rely on inductive reasoning — specifically some form of Mill's methods of agreement and difference. Firstly, they enumerate a number of properties of each movie. They say that *d. Nie lubię poniedziałku* and *a. Rejs* are older movies (from 70s), and *b. Chłopaki Nie Płaczą* as well as *c. Kiler* are newer ones (from 90s) and the form of language and words (some kind of slang) uttered in the quote cited in the question does not fit older movies, so they reject those two answers (i.e., options *a.* and *d.*). Therefore, contestants are left with *b. Chłopaki Nie Płaczą* and *c. Kiler*. So now, the set of possible answers is {b, c} between which players allocate all the money.

2. Link to the episode on Vod platform: the same as above (in number 1).
   Level of difficulty: 3 answers
   Question: *Który narząd człowieka waży przeciętnie najwięcej?*
   a. mózg, b. serce, c. wątroba

   Here participants employ inductive reasoning, specifically Mill's method of difference to reject one (incorrect) answer at a time. Firstly, they conclude that the correct answer cannot be *b. heart*, as a human heart is much smaller than the other two organs, and therefore this is the reason why its weight is lower than the other two organs. Secondly, they recall that although a brain and a liver have similar size, the former has very different consistency than the latter — a liver being firm, solid and a brain being more jelly-like. This lets them infer that brain's weight is lower than liver's weight, thus making the participants choose *c. liver* as the correct answer.

3. Link to the episode on Vod platform: `https://vod.tvp.pl/video/postaw-na-milion,23102021,56176016`
   Level of difficulty: 4 answers
   Question: *Do której rodziny roślin botanicy zaliczają sałatę?*
   a. bobowatych, b. astrowatych, c. kapustowatych, d. selerowatych

   In this example participants employ reasoning from analogy as they recall that a lettuce reminds a cabbage (probably in terms of appearance — like green leaves). Here participants explicitly say that they are not one hundred percent sure of their choice. So, they most probably rely only on some shallow characteristics of those plants and plant families instead of some in depth knowledge about them.

4. Link to the episode on Vod platform: `https://vod.tvp.pl/video/postaw-na-milion,16102021,56054442`
   Level of difficulty: 4 answers
   Question: *Według badania "Aspiracje dziewczynek w Polsce" najwięcej nastolatek marzy o karierze youtuberki. Na drugim miejscu jest:*
   a. trenerka personalna, b. prawniczka, c. księgowa, d. projektantka stron internetowych

   Here, participants make use of reasoning from analogy as they assume that career as a Youtuber (video creator on the YouTube platform) is similar to career as a fitness trainer which in today's popularity of social media sites is also a content creator on platforms such as Instagram and YouTube. Therefore, making Youtubers and fitness trainers similar professions to a certain degree. Additionally, they could rely on some form of Mill's methods of induction — especially the method of difference in rejecting answers *b. lawyer* and *c. accountant* which are rather old, traditional professions, while a Youtuber and a fitness trainer are the new ones. However, they completely omit the fourth answer (*d. web designer*) and do not take it into consideration.
   Worth mentioning here is the fact that this example is the perfect one showing that conclusions drawn with the use of inductive reasoning are only probable but not certain (as in the case of deductive reasoning). Therefore, conclusions could be either true or false given a set of true evidence. Thus, in this example participants lost in the game as their conclusion turn out to be false.

5. Link to the episode on Vod platform: the same as above (in number 4).
   Level of difficulty: 4 answers
   Question: *Ile uderzeń na minutę wykonywało serce Neila Armstronga, gdy lądownik misji Apollo 11 osiadł na powierzchni Księżyca?*
   a. 220, b. 182, c. 150, d. 75

   At the beginning, players assume that normal heart rate ranges from 75 to 200 beats per minute, so they reject both the low and high ends of this range, and are left with two options — *b. 182* and *c. 150*. However, one can notice uncertainty in one of the participants verbal report — to be more precise, he wonders for some time about the low edge of this spectrum. Participant's verbal report lets me suppose, that he bases his assumptions on some form of bias, such as recency bias[1] or availability heuristic[2], instead of hard and certain knowledge about normal human heart rate — in his verbal report he does not, for example, cite any source of medical information to back up his assumption. Anyway, from the set of four possible answers a, b, c, d they say it is not *a* and not *d*, and therefore reason that the right answer is *b* or *c*, so they split the money between these two answers.

6. Link to the episode on Vod platform: the same as above (in number 4).
   Level of difficulty: 4 answers
   Question: *Największą różnica czasu jest pomiędzy Warszawą a:*
   a. Lizboną, b. Tokio, c. Dubajem, d. Moskwą

   One of the participants remembers that there is eight hours time difference between Tokyo and Warsaw. They immediately disregard Lisbon as it is located very close to the time zone of Warsaw. Then one says that "Moscow is closer to Warsaw [than Tokyo] (...) Dubai is underneath". In their reasoning firstly they follow *modus ponendo ponens*:
   scheme: $((p \to q) \wedge p) \to q$
   If Moscow is closer to Warsaw than Tokyo, then there is less time difference between Moscow and Warsaw than between Tokyo and Warsaw.
   Moscow is closer to Warsaw than Tokyo.
   Then, there is less time difference between Moscow and Warsaw than between Tokyo and Warsaw.

   Further, their reasoning follows something like the law of transitivity (also known as the law of syllogism) and can be reconstructed as follows:
   scheme: $((p \to q) \wedge (q \to r)) \to (p \to r)$
   If Dubai is located closer to Warsaw than Moscow, then Dubai is located closer to Warsaw than Tokyo.
   If Dubai is located closer to Warsaw than Tokyo, then there is less time difference between Dubai and Warsaw than Tokyo and Warsaw.
   Therefore, if Dubai is located closer to Warsaw than Moscow, then there is less time difference between Dubai and Warsaw than Tokyo and Warsaw.

[1] People make use of observations and information encountered more recently rather than the ones encountered long time ago

[2] Availability heuristic describes people's tendency to use information that comes to mind quickly and easily when making decisions

Thus, the correct answer is *b. Tokyo* as both Moscow and Dubai are located closer to Warsaw than Tokyo, so the biggest difference in time is between Warsaw and Tokyo.

3 Conclusions

At the beginning it is worth to mention that present study without a question possesses some limitations — the main ones being a small sample size of analyzed examples — which restricts the set of inferences possible to be drawn from results, and availability of verbal reports — which influences the accuracy of reasoning reconstruction and classification.

Nonetheless, results of my study indicate that in *Who Wants to Be a Millionaire?* as well as in *The Money Drop* some participants employ deductive reasoning and some make use of inductive reasoning. The following types of inductive reasoning were found to be employed by game shows participants in my research sample: reasoning from analogy, statistical reasoning and John Stuart Mill's inductive methods (specifically the method of agreement and the method of difference). I found also a few cases of employing deductive rules of inference, more precisely *modus ponendo ponens*, *law of syllogism*, *modus tollendo ponens* and *modus tollens*. In one case participant relied on heuristics.

Question difficulty seems to does not have an influence on the type of reasoning employed by participants. That is because in the case of *Who Wants to Be a Millionaire?* players used deduction as well as induction in questions with lower as well as with higher money prizes. For the second game show, it is even more difficult to make any sound inferences as participants in five out of six examples are in the first stage of the show with all four possible answers, so there is not much variation in the difficulty level of questions.

The final note would be that participants did not commit any logical fallacies, although some of them, such as negation of the antecedent and affirmation of the consequent, commonly occur in everyday reasoning tasks as well as in experimental studies on human reasoning.

References

J. E. Adler and L. J. Rips. Reasoning: Studies of human inference and its foundations. 2008.

W. B. Arthur. Inductive reasoning and bounded rationality. *The American economic review*, 84(2): 406–411, 1994.

J. S. B. Evans, S. E. Newstead, and R. M. Byrne. *Human reasoning: The psychology of deduction*. Psychology Press, 2019.

J. Y. Halpern. *Reasoning about uncertainty*. MIT press, 2017.

IEP Staff. Deductive and Inductive Arguments. In *The Internet Encyclopedia of Philosophy*. 1995.

P. N. Johnson-Laird. *Human and machine thinking*. Psychology Press, 2013.

J. Joyce, A. Mele, and P. Rawling. The oxford handbook of rationality. 2004.

S. Krauss and X.-T. Wang. The psychology of the monty hall problem: discovering psychological mechanisms for solving a tenacious brain teaser. *Journal of Experimental Psychology: General*, 132 (1):3, 2003.

K. Lam, D. M. Pennock, D. Cosley, S. Lawrence, et al. 1 billion pages= 1 million dollars? mining the web to play" who wants to be a millionaire?". *arXiv preprint arXiv:1212.2477*, 2012.

C. Manfredotti and P. Viappiani. A bayesian interpretation of the monty hall problem with epistemic uncertainty. In *The 18th International Conference on Modeling Decisions for Artificial Intelligence*. Springer, 2021.

M. Oaksford and U. Hahn. Induction, deduction and argument strength in human reasoning and argumentation. *Inductive reasoning*, pages 269–301, 2007.

L. M. Osbeck. Scientific reasoning as sense-making: Implications for qualitative inquiry. *Qualitative Psychology*, 1(1):34, 2014.

J. C. Overholser. Elements of the socratic method: Ii. inductive reasoning. *Psychotherapy: Theory, Research, Practice, Training*, 30(1):75, 1993.

D. Ross. Aristotle: the nicomachean ethics. *Philosophy*, 31(116), 1956.

C. Spiel and J. Glück. A model-based test of competence profile and competence level in deductive reasoning. *Assessment of competencies in educational contexts*, pages 45–65, 2008.

K. E. Stanovich and R. F. West. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and brain sciences*, 23(5):645–665, 2000.

E. W. Tsang and J. N. Williams. Generalization and induction: Misconceptions, clarifications, and a classification of induction. *MIS quarterly*, pages 729–748, 2012.