A Study Assessing The Validity And Reliability Of The General Misconceptions Measure

Name: Martijn de Man

Student number: 2010089

Education program: Bachelor Psychology

Major: Economic Psychology

University: Tilburg University

Supervisor: Willem Sleegers

Second reader: Anneloes Kip

Date of submission: June 30, 2021

Abstract

The majority of studies on misconceptions have used tests for measuring misconception endorsement with methodological shortcomings. This causes problems when interpreting the results of these studies, since there is a possibility of issues with response bias and there is a lack of guess correction. To provide a solution to these problems, Bensley, Lilienfield, & Powell (2014) developed and validated the Test of Psychological Knowledge and Misconceptions (TOPKAM). In this study, we validated the General Misconceptions Measure (GMM) in the same way the TOPKAM was validated. A total of 113 participants were included in the sample. The sample consisted of 38 males, 73 females, and 2 non-binary, with a mean age of 25 years (*SD* = 10.6). We correlated the GMM to measures that were supposed to be related to misconception endorsement, such as conspiracy beliefs, paranormal beliefs, bullshit receptivity, and objectivism. We compared the mean certainty scores for the correct and incorrect answers, since we expected people to be more confident of their incorrect answers than of their correct answers. This would be in line with previous research and allows for further validation of the GMM. The number of misconceptions endorsed by participants on the GMM was positively correlated with measures of conspiracy beliefs, paranormal beliefs, and bullshit receptivity, but no significant correlation was found for objectivism. The difference in mean certainty ratings for the correct and incorrect answers was not significant. The results suggested that the GMM is a valid test for measuring belief in poorly-supported claims (i.e. misconceptions), but it does not measure underlying constructs like objectivism.

Keywords: misconceptions, conspiracy beliefs, paranormal beliefs, objectivism, bullshit receptivity

A Study Assessing The Validity And Reliability Of The General Misconceptions Measure

There are people who believe in claims that are demonstrably wrong. For example, some people believe that Napoleon was short (Wilde, 2020). The reason for this is the work of British cartoonist James Gillray, who pictured Napoleon as a short man. His cartoons became popular and people actually thought Napoleon was short, while in fact his height was average for that time (McIlvenna, 2019). Another popular but false belief is the causal relationship between vaccines and autism (Bordon, 2020). A study suggested that there was a link between getting a vaccine and autism (Wakefield et al., 1998). However, it was later revealed that the data was falsified and the study was refuted multiple times by other researchers (Madsen et al., 2002; Hviid, Hansen, Frisch, & Melbye, 2019). These beliefs that are held contrary to known evidence are defined as misconceptions and they can have harmful consequences (Taylor & Kowalski, 2004).

There are several reasons why misconceptions can be harmful to society. For instance, the misconception about the link between vaccines and autism can be dangerous for public health. A number of people do not want a vaccine, because they believe there is a link with autism. This is problematic, because vaccines are only effective in eliminating diseases if enough people get a vaccine (Gordis, 2013). Furthermore, the acceptance of incorrect beliefs can be an obstacle to effective science teaching (Bensley et al., 2014). This means that students, for example, will hold on to their own views instead of changing their views according to scientific evidence. Additionally, misconceptions are persistent, meaning that people who endorse misconceptions will not change their minds easily (Özgur, 2013). These findings suggest that misconceptions can have harmful consequences for people in general if a large proportion of the population believes in them.

Misconception endorsement is a frequently studied topic. The results of various studies on misconceptions reveal something about the prevalence of misconception endorsement. Research has shown that endorsement of psychological misconceptions is prevalent among the general public, educators, and students (Hughes, Lyddy, & Lambe, 2013). Several studies have shown that a proportion of psychology students endorse psychological misconceptions (Arntzen, Lokke, Lokke, & Eilertsen, 2010; Taylor & Kowalski, 2004; Vaughan, 1977). However, misconceptions are not only limited to psychology. Studies suggested that a proportion of students in other disciplines also endorse misconceptions about their field of study, such as biology (Kummer, Whipple, & Jensen, 2016; Odom, 1995) and physics (Abrahams, Homer, Sharpe, & Zhou, 2015). Apart from students, research has shown that there is also a proportion of the general public who endorse science-related misconceptions (Swami, Stieger, Pietschnig, Nader, & Voracek, 2012). These findings suggest that misconception endorsement is quite prevalent.

To assess the prevalence of misconception endorsement, it is required that we can accurately measure misconceptions, but there are important limitations of how misconceptions are commonly studied. Most studies use a true-false response format for measuring misconception endorsement. Generally, participants are presented with a list of misconceptions, where they have to rate each item true or false. Since all of the items are misconceptions, ‘false’ is the correct answer for each question. Less correct answers equals more endorsement of misconceptions. There are several limitations when using this true-false format for measuring misconceptions.

Firstly, participants can recognize a pattern in the questions. If a participant recognises that the statements are actually misconceptions, they might know that the items are all false.

They will then answer 'false' to each question. When this is the case, the answers from the participant do not represent the endorsement of the misconceptions by this participant. This will result in an underestimation of the prevalence of misconception endorsement.

Secondly, the use of true-false answers does not take into account uncertainty. Participants might not know an answer, but if they are forced to choose between true and false, their answer does not represent what they really think is the right answer. Some people may respond 'true' when they do not know an answer (yea-saying), which leads to inflated estimates of misconception endorsement. On the other hand, some people may respond 'false' when they do not know an answer (nay-saying), which leads to deflated estimates of misconception endorsement (Bensley et al., 2014).

Lastly, misconceptions are not always entirely false. For example, there is a misconception that people only use 10% of their brain. It is true that people do not always use 100 percent of their brains all the time (Wanjek, 2003; Quest, 2013). Therefore, if people know this and they are presented with this misconception, they might think it is true since we do not use 100 percent of our brain all the time and 10 percent may seem acceptable to them. In other words, some participants might answer 'true' on an item that is only partially true. This will result in an overestimation of misconception endorsement.

TOPKAM

Bensley et al. (2014) developed the Test of Psychological Knowledge and Misconceptions (TOPKAM) to eliminate the problems with other measures of misconceptions. They replaced the true-false response format with a forced-choice two-option response format, where one option contains evidence-based psychological knowledge and the other option contains the misconception. By using a two-option response format instead of a true-false

response format, the problem of recognizing a pattern by participants is eliminated. Furthermore, the two-option response format eliminates the problem of yea-saying and nay-saying response styles. Participants have to choose between two options and cannot answer only ‘true’ or only ‘false’ for each question they do not know. This will prevent inflated or deflated estimates. Since psychological knowledge is not always completely true or false, participants are instructed to choose the answer that is ‘most correct’. For each question, participants have to indicate how certain they are about their answer to be able to distinguish between participants who are guessing and participants who are certain of their answer. More correct answers result in a higher score on the TOPKAM. Higher scores on the TOPKAM equals less misconception endorsement.

Bensley et al. (2014) validated the TOPKAM by correlating it to several measures of psychological constructs that should be related to the probability to endorse misconceptions. They proposed that the TOPKAM should be related to measures of paranormal belief, faith in intuition, the ability to distinguish scientific fields and practices from pseudoscientific ones, critical thinking, and SAT scores. Paranormal beliefs are beliefs that violate scientifically founded principles of nature and are held contrary to scientific evidence (Lindeman, 2017). A negative relation was found between paranormal beliefs and the number correct on the TOPKAM, so if participants endorsed more misconceptions, they had more paranormal beliefs. Faith in intuition refers to an individual's trust in their own intuition as opposed to thinking rationally (Alós-ferrer & Hügelschäfer, 2012). There was a negative relation between the number correct on the TOPKAM and faith in intuition scores, this means that when participants endorsed more misconceptions, they had more faith in intuition. Furthermore, participants that were better in distinguishing scientific fields and practices from pseudoscientific ones had a higher score on TOPKAM, thus they endorsed fewer misconceptions. There was a positive relationship between

TOPKAM performance and critical thinking, this means that when a participant endorsed more misconceptions, they had a lower score on critical thinking. Lastly, SAT scores (academic ability) were positively correlated to the TOPKAM scores, meaning the lower the academic performance, the more misconceptions a participant endorsed.

Although it solves the problems when measuring misconceptions, TOPKAM itself also has a limitation, it only uses psychological misconceptions. Psychological misconceptions are only a small part of all the misconceptions. Therefore, non-psychological misconceptions should be included to have a better and more complete overview of misconception endorsement. When a measure represents a more complete overview of misconception endorsement, the prevalence of misconception endorsement, in general, can be estimated more accurately. Furthermore, adding general misconceptions to the current measure would be a further validation of the two-option response format for measuring misconception endorsement.

The Present Study

To provide a possible solution to the limitations of the TOPKAM, we will validate a new measure of misconceptions, the General Misconceptions Measure (GMM). The format of the GMM is similar to the format of the TOPKAM, consisting of a two-option response format where both the misconception and the evidence-based knowledge are presented. The GMM does contain more general misconceptions, instead of psychological misconceptions.

To validate the GMM, we will relate it to measures of conspiracy beliefs, paranormal beliefs, objectivism, and bullshit receptivity. Conspiracy beliefs are beliefs that are usually unsubstantiated and implausible (Brotherton, French, & Pickering, 2013). Therefore, people who believe in conspiracy theories should also be susceptible to misconceptions. Moreover, paranormal beliefs are beliefs held contrary to scientific evidence (Lindeman, 2017). People who

hold these paranormal beliefs will probably be susceptible to misconceptions as well, since both paranormal beliefs and misconceptions lack scientific evidence. Objectivism can be defined as the tendency to base one's judgments and beliefs on empirical information and rational considerations (Leary, Shepperd, McNeil, Jenkins, & Barnes, 1986). This means that people who are objective will seek relevant information, suggesting that they are less likely to endorse misconceptions as they do not just take something for granted. Finally, bullshit receptivity refers to one's receptivity to randomly generated, unsupported claims (Pennycook, Cheyne, Barr, Koehler, & Fugelsang, 2015). People who believe in bullshit sentences should be more susceptible to believing in misconceptions.

To further validate the GMM, we will look at the certainty ratings for the correct and incorrect answers. Previous research found that psychology students are less certain of their correct answers and more certain of their incorrect answers. People are less certain of their answer when they question the truthfulness of the misconception, because it is the opposite of what is often seen as general ‘knowledge’ in society (Landau & Bavaria, 2013). Therefore, we expect the mean certainty ratings of the incorrect answers to be higher than the certainty ratings of the correct answers.

Hypotheses

Our hypothesis is that the GMM is a valid measure, and therefore the score on the GMM (the proportion of misconceptions endorsed by a participant) should be positively correlated to conspiracy beliefs, paranormal beliefs, and bullshit receptivity, and negatively correlated to objectivism. Furthermore, we will examine whether there is a difference in certainty ratings between correct and incorrect answers to further validate the GMM. We expect that participants

are more certain of their incorrect answers and less certain of their correct answers, which would be in line with previous research.

Method

Design And Procedure

We created a survey where we presented the GMM along with measures of several constructs thought to be related to misconceptions, including conspiracy beliefs, paranormal beliefs, objectivism, and bullshit receptivity. We examined the validity of the GMM, by correlating the measures with each other. Participants were asked to fill in an online survey. First, participants completed the GMM. After that, they were asked to fill in measures of conspiracy beliefs, paranormal beliefs, objectivism, and bullshit receptivity in random order. Then, the participants indicated whether they were familiar with the misconceptions presented in the GMM. However, the familiarity ratings were not used in this study. Finally, there were some demographic questions, including a question asking whether they looked up the answers of the GMM. On average, participants spent around 35 minutes answering the survey.

Materials

General Misconceptions Measure

The GMM consists of 30 questions, where each question has two options. One option contains evidence-based information (the correct answer), the other option contains the misconception (the wrong answer). Participants were told that they take part in a knowledge test. Some examples of the items are “People only use about 10% of their brains” versus “People use the entirety of their brains, but not all at once” and “Lightning never strikes twice in the same spot” versus “Lighting can strike more than one time in the same spot”. For each question, participants have to choose between the two options and rate how certain they are about their

answer on a 5 point Likert scale, where 1 = not at all certain and 5 = completely certain. After reverse coding some items, the scores on the GMM were calculated by dividing the sum of all the wrong items on the GMM by the total number of items. This generated a number between 0 and 1, representing the proportion of misconceptions endorsed by a participant. The reliability of the GMM was sufficient according to this study (α = .78).

Generic Conspiracist Beliefs Scale

To assess the belief in conspiracy theories, we used the Generic Conspiracist Beliefs Scale (Brotherton et al., 2013). The scale contains 15 items, each presenting a different conspiracy belief. Some examples of the items are, "The spread of certain viruses and/or diseases is the result of the deliberate, concealed efforts of some organization" and "The power held by heads of state is second to that of small unknown groups who really control world politics". Participants have to rate each item on a 5 point Likert scale, where 1 = definitely not true and 5 = definitely true. To calculate the scores on the Generic Conspiracist Beliefs Scale, we calculated the mean score. Brotherton et al. (2013) reported extremely high reliability (α = .93) for the scale. In this study, we found a reliability of α = .92 for the scale.

Revised Paranormal Beliefs Scale

To assess paranormal beliefs, we used the Revised Paranormal Beliefs Scale (Tobacyk, 2004). The scale contains 26 items, each item presenting a different paranormal belief. Some examples of the items are "The number '13' is unlucky" and "A person's thoughts can influence the movement of a physical object". Participants have to rate each item on a 7 point Likert scale, where 1 = strongly disagree and 7 = strongly agree. To calculate the scores on the Revised Paranormal Beliefs Scale, we calculated the mean score. Tobacyk (2004) reported a four-week

test-retest reliability of .92 for the scale. In this study, we found a reliability of α = .94 for the scale.

Objectivism Scale

To assess objectivism, we used the Objectivism Scale (Leary et al., 1986). The scale contains 11 items on how objective or rational a person is. Some examples of the items are, "I seek as much information as possible before making decisions" and "I see myself as a rational and objective person". Participants have to rate each item on a 5 point Likert scale, where 1 = not at all characteristic of me and 5 = extremely characteristic of me. To calculate the scores for the Objectivism Scale, we calculated the mean score after reverse coding some of the items. The reliability of the scale was acceptable according to this study (α = .77).

Pseudo-Profound Bullshit Receptivity Scale

To assess bullshit receptivity, we used a scale to measure Pseudo-Profound Bullshit Receptivity (Pennycook et al., 2015). The scale contains 10 items, each item is a randomly generated bullshit sentence. Some examples of the items are "Today, science tells us that the essence of nature is joy" and "Imagination is inside exponential space time events". Participants have to rate each item on a 5 point Likert scale, where 1 = not at all profound and 5 = very profound. To calculate the scores for the Pseudo-Profound Bullshit Receptivity scale, we calculated the mean scores. According to this study, the scale has a reliability of α = .85.

Data Analysis

To examine the concurrent validity of the GMM, we calculated Pearson correlations for the GMM with measures of conspiracy beliefs, paranormal beliefs, objectivism, and bullshit receptivity in SPSS. We also investigated whether there was a difference between certainty ratings for correct and incorrect answers on the GMM. To do this, we calculated the mean

certainty score for all the correct items on the GMM and we also calculated the mean certainty score for all the incorrect items on the GMM. After that, we compared the two mean scores using a paired sample t-test to see if the mean scores significantly differed from each other. To investigate whether we have collected enough data, we looked at the confidence intervals for the correlations.

Participants

The survey was distributed among first-year psychology students, who received course credit for taking part in the study. Since the number of respondents was quite low, we decided to also send the survey to friends and family. A total of 209 people participated in the study. After excluding participants that did not consent (n = 35), did not complete the survey (n = 66), did look up answers (n = 11), or were not good at English (n = 0), the total number of participants included in the sample was 113. The sample consisted of 38 males, 73 females, and 2 non-binary, with ages ranging from 16 to 63 years (M = 25.4, SD = 10.6). There were little to no risks when participating in this study, as it was very unlikely that answering the question affected participants emotionally or otherwise. The study was approved by the Ethical Review Board.

Results

Validation Of The GMM

To examine the concurrent validity of the GMM with measures related to conspiracy beliefs, paranormal beliefs, objectivism, and bullshit receptivity, we calculated Pearson correlations between the GMM and these other measures. The score on the GMM (representing the amount of misconceptions endorsed by a participant) was positively correlated with measures of paranormal beliefs (r(112) = .277; p = .001), conspiracy beliefs (r(112) = .241; p = .005), and bullshit receptivity (r(111) = .397; p < .001). This suggests that participants who endorsed more

misconceptions (thus having a high score on the GMM), believed more in paranormal claims, conspiracy theories, and bullshit (see Table 1). No significant correlation was found between the GMM and objectivism ($r(112) = -.042$; $p = .329$).

Table 1

Descriptive Statistics, Pearson Correlations, and 95% Confidence Intervals for the General Misconceptions Measure and other measures.

Measure	1	2	3	4	5	M	SD
1. GMM	-	[.097, .439]	[.059, .408]	[.228, .543]	[-.225, .144]	.497	.183
2. Paranormal	.277**	-	[.391, .657]	[.275, .577]	[-.473, -.139]	2.56	0.98
3. Conspiracy	.241**	.547**	-	[.211, .530]	[-.349, .009]	2.33	0.80
4. Bullshit	.397**	.438**	.382**	-	[-.377, -.021]	2.55	0.74
5. Objectivism	-.042	-.316**	-.176*	-.206*	-	3.53	0.54

Note. GMM = General Misconceptions Measure; Paranormal = Revised Paranormal Beliefs Scale of Tobacyk (2004); Conspiracy = Generic Conspiracist Beliefs Scale of Brotherton et al. (2013); Bullshit = Pseudo-Profound Bullshit Receptivity of Pennycook et al. (2015); Objectivism = Objectivism Scale of Leary et al. (1986).
* $p < .05$
** $p < .01$

Reliability Of The GMM

Since each item of the GMM contains a different misconception, our expectation was that the internal consistency would not be extremely high. If someone endorsed one misconception (e.g., “Napoleon is short”), it does not mean that this person also endorsed another misconception (e.g., “SOS stands for Save Our Souls”). We expected the internal consistency of the GMM to be similar to the internal consistency of the TOPKAM (*KR-20* = .74), because the GMM measured misconception endorsement in a similar way as the TOPKAM. The internal consistency of the GMM was calculated using Cronbach's alpha ($\alpha = .782$), which is a little higher than the internal consistency of the TOPKAM. The fact that the reliability of the GMM

was similar to the reliability of the TOPKAM and the fact that we did not expect the internal consistency to be extremely high, indicated that the reliability of the GMM was sufficient.

Prevalence Of Misconceptions

To investigate the prevalence of misconception endorsement among participants, we calculated the score on the GMM and divided it by the total number of questions. The number that was calculated represents the proportion of misconceptions endorsed by a participant. Participants had an average score of 0.497 (*SD* = 0.183) on the GMM. This means that, on average, participants endorsed 49.7% of the misconceptions that were presented. Table 2 shows the percentage of participants who endorsed a misconception for each item. These findings suggest that misconception endorsement is quite prevalent among participants of this study.

Table 2

Percentage of participants who endorsed the misconceptions.

Misconception	Fact	% Incorrect
People only use about 10% of their brains.	People use the entirety of their brains but not all at once.	28
Cracking one's knuckles may cause arthritis and other harm to your joints.	Cracking one's knuckles causes the release of gas bubbles and is not harmful.	45
Shaving causes hair to grow back faster, stiffer, and darker.	Shaving cuts hair tips causing them to feel coarser.	33
Eating shortly before swimming increases the risk of experiencing muscle cramps and can be dangerous.	Eating shortly before swimming can be unpleasant.	50
It takes seven years to digest chewing gum.	Chewing gum is mostly indigestible.	14
You cannot survive without your appendix.*	It is possible to survive without your appendix.*	9

Ostriches stick their heads in the sand to hide when they sense danger.	Ostriches who sense danger flop down with their head and neck flat against the ground.	42
Bulls are enraged by the red color of the matador's cape.	Bulls are enraged by the movement of the matador's cape.	23
Color change in chameleons is predominantly used for camouflage to hide from possible danger.	Color change in chameleons is predominantly used for social signaling and temperature regulation.	77
Bears hibernate during the winter to save energy.	Bears rest more during the winter.	71
Bats are blind and can only navigate through echolocation.	Bats have eyes and are capable of sight.	63
An octopus has six limbs.*	An octopus has eight limbs.*	8
The seasons are determined by the Earth's distance from the Sun.	Seasons are caused by the Earth's axial tilt.	37
People have five senses.	People have around 20 senses.	70
Lightning never strikes twice in the same spot.	Lighting can strike more than one time in the same spot.	13
Bananas grow on trees.	Bananas grow on bushes.	89
When a coin falls from a high altitude it reaches a high enough speed to be able to kill someone.	Coins dropped from a high altitude are not able to kill someone due to wind resistance.	37
The Solar System contains nine planets.*	The Solar System contains eight planets.*	30
Christopher Columbus discovered North-America.	Leif Erikson discovered North-America.	58
During the Middle-Ages people were expected to live until about 30 years old.	During the Middle-Ages people lived to be older than 30 years but many children died during childhood.	29
Vikings wore horned helmets in battle.	Vikings wore round helmets in battle without horns.	50
Napoleon was an exceptionally small Frenchman.	Napoleon was of average height for French men.	53

Medieval harnesses were so heavy that knights had to be hoisted on their horses.	Medieval harnesses were light enough for knights to run and mount horses without help.	76
The Great War refers to World War 2.*	The Great War refers to World War 1.*	27
You need to wait 24 hours before you can report someone as missing.	If there is evidence of violence or unusual absence you can immediately inform the police about a missing person.	22
Eskimos languages have an exceptional number of words for snow.	Eskimo languages do not have more words for snow than the English language has.	82
"SOS" stands for "Save our Souls".	"SOS" is no abbreviation.	61
The Great Wall of China is the only man-made structure visible from space with the naked eye.	Man-made structures such as highways, dams, cities are visible from space with the naked eye.	73
Big Ben is the name of the clock tower at the Palace of Westminster in London.	Big Ben is the name of the great bell hanging inside the clock tower in London.	44
Danes are the tallest people in the world.*	Dutch people are the tallest in the world.*	35

*Filler items, these items were not included in the calculation of the GMM scores.

Confidence Ratings

To examine whether there was a difference between certainty ratings for correct and incorrect answered items, we calculated the mean confidence ratings for each participant's correct and incorrect answers. The difference between the mean confidence ratings was not significant (p = .098). The mean confidence rating for the correct answers was 3.21 (SD = 0.61) and the mean confidence rating for the incorrect answers was 3.30 (SD = 0.55).

Discussion

The goal of introducing this new measure of misconceptions was to provide a possible solution for the problems when using a true-false response format for measuring misconceptions. There are several problems with measures of misconceptions that have used the true-false response format, since there is a possibility of response style bias and there is a lack of guess correction. The GMM and the TOPKAM provided a possible solution to the problems with the true-false response format by replacing it with a two-option response format presenting both evidence-based information and the misconception. Since misconceptions are not always entirely false, participants were asked to choose the answer that is most correct. For each item, participants indicated how certain they were about their answer. In this study, we assessed the validity and reliability of the General Misconceptions Measure (GMM).

The GMM was validated by correlating it to measures that are thought to be related to misconception endorsement, including conspiracy beliefs, paranormal beliefs, objectivism, and bullshit receptivity. The scores on the GMM showed a significant positive correlation with measures of conspiracy beliefs, paranormal beliefs, and bullshit receptivity. This means that the more misconceptions were endorsed by a participant, the more they believed in paranormal claims, conspiracy theories, and bullshit. The correlation between the GMM and a measure of objectivism was not significant, indicating that there was no relationship between the GMM and objectivism. The mean certainty ratings for the correct answers were not significantly different from the mean certainty ratings for the incorrect answers. The reliability of the GMM was similar to the reliability of the TOPKAM. It was not extremely high, which is in line with the expectations.

The positive correlation between the GMM and measures of conspiracy beliefs, paranormal beliefs, and bullshit receptivity suggests that the GMM is a valid measure of belief in

claims that lack scientific evidence, also known as misconceptions. The absence of a significant correlation between the GMM and objectivism does not affect the fact that the GMM is a valid measure. Objectivism is defined as the tendency to base one's judgments and beliefs on empirical information and rational considerations (Leary et al., 1986). This is different from measures of paranormal beliefs, conspiracy beliefs, and bullshit receptivity, as they measure belief in poorly-supported claims. The fact that there is no significant correlation between the GMM and objectivism, suggests that the GMM does not provide information about the underlying constructs that should be related to misconception endorsement. Instead, the GMM predominantly measures participants' belief in poorly-supported claims, such as paranormal claims, conspiracy theories, and bullshit. The fact that there was no significant difference between the certainty ratings for the correct and incorrect answers, does not affect the validity of the GMM for measuring belief in poorly-supported claims. The certainty ratings only provide information about how certain participants are of their answers and not whether they believe in poorly-supported claims. Although we found some non-significant results, we can still conclude that the GMM is a valid measure, because it shows a significant correlation with measures related to belief in poorly-supported claims.

The results of this study are in line with previous research by Bensley et al. (2014). Bensley et al. (2014) introduced a new measure of psychological misconceptions, called the Test of Psychological Knowledge and Misconceptions (TOPKAM). The TOPKAM provided a solution to the problems with other measures of misconceptions by replacing the commonly used true-false response format when measuring misconceptions, with a two-option response format presenting evidence-based psychological knowledge and a psychological misconception. The TOPKAM was validated by correlating it to measures that should be related to misconception

endorsement, such as academic background, critical thinking, objectivism, and acceptance of paranormal and pseudoscientific claims. Similar to the GMM, the correlation between the TOPKAM and objectivism was not significant. A significant correlation with objectivism is not needed for the GMM and TOPKAM to be a valid measure, as the GMM and TOPKAM measure belief in poorly-supported claims and not the underlying constructs that should be related to the probability to endorse misconceptions. The results of that study suggest that the two-option response format used for the TOPKAM is a valid way of measuring misconception endorsement. In contrast to the findings of Landau and Bavaria (2013), no significant difference was found between the mean certainty ratings of the correct and incorrect answers. However, this does not affect the validity of the GMM for measuring belief in poorly-supported claims. The certainty ratings only provide information about the certainty of participants' answers and not whether they believe in poorly-supported claims. The results of this study and the study of Bensley et al. (2014) provide evidence for the validity of this new method for measuring misconceptions.

The results of this study can be helpful for future research on misconceptions, since this testing method, where both a misconception and evidence-based knowledge are presented, has better validity than other measures that were used previously. Research like this is useful, because it is important to have a valid and reliable measure to have a better understanding of misconception endorsement. Misconceptions can have harmful consequences if a large proportion of the population believes in them. For instance, if people do not want a vaccine because they think it will cause autism, vaccination rates will fall and dangerous diseases will make a comeback. Therefore, it is important to further investigate the effects of misconceptions and to create interventions to reduce misconception endorsement. That is why it is important to have a valid and reliable test for measuring misconception endorsement. The scales used for

measuring misconceptions should not result in an overestimation or underestimation of misconception endorsement. Eventually, this measure can be used for research on how to deal with misconceptions and evaluate interventions to reduce misconception endorsement.

While the GMM provides a solution for problems with previous measures of misconceptions, there are some suggestions for future research regarding misconceptions. The GMM was used to measure general misconceptions, while the TOPKAM was used to measure psychological misconceptions. However, these are only two types of misconceptions. Future research should include other types of misconceptions to see if similar relations are found. For example, misconceptions that are important and have real life implications could be included in a test to see if they have an effect on the relations between the measures. People will probably think more thoroughly about important misconceptions (e.g., "Vaccines cause autism"), than about trivial misconceptions (e.g., "Napoleon was small"). When people are faced with important information, they rely more on system two, their rational thinking system (Kahneman, 2011). It would be interesting to see if we still find the same relations with conspiracy beliefs, paranormal beliefs, objectivism, and bullshit receptivity when people rely more on system two while completing the questionnaire.

Another type of misconceptions to include in future research are misconceptions that challenge people to think logically to come to the correct answer instead of testing their knowledge. An item that could be included is "Dan dreams of becoming a doctor. ... After ending his army service, Dan registered at the university. Which seems to you to be more likely?", with answer options "Dan is a student of the medical school" (misconception) and "Dan is a student" (correct answer). According to the available information, there are people who would choose the first option (Fischbein & Schnarch, 1997). However, if someone thinks logically about this

question, it should be clear that the second option is most likely to be true. The chance of option A is greater than the chance of option A and B. Another item that could be included is "Which pattern is more likely to occur when flipping a coin six times?", with answer options "HTTHTH", "TTTHHH", "HHHHHH", and "All options are equally likely" (H = heads, T = tails). The first option ("HTTHTH") looks most irregular, therefore there are people who think that it is the correct answer (Smith, 1998). In fact, all patterns are equally likely. The probability is the same for each pattern, namely $(.5)^6 = .015625$. When thinking logically about this item, it should be clear that only the last answer is correct. This is often referred to as the representativity heuristic. It would be interesting to see if a measure of misconceptions where people are challenged to think logically to find the correct answer (instead of testing their knowledge), shows different relations with measures of paranormal beliefs, conspiracy beliefs, objectivism, and bullshit receptivity.

While we found significant results in this study, there were also some limitations. First, answering the questions of the GMM can lead to a change in misconception endorsement. There is a possibility that someone endorsed a misconception before they filled in the questionnaire, but since the correct answer was also given, the participant may have realized that the correct answer sounds more logical than the initial misconception and they do not endorse the misconception anymore. For example, when filling in the GMM, participants were presented with the misconception "During the Middle-Ages people were expected to live until about 30 years old" and the correct answer "During the Middle-Ages people lived to be older than 30 years but many children died during childhood". There is a possibility that participants initially endorsed the misconception, since they probably did not take into account that the death of many children could lower the estimated life expectancy when they encountered the misconception for

the first time. So when they see the correct answer, there is a possibility that their misconception endorsement changes, because the correct answer may seem more logical than the misconception they initially endorsed. When this occurs, the participant's misconception endorsement has changed as a result of answering the question.

Another limitation of this study was the small sample size. We are interested in the exact correlations, but the 95% confidence intervals for the correlations are quite wide. Therefore, we do not know the correlations precisely. It was expected that the correlations found in this study were similar to the correlations found by Bensley et al. (2014) for the TOPKAM. The correlation between the TOPKAM and a measure of paranormal beliefs was r = -.22 (this is a negative correlation, because TOPKAM scores represent the proportion correct instead of the proportion of endorsed misconceptions). Since the range of the correlation between GMM and paranormal beliefs goes from .097 to .439 it is unclear whether the correlation found in this study is similar to the correlation found by Bensley et al. (2014). Furthermore, the wide confidence interval makes it difficult to interpret the size of the correlations. When assessing the correlations between the GMM and measures of paranormal beliefs and conspiracy beliefs, the lower bounds of the confidence intervals suggest that the correlations are small (r < .1), while the upper bounds of the confidence intervals suggest that the correlations are medium to large (r > .4). This makes it difficult to decide whether the correlations are small, medium, or large, which indicates that the confidence intervals of the correlations are too wide. A larger sample size allows for a smaller confidence interval, therefore more participants should be included to have a better estimation of the size of the correlations.

Conclusion

We correlated the GMM to several measures that should be related to misconception endorsement. The GMM showed a positive correlation with measures of conspiracy beliefs, paranormal beliefs, and bullshit receptivity. This means that a higher score on the GMM (equals more endorsement of misconceptions) is accompanied by higher scores on measures of conspiracy beliefs, paranormal beliefs, and bullshit receptivity. No significant correlation was found between the GMM and objectivism, but this should not be a problem for the validity of the GMM. The GMM measures the belief in poorly-supported claims and not the underlying constructs that are related to the probability to endorse misconceptions. The difference between the mean certainty ratings for correct and incorrect answers was not significant. However, this does not affect the validity of the GMM for measuring belief in poorly-supported claims. The certainty ratings only provide information about how certain participants are of their answers and not whether they believe in poorly-supported claims. These results suggest that the GMM is a valid measure of belief in poorly-supported claims (i.e. misconception endorsement) and the two-option response format, where one option contains the misconception and the other contains evidence-based knowledge, is a valid way for measuring misconception endorsement.

References

Abrahams, I., Homer, M., Sharpe, R., & Zhou, M. (2015). A comparative cross-cultural study of the prevalence and nature of misconceptions in physics amongst English and Chinese undergraduate students. *Research in Science & Technological Education. 33(1)*. 111–130. doi:10.1080/02635143.2014.987744

Alós-ferrer, C., & Hügelschäfer, S. (2012). Faith in intuition and behavioral biases. *Journal of Economic Behavior & Organization. 84(1).* 182–192. doi:10.1016/j.jebo.2012.08.004

Arntzen, E., Lokke, J., Lokke, G., & Eilertsen, D. (2010). On misconceptions about behavior analysis among university students and teachers. *The Psychological Record. 60(2).* 325–336. doi:10.1007/BF03395710

Bensley, D. A., Lilienfield, S. O., & Powell, L. A. (2014). A new measure of psychological misconceptions: Relations with academic background, critical thinking and acceptance of paranormal and pseudoscientific claims. *Learning and Individual Differences. 36.* 9–18. doi:10.1016/j.lindif.2014.07.009

Bordon, Y. (2020, September 28). Public trust in vaccines. *Nature.* Retrieved from https://www.nature.com/articles/d42859-020-00024-5

Brotherton, R., French, C., Pickering, A. (2013). Measuring belief in conspiracy theories: The generic conspiracist beliefs scale. *Frontiers in Psychology. 40.* 279. doi:10.3389/fpsyg.2013.00279

Fischbein, E., Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education. 28(1).* 96–105. doi:10.5951/jresematheduc.28.1.0096

Gordis, L. (2013). *Epidemiology* (5th Edition). Saunders.

Hughes, S., Lyddy, F., & Lambe, S. (2013). Misconceptions about psychological science: A review. *Psychology Learning and Teaching. 12(1).* 20–31. doi:10.2304/plat.2013.12.1.20

Hviid, A., Hansen, J. V., Frisch, M., & Melbye, M. (2019). Measles, mumps, rubella vaccination and autism: A nationwide cohort study. *Annals of Internal Medicine. 170(8).* 513–520. doi:10.7326/M18-2101

Kahneman, D. (2011). *Thinking, Fast and Slow.* Farrar, Straus and Giroux.

Kummer, T. A., Whipple, C. J., & Jensen, J. L. (2016). Prevalence and persistence of misconceptions in tree thinking. *Journal of Microbiology & Biology Education. 17(3).* 389–398. doi:10.1128/jmbe.v17i3.1156

Landau, J. D., & Bavaria, A. J. (2003). Does deliberate source monitoring reduce students' misconceptions about psychology? *Teaching of Psychology. 30(4).* 311–314.

Leary, M. R., Shepperd, J. A., McNeil, M. S., Jenkins, T. B., & Barnes, B. D. (1986). Objectivism in information utilization: Theory and measurement. *Journal of Personality Assessment. 50.* 32–43. doi:10.1207/s15327752jpa5001_5

Lindeman, M. (2017). Paranormal beliefs. *Encyclopedia of Personality and Individual Differences.* doi:10.1007/978-3-319-28099-8

Madsen, K. M., Hviid, A., Vestergaard, M., Schendel, D., Wohlfahrt, J., Thorsen, P., Olsen, J., & Melbye, M. (2002). A population-based study of measles, mumps, and rubella vaccination and autism. *The New England Journal of Medicine. 347(19).* 1477–1482. doi:10.1056/NEJMoa021134

McIlvenna, U. (2019, November 13). Was Napoleon short? Origins of the 'Napoleon Complex'. *History.* Retrieved from https://www.history.com/news/napoleon-complex-short

Odom, A. L. (1995). Secondary and college biology students' misconceptions about diffusion and osmosis. *The American Biology Teacher. 57(7).* 409415. doi:10.2307/4450030

Overvelde, B. (2013, Juli 3). Gebruiken we echt maar 10 procent van onze hersenen? *Quest.* Retrieved from https://www.quest.nl/mens/psychologie/a25455825/hersencapaciteit-10-procent/

Özgur, S. (2013). The persistence of misconceptions about the human blood circulatory system among Students in different grade levels. *International Journal of Environmental & Science Education. 8(2).* 255–268. doi:10.12973/ijese.2013.206a

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgement and Decision Making. 10(6).* 549–563.

Saher, M., & Lindeman, M. (2005). Alternative medicine: A psychological perspective. *Personality and Individual Differences. 39*. 1169–1178. doi:10.1016/j.paid.2005.04.008

Smith, H. D. (1998). Misconceptions of chance: Developmental differences and similarities in use of the representativeness heuristic. *Psychological Reports. 83(6).* 703–707. doi:10.2466/PR0.83.6.703-707

Swami, V., Stieger, S., Pietschnig, J., Nader, I. W., & Voracek, M. (2012). Using more than 10% of our brains: Examining belief in science-related myths from an individual differences perspective. *Learning and Individual Differences. 22.* 404–408. doi:10.1016/j.lindif.2011.12.005

Taylor, A. K., & Kowalski, P. (2004). Naïve psychological science: The prevalence, strength, and sources of misconceptions. *The Psychological Record. 54.* 15–25. doi:10.1007/BF03395459

Tobacyk, J. J. (2004). A revised paranormal belief scale. *The International Journal of Transpersonal Studies. 23.* 94–98.

Vaughan, E. D. (1977). Misconceptions about psychology among introductory psychology students. *Teaching of Psychology. 4(3).* 138–141. doi:10.1207/s15328023top0403_9

Wilde, R. (2020, January 29). Was Napoleon Bonaparte really short? *European History.* Retrieved from http://europeanhistory.about.com/od/bonapartenapoleon/a/napoleon height.htm

Wakefield, A. J., Murch, S. H., Anthony, A., Linell, J., Casson, D. M., Malik, M., Berelowitz, M., Dhillon, A. P., Thomson, M. A., Harvey, P., Valentine, A., Davies, S. E., & Walker-Smith, J. A. (1998). Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet. 351(9103).* 637–641. doi:10.1016/S0140-6736(97)11096-0

Wankjek, C. (2003). *Bad Medicine: Misconceptions and Misuses Revealed, from Distance Healing to Vitamin O.* John Wiley & Sons.