Weaponized Meta-Memetics

Measuring the Effectiveness of Internet Memes as Ideological Propaganda

Student Name:	Nicolas Koutonias
Student Number:	430273
Supervisor:	Dr. João Fernando Ferreira Gonçalves

Master Media Studies – Media, Culture & Society

Erasmus School of History, Culture and Communication

Erasmus University Rotterdam

MA Thesis

September 25th, 2020

18 375 words

ABSTRACT

Over time, internet memes have grown beyond the jokes and humorous cultural artefacts of once fringe online communities. Nowadays, they've become so widespread as to be a self-reinforcing genre of online communication, and they have therefore been the subject or increasing scrutiny and study, both in academia and in popular media sources. Although the process of how internet memes disseminate information can be applied to both positive and negative information, the efficiency through which they can spread information has made them the ideal vehicles for bad faith actors to use them to spread misinformation. Misinformation-carrying and offensive memes have, as well, contributed to the radicalization of people. Through informational overload inherent to Web 2.0, ideologically harmful internet memes may have also played a role in the normalization of extreme, previously fringe ideologies. There is a lack of quantitative experimental research into the extent to which harmful (racist, for example) internet memes can lead viewers to later consider that same harmful ideological position as *more normal* than if they hadn't been first exposed to it through memes. Overexposure to a harmful ideology through memes could be theorized to *prime* people into being more accepting of more extreme forms of prejudice later, although one's level of knowledge and literacy in 'internet meme language' could *inoculate* them against this negative effect. This study attempts to demonstrate this normalizing effect, but ultimately fails to do so, in part due to the relatively small sample being insufficient to achieve statistically significant results. Although this study's results lack significance, the theoretical background and how it motivates the study design could prove to have value, and would benefit from being replicated on a larger scale, with more respondents.

KEYWORDS

Experiment – Propaganda –Racist – Meme – Acceptability

TABLE OF CONTENTS

1. INTRODUCTION

1.1 General Introduction

Although they have existed in one form or another since the early days of the internet, internet memes have gone through a process of mainstreaming since their proliferation was facilitated under Web 2.0's focus on online *participation* (Shifman, 2014, p. 18). These memes have grown from often humorous, always self-referential images spread through social networks to a significantly intricate genre of communication between participants of various online spaces (Wiggins & Bowers, 2015).

Much academic and non-academic research has been conducted into the creation, understanding, and use of internet memes, and even more into how culture and cultural artefacts spread; the field of memetics, from which online culture appropriated the name assigned to the unit of culture, the meme. As internet memes have developed, however, a growing focus in literature has been into the significant socio-political power they may hold (Bertazolli, 2019; Huntington, 2019; Moreno-Almeida, 2020), as a concise form of ideological communication.

Internet memes, and their ease of dissemination, can have a positive impact on society by allowing information to spread, contributing for instance to the construction and reinforcement of collective identity of marginalized groups (Gal, Shifman & Kampf, 2016). They can also, however, have a negative impact, by allowing misinformation and disinformation to spread and catch on just as easily. The recent phenomenon of fake news, for instance, owes a lot of its intensity and consequences to the ease though which internet memes can disseminate false information (Smith, 2019). Similarly, internet memes can also spread harmful ideas, of a racist, sexist, of homophobic nature for instance, while maintaining a relaxed detachment towards the harmfulness of the carried information (Drakett, Rickett, Day & Milnes, 2018). A racist meme, for example, will have its offensiveness defended under the guise of it having been "only a joke" (Williams, Oliver, Aumer & Meyers, 2016, p. 425).

In the following, I will argue that while certain offensive memes can indeed be perceived as simply jokes – in poor taste, hurtful, but jokes nonetheless – other offensive memes can have a more serious, more insidious negative effect on those who view them.

Someone who does not hold racist beliefs will not necessarily have their minds changed from exposure to a racist meme, but the harmful information it carries may initiate a process of subconscious normalization in that person, and they may later encounter a more extreme form of racism, in a meme or a different format, and not perceive it to be as harmful as it actually is.

Through experimental conditions in an online survey, I will expose consenting participants to different formats of offensive racist information, in a small sample of memes and a short paragraph of text. Different groups will be exposed to memes and text in different orders, to experimentally determine if exposure to a particular harmful idea in an internet meme *first* can influence later consideration of the same idea in a textual format, with respondents viewing it as more credible, less harmful, and therefore more 'normal'. My first research question, then, will be as follows:

> **To what extent does the exposure to a harmful idea through internet memes influence the later acceptance of that same idea as normal, in different formats?**

The online communities where internet memes generally proliferate, however, give involved members a certain level of comprehension exclusivity. Memes are, at their most fundamental level, inside jokes shared between friends. Even when these small groups have grown far beyond any reasonable scope of personal acquaintance they may have held in the internet's early days, the dynamic of memes being most understandable by those 'in the know' has remained. As such, I will conceptualize and attempt to measure such 'internet meme literacy' in my respondents prior to exposing them to the stimuli (in whichever order), to see if intimate knowledge of the 'language' of internet memes can somehow protect respondents from their negative ideological effect of normalization. The second research question, therefore, will be:

> **To what extent does high internet meme literacy protect viewers from accepting a harmful idea as normal, if initially exposed to it through internet memes?**

In short, with the present research project, I will experimentally measure the effectiveness of offensive memes as forms of ideological propaganda, leading respondents to accept the existence of harmful ideologies more willingly, through a process of normalization in current internet meme culture.

1.2 Relevance

The present research will be of an exploratory nature but is nonetheless not completely novel. Some research has already been conducted into the role that online communities – where memes proliferate – have played into radicalizing certain groups and people (Albrecht, Fielitz & Thurston, 2019; Benkler, Faris & Roberts, 2018; Christensen, Spahiu, Wilson & Duval, 2015). This research, however, has not focused on the specific impact of internet memes. While they were mentioned as taking a role, there were never the focus of the research. In addition, these studies have tended to be qualitative in their method, and 'after the fact' in their exploration. Their focus has been on things that happened.

My research, in contrast, will instead focus on things that *might happen*, by focusing on the process itself, having broken it down to its basest elements. Similar research has already been conducted by Huntington (2019), but whereas she focused on political misinformation through memes specifically, I instead aim to focus on a more general ideological, rather than political, influence. In this way, it will be adopting a quantitative research method, something not often seen in online radicalization research. Further yet in the exploratory aspect, my research will be one with an experimental design, with clear conditions, stimuli, and statistically measured effects acquired in a clear testing environment unaffected by potential outside influence.

Overall, existing online radicalization research, even when pertaining more specifically to the role played by internet memes, has nevertheless approached the topic in an after-the-fact manner through qualitative methods. A focus on the meta-memetic process of information dissemination (Merrin, 2019) can be more useful than a strict focus on internet memes when dissecting 'what happened', but tactics devised to combat this negative influence could then be countered by bad faith online actors, and put in simplest terms, the damage has 'already been done' by that point. A stricter focus on the influence offensive internet memes can have *specifically*, free of the influence of other contributing factors of online radicalization, could prove useful in the eventual development of a method to prevent this radicalization from growing beyond reasonable control. As meta-memetically-disseminated misinformation and disinformation has spread in recent years, the social utility of my research could provide a starting point for developing educational methods to combat 'meme fake news': By showing how the process of informational dissemination occurs in an

experimental context, it would be possible to extrapolate to how it occurs on the internet at large, and methods to combat this process before it properly begins may then be conceptualized to protect the increasing number of people who are becoming active users of the participatory web.

1.3 Structure

The present research project will be exploring a significant amount of literature to justify its purpose and utility. To the average layperson, internet memes may not seem as posing a significant risk to those who view them. In order to credibly establish the risk they *could* pose, it is important to begin research with the foundational texts that will inform and legitimize the idea that internet memes can be powerful communication tools. As such, the following section of this paper will give an overview of existing theories of how ideas and culture spread, by exploring the development and legitimization of (classical) memetics as a theory of human culture. With this foundation laid out, I will then apply the research lens of memetics to the topic of internet memes, showing how wider culture and internet memes are similar, but also how they are different. In doing so, I will conceptualize a *meta-memetic* frame of analysis to be applied to internet memes, considering their highly self-referential (or *meta*-referential) and intertextual nature.

Having justified the value of studying internet memes beyond the scope of humorous online cultural artefacts, and justified their validity as a new genre of online communication, I will show how their dissemination process can be 'hacked' to spread information, whether positive or negative. Through this hacking of the memetic process, I will demonstrate how internet memes can be powerful propaganda tools. In essence, I will show how the process of normalization I study *has already* happened, to justify the value of testing said process experimentally.

The corpus of selected academic literature will provide a solid base upon which to build my study design. This research being quite exploratory, some leaps may be taken in the conceptualization of measures to measure internet meme literacy and perception of content as more 'normal' than other, but the general study design will be sound enough to allow future research to fine-tune and add to it, and attempt to replicate results. In the final sections of this research report, I will describe my dataset, the steps taken to prepare it for analysis, and the statistical tests that were ran, as well as their results. These results will be discussed within

the context of the existing literature, and their limitations explored, before giving suggestions for future research into the topic of internet-meme-driven ideological propaganda.

2. THEORY

2.1 Introduction

Today, it could be reasonably stated that the term ‘meme’ has taken on a meaning that is quite different from originally intended. Indeed, prior to the colloquialization of the word to refer to humorous self-referential images under Web 2.0, *meme* would be used in the academic field of *memetics*, or in the study of how human culture spreads and is disseminated.

Consider the process of genetic reproduction, in its simplest form. An individual possesses a particular genetic trait, and if they form a relationship with another individual and produce offspring, the gene will be passed down to their descendant, or it may not, depending on how the genetic material of both parents will influence the development of the child. This process happens to all living creature on earth. One must wonder, then, how non-genetic artefacts are passed down from one generation to the next. One such artefact is that of culture.

Culture is not embedded genetically, and yet, it maintains itself throughout the generations. It changes, but if one were to trace back any culture throughout the years, they may see a logical progression. How, then, does such a transfer occur if cultural artefacts are not genetic? In one chapter of his 1976 seminal work *The Selfish Gene*, evolutionary biologist Richard Dawkins (1976) sought to explore how such a non-genetic artefact as culture may spread between generations. To describe such a process, he wished to create a word that would be analogous – and therefore seemingly related – to *gene*, but used to describe non-genetic inheritance. From the Greek lexical root *mimeme* ‘to imitate’, Dawkins thought an appropriate monosyllabic word that would rhyme with *gene* was *meme* (p. 192).

Before returning to the current, colloquial meaning of meme as online cultural artefact, we must briefly examine how the term, as well as the field of memetics, came to be relevant.

2.2 Classical Memetics

The focus of memetics is the exploration of how culture – a non-genetic artefact – is disseminated, both vertically across generations, and horizontally between peers. Citing

previous research (Jenkins; referenced p. 189), Dawkins (1976) explains how the male saddleback bird native to New Zealand sings songs that are unique to certain groups or families of birds. These birds are part of different 'dialect groups', and specific songs are passed down from father to son, in a manner that can only be described as cultural reproduction. On occasion, errors are made in the process of learning songs, are not corrected, and these random mutations give rise to new songs, in a way that could be seen as similar to how human language and culture evolve (p. 190).

Culture and cultural reproduction, therefore, are not unique to humans. What makes human culture unique, Dawkins (1976) argues, is its complexity. The complexity of human culture is also why he expresses his dissatisfaction in his fellow geneticists' explanation of human behaviour and culture as existing to ensure group survival through "reciprocal altruism" (p. 191) between members of small kin groups. This conception of cultural beliefs as facilitating biological (and therefore, genetic) survival fails to account for the complex differences between various ancient human cultures across the world. Culture, therefore, must be studied on its own terms, although the evolutionary process can provide a valid conceptual framework to study it.

At its fundamental level, biological reproduction is a process of genetic *imitation*, the smallest *unit* of which is the gene. Cultural reproduction can then be reasonably assumed to *also* be a process of imitation, the smallest unit of which is the meme. Memes can be "tunes, ideas, catch-phrases, clothes, ways of making pots or of building arches" (Dawkins, 1976, p. 192). Similarly to how genes "propagate in the gene pool by leaping from body to body" via biological material, memes "propagate themselves in the meme pool by leaping from brain to brain" via imitation (p. 192). Memes can be more successful in propagating through the meme pool of human culture if they 'catch on', and their success at remaining in the meme pool can be measured along three different criteria; longevity (how long they stay relevant), fecundity (how 'catchy' they are), and copying-fidelity (how easily and faithfully they can be reproduced) (p. 194).

In later writings, Dawkins (1999) himself will admit that his original conception of memes, cultural reproduction and memetics were quite "modest" (p. xvi), and that his original goal with the last chapter (in the original 1976 edition) of *The Selfish Gene* was not to provide a concrete "theory of human culture" (p. xvi), but to instead illustrate how the overall process of evolution was one of *replication*, and that many units of replication were

possible, it just so happened that biological replication through genes came to be the dominant method on our planet. Memes were given as an example of a different kind of replicator, to illustrate his more general theory of 'universal Darwinism', or the "application of Darwinian thinking beyond the confines of biological evolution" (Blackmore, 1999, p. 5).

Ironically, the idea of *meme* caught on memetically, and memetics developed into a field that could potentially constitute a legitimate theory of human culture. Culture, as Dawkins (1976) explained in 1976, is not unique to humans, but innate imitation might be, as Susan Blackmore (1999, pp. 3-4) argues in her 1999 contribution to the developing field of memetics, *The Meme Machine.* In a simple but powerful example, she writes that if one "blinked, waved, or 'goo gooed', or even just smiled" (p. 3) at a baby, they would often imitate the behaviour right back, in a manner that comes so easily only to humans. When imitation occurs, *something* is passed on, to then "be passed on again, and again, and so take on a life of its own" (p. 4). That something is the meme. Memes, although carried and transmitted by hosts, tend to spread indiscriminately of their utility. They spread as if having a life of their own, whether they are "useful, neutral, or positively harmful to us", the latter of which can for instance be "chain letters, pyramid selling, new methods of fraud and false doctrines, ineffective slimming diets and dangerous medical 'cures'" (p. 7).

Although writing over 30 years after Dawkins' (1976) initial conceptualization of the concept and field of study, Blackmore (1999) did agree there still were problems to be resolved in the ongoing science of memetics, most important of which was giving an answer to the question "what *really* is the unit of the meme?" (p. 54). The answer to that, in short, is that no one really knows. Fellow memeticist Dennett (1995, p. 344; referenced p. 54) defined the unit of the meme as "the smallest elements that replicate themselves with reliability and fecundity", but Blackmore (1999) argues against this, with the example of Beethoven's *Fifth Symphony*. Is the meme the whole symphony, or its first – highly identifiable – four notes? The answer to this question is a socially defined one: Different people, with different life experiences and expertise, will consider the memetic unit of Beethoven's *Fifth* differently; musicians would think one way, laypeople another, and both can be correct. Just because an academic theory (such as memetics) is *new* and still unclear does not render it invalid or unworthy of further research (p. 56).

At this point in the existing literature, the important points to remember are thus: Memes are socially defined, spread not necessarily because they are true but because they are

catchy, and finally, they can group up to form self-reinforcing 'co-adapted meme complexes' (or *memeplexes*) that self-perpetuate throughout the meme pool. A few examples of successful memeplexes are religions, long-lasting scientific theories, or political ideologies (Blackmore, 1999, p. 65).

The above characteristics of memes can be used to explain, for example, how superstitious beliefs can form. To strengthen the argument Blackmore (1999) makes about this process, we must make a quick aside to the social psychology research theory of *priming*, used since the late 1970s to study "person perception, attitude activation, and stereotyping" (Dillman Carpentier, Roskos-Ewoldsen & Roskos-Ewoldsen, 2008, p. 186). As a theory rooted in "network models of semantic memory" (p. 187), priming experiments have shown that respondents who were exposed to a certain stimuli that activated one concept in their mind then had other, related concepts activated, which went on to influence information processing when considering later stimuli. In one such experiment, conducted by Srull and Wyer (1979), respondents were given four words ('he', 'Sally', 'hit', and 'kicked') and asked to construct a sentence, for which they only had two options ('He hit Sally' and 'He kicked Sally'). This process was shown to prime respondents by activating memory concepts of hostility and aggression, and they tended, in the next, seemingly unrelated experimental task, to skew towards hostile impressions of an ambiguously described person (Dillman Carpentier et al., 2008, p. 187). Later experiments in the field went on to demonstrate the networked nature of conceptual memory, where certain ideas can be seen to activate related ideas and influence later information processing. This network model of ideas and concepts, not unlike memeplexes, gives us a glimpse of how the memetic process can be 'hacked' to lead people to believe certain things. The experiment described above illustrates the main idea behind priming for the sake of clear explanation, but the process itself can take place more subtly: Gerend and Sias (2009), for example, demonstrated how the colour that surrounds a piece of information in the format in which it is presented can influence how respondents will process it. They also showed how framing a piece of information in a positive or negative light can change how later stimuli will be considered, even if the information itself was not modified. How an individual processes and interprets information can therefore be influenced by the framing of the information, as well as the context and format through which the information is presented, in addition to the processing influence exerted by that individual's knowledge and previous life experiences.

Blackmore (1999) gives one example of a successful memeplex disseminated despite not being true in the case of alien abductions. Sleep paralysis has been a common human experience across the ages, perceived by sufferers as "buzzing and humming noises, vibrations of the body or bed, a powerful sense that there is somebody or something in the room with you, and strange lights floating about" (p. 176). Sexual arousal that sometimes happens in dreams can persist throughout the paralysis. Different cultures through history offered different (memetically successful) explanations of the event, one example being the "incubus and succubus of medieval times, […] evil spirits sent to tempt the wicked into sexual activities" (p. 177). Within highly religious social contexts of the medieval era, the superstitious explanation for sleep paralysis was that of demons; in today's context of pop culture saturation of "outer space, spaceships, UFOs, and sinister aliens" (p. 177), the superstitious explanation becomes alien abductions. As such, previous knowledge (or lack thereof, in the case of scientific explanations for sleep paralysis or near-death experiences, another successful memeplex linked to religion), experiences, and general social and societal context can *prime* individuals to consider certain memes a certain way. The memetic process, then, can be taken advantage of, in this case by television companies looking to make money off sensational programming which plays into peoples' conspiratory and superstitious mindsets (p. 178).

Television executives have harnessed the power of the memetic process for the sake of increasing profits. On a more basic level than the one mentioned above, this harnessing can be as simple as an advertising jingle being catchy and memorable (Blackmore, 1999, p. 54). Other examples given by Blackmore (1999) of taking advantage of the memetic process for the sake of profits are divination and fortune-telling, astrology, or alternative medicine (pp. 182-186). Do, however, remember that memetic spread can happen independently of the meme carriers, so while the above entities may have taken advantage of the process, they may have done so unconsciously: If something worked, they simply kept doing it, even without understanding the intricacies behind it (p. 183). Further along the development of memetics as a scientific field, Conte (2001) rejected previous considerations of meme carriers (people) as passive and assigned to them a more active role in the memetic process. Rather than passive agents taking advantage of a process they may not understand to further their own goals, active agents can instead be seen as having *weaponized* the memetic process for personal benefit, or for the benefit of those they care about and wish to see succeed. In the analogy of television executives mentioned above, they can longer be seen as 'giving the

people what they want' and profiting from it, but instead 'telling the people what they want', and the people accepting it, in a manner not unlike that of propaganda. Later explaining Conte's (2001) writing within the context of research into internet memes, Shifman (2014) explains that earlier memeticists saw individuals as simple *vectors* of cultural/memetic transmission, which happened regardless of their intentions, but later research reframed them as *actors* behind the process (p. 12), keenly aware of the process itself, and using that knowledge to their advantage.

Knowing how the process of memetic dissemination takes place, how it can be weaponized by actors, and how priming theory tells us that different types of stimuli can influence processing and interpretation of later stimuli, we can begin to see the possible negative consequences for viewers of internet memes. Let us reconsider the experiment conducted by Srull and Wyer (1979): They showed that an initial stimuli which activated mental models of hostility and aggression affected how respondents considered later stimuli. With this process in mind, let us now consider how something similar could happen with internet memes: If an individual on the internet is part of an informal community where internet memes are strongly present, these memes would similarly activate certain mental models that could then affect how the individual user processes later information. If internet memes derive their humour from disparaging minority racial groups, for example, those who view them – *even if* they remain aware of the memes being 'jokes' – might have the offensive racist elements mentally normalized to a certain extent, and may therefore be *primed* to consider more extreme forms of racism (in internet memes or in other formats) as less offensive/harmful, and therefore 'more normal'. Later sections of the present report will clarify the legitimacy of risk that offensive internet memes could pose, but we must nevertheless present here the first hypothesis of this research, as built on the first research question[1], in order to justify the importance of priming as an overarching theory of weaponized memetic influence.

> **<u>Hypothesis 1</u>: Respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were primed to perceive it as *more normal*, when compared to other groups.**

[1] 'To what extent does the exposure to a harmful idea through internet memes influence the later acceptance of that same idea as normal, in different formats?'

A detailed explanation and justification of the operationalization process will be given in later methodological sections, but it must be mentioned that the *perceived normality* in hypothesis 1 was conceptualized as *higher perceived credibility of information* and *lower perceived harmfulness of information.* Hypothesis 1, therefore, may be split thusly:

> **<u>Hypothesis 1A</u>: Respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were primed to perceive it as more credible, when compared to other groups.**
>
> **<u>Hypothesis 1B</u>: Respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were primed to perceive it as less harmful, when compared to other groups.**

With the internet becoming more widely used, Blackmore (1999) predicted that the memetic process would accelerate. She argues that as technology progressed, the process of memetic spread had perfected itself; "books, telephones, and fax machines" (p. 205) allowed greater fidelity, fecundity, and longevity of memetic artefacts. The internet, then, with its immateriality of storage and near-instantaneous transfer of information, seems like the next logical step in technological affordances for memetic spread. Later examining how the internet has changed the memetic process, Knobel (2006) gives us a glimpse of what would go on to encapsulate 'internet meme culture', when applying a memetic frame of analysis to early viral videos and internet memes. One of Knobel's (2006) key contribution to memetics as adapted to internet culture is an even greater focus than Conte's (2001) on the agency of the memes' disseminating agents. As described, one characteristic of successful memes is the sense of meaningfulness; "units of information that make more sense or are meaningful to a person and can be successfully imitated or reproduced will more easily become memes than units of information that are not easily copied or understood" (Knobel, 2006, p. 413). This characteristic can tell us two things: First, the more concise and easily transmittable a meme is, the more chances of success it has, and second, successful memes have a higher chance of success in groups that understand them more.

The internet, with its affordances of any layperson becoming a meaningful participant of 'the conversation', has reinforced the possibilities of in-group cohesion for people with fringe beliefs who found "people like me" (Knobel, 2006, p. 415), and in those informal belief-driven informal groups, memes that are transmitted by "trustworthy others" (Brodie, 1996, p. 152; referenced p. 415) who were also part of the group had a higher chance of

proliferating. The online platforms these informal groups occupy can be referred to as "affinity spaces" (p. 415) for ideologically like-minded individuals, who through the space they shared, could feel more closely related than simply being two strangers on the internet.

2.3 Meta-Memetics

In the affinity spaces that Knobel (2006) describes, viral online content of the nature of what would eventually be considered an 'internet meme' proliferated. As she explains, "online, contributing directly to spreading a new, popular and catchy meme is considered cool, and generating an entirely new successful meme is even cooler" (p. 416). Most importantly, however, "being among the first to spot a new, popular and catchy online meme is perhaps coolest of all" (p. 416). Through the fringe nature of early internet-meme-rich online affinity spaces and communities, as well, successful memes were "often absurdist in nature" and "closely akin to shared jokes between friends" (p. 416), exemplifying the strong communal aspect of these groups.

Knobel (2006) continues by giving two examples of successful pre-Web 2.0 internet memes, but before explaining them, it is important to clarify a logistical issue of the present text. In the previous section, examining the birth and early aspects of memetics, 'meme' was used to describe the meme as *unit of cultural imitation*, with 'internet meme' used to refer specifically to *humorous self-referential units of online cultural imitation*. As we move closer to present day in the literature, we must switch the definitions for the sake of reading and comprehension convenience. From this point forward, 'meme' will be used to refer to *internet memes*, and 'cultural meme' to refer to earlier conceptualizations of the term. 'Memetics' will continue to be used as it has been thus far.

In the foreword of *The Meme Machine*, Dawkins (1999) summarised one of Blackmore's (1999) important contributions to memetics; differentiating between two types of cultural memetic imitation, *copying-the-product* and *copying-the-instructions* (p. xi). He asks readers to imagine an experiment: Two groups of twenty children are tasked with folding a sheet of paper in certain ways to create a paper crane. In group A, the first child makes the crane according to instructions, and the second child must replicate it *without* being given instructions, they can only look at the first child's completed crane. Each of the study participants must complete the crane only by looking at the previous participant's completed crane. This is copying-the-product. In group B, the first child is given instructions,

completes their crane, then *explains the instructions* to the next child, and so on. This is copying-the-instructions. Dawkins (1999) argues that if such an experiment were conducted, the crane made by the last child of group A will plausibly look nothing like the first: A random error (mutation) along the way will have completely altered the final product. In group B, however, random mutations will not have been carried over, and the final crane should look mostly like the first. In copying-the-instructions memetic transfer, "the instructions are self-normalising" (p. xii), and a certain 'essence' of the product is transferred, unaffected by random errors and mutations. This separation helps us in understanding the two successful memes that Knobel (2006) explains.

When focusing on cultural memes, it seems natural to be thinking about the potential for random errors to irrevocably change the final product, although with internet memes, the point seems irrelevant. Copying-the-product becomes as simple as sharing an image or forwarding an email exchange, and other than some potential image compression artefacts, the final meme of this line of transmission will be the same as the first, so copying-the-product can therefore be reframed as sharing something 'as-is'. One such successful meme was an email exchange between student Jonah Peretti and Nike's email support (Knobel, 2006, p. 416). In 2001, the company ran a campaign of customer personalization, and Peretti placed an order for running shoes with the word 'sweatshop' embroidered on them. The sportswear company denied his request, and in the subsequent email exchange, the support worker seemingly avoided giving Peretti a specific reason as to why. After he shared it with a couple of friends, who shared it with a few people, who did the same, etc., the email exchange went viral[2], seen by millions and giving Peretti media notoriety as having commentated on corporate worker exploitation in a manner that can also be described as "culture jamming" (Lievrouw, 2011, p. 79), a genre of new media activism which "takes the form of popular culture", yet aims to "subvert[ing] and critique[ing] that culture" (p. 73). This meme, shared and disseminated as-is, exemplifies copying-the-product in an internet context.

The other successful meme described illustrates copying-the-instructions as applied to internet meme culture. In 2004, a hand-drawn flier was posted in the streets of Seattle. A 16-year-old autistic boy had made the flier after losing his toy frog. In September of that same year, an online user posted an picture of the flier, and members of that online space went on

[2] "(of an image, video, piece of information, etc.) circulated rapidly and widely from one internet user to another: a viral ad campaign; the video went viral and was seen by millions." Viral, *Oxford English Dictionary.*

to create their own versions of the flier, pushing the meme into widespread virality, where all iterations collectively "narrate massive, albeit fictional, citizen mobilization in ongoing search for Hopkin Green Frog" (Knobel, 2006, p. 419). All these memes were different, but they all adopted the same 'format', in a sense all following the same instructions, yet remaining recognizable as individual instances of the 'same' meme.

The recognizability of different iterations (or versions) of a meme as still all being 'the same meme', as well as the copy-the-instructions production process, seem to announce later research into the development of internet memes as a full-blown *genre* of online communication, with "its own set of rules and conventions" (Wiggins & Bowers, 2015, p. 1888), and which "obfuscate the consumer/producer binary" in a manner that is representative of Web 2.0 participatory culture as a whole (Jenkins, 2007; referenced p. 1888). Offering a higher degree of specificity in the analysis of online meme culture, Wiggins and Bowers (2015) differentiate between three 'phases' of a meme. At the earliest stage, there is a piece of digital content, one that can be easily shared, that becomes popular and widespread enough to be described as having gone 'viral'. At this stage, the content is described as "*spreadable media*" (p. 1897) and along the three memetic characteristics Dawkins (1976) described, it has low longevity (it may become incredibly popular for a short time, then fade back into obscurity), high fecundity (spreading fast and far), and high copy-fidelity (being shared as-is and not remixed or modified). Knobel's (2006) example of Jonah Peretti's email exchange would here be classified as spreadable media. The next stage is that of "*emergent meme"* (Wiggins & Bowers, 2015, p. 1897), where the initial content has been remixed, parodied, or modified in some way. The emergent meme has similar longevity and fecundity as spreadable media, but lower copy-fidelity, being a modification of the original content (though still recognizably related). The original Hopkin Green Frog flier would be classified as spreadable media, and its first few remixes as emergent meme. What moves content from *emergent meme* to "*meme"* (p. 1899) is its potential for further modification and iteration. When the emergent meme itself goes viral and is further iterated upon, through remix or parody, the intertextuality that is inherent to online meme culture (Shifman, 2014, p. 2) is intensely applied: Various disparate elements of existing memes are mixed in with every iteration of the new meme as it transitions out of its emergent status, thus cementing it as part of the 'meme canon' of the Internet. This meme, as part of the canon, will have high longevity, high fecundity, and simultaneous low *and* high copy-fidelity.

The memetic process is “a product of the human capability to separate ideas into two levels – content and structure” (Rintel, 2013, p. 256), and this separation, echoing copy-the-product and copy-the-instructions dissemination, can illustrate how memes can be both high and low copy-fidelity. The structure of popular memes can give birth to new meme *templates*, which are specific formats used to structure a particular meme. The format or template of a meme, therefore, can be reproduced and disseminated with *high* copy-fidelity, even if the content itself is completely different each time (*low* copy-fidelity). One such structural element, Brideau and Berret (2014) explain, is the font that came to be the ‘default’ font of image macro memes. Image macros were one of the earliest forms of internet memes, which were simply images with a super-imposed text, text that would often, when combined with the image, be strongly representative of a particular emotional state (Rintel, 2013, p. 257).

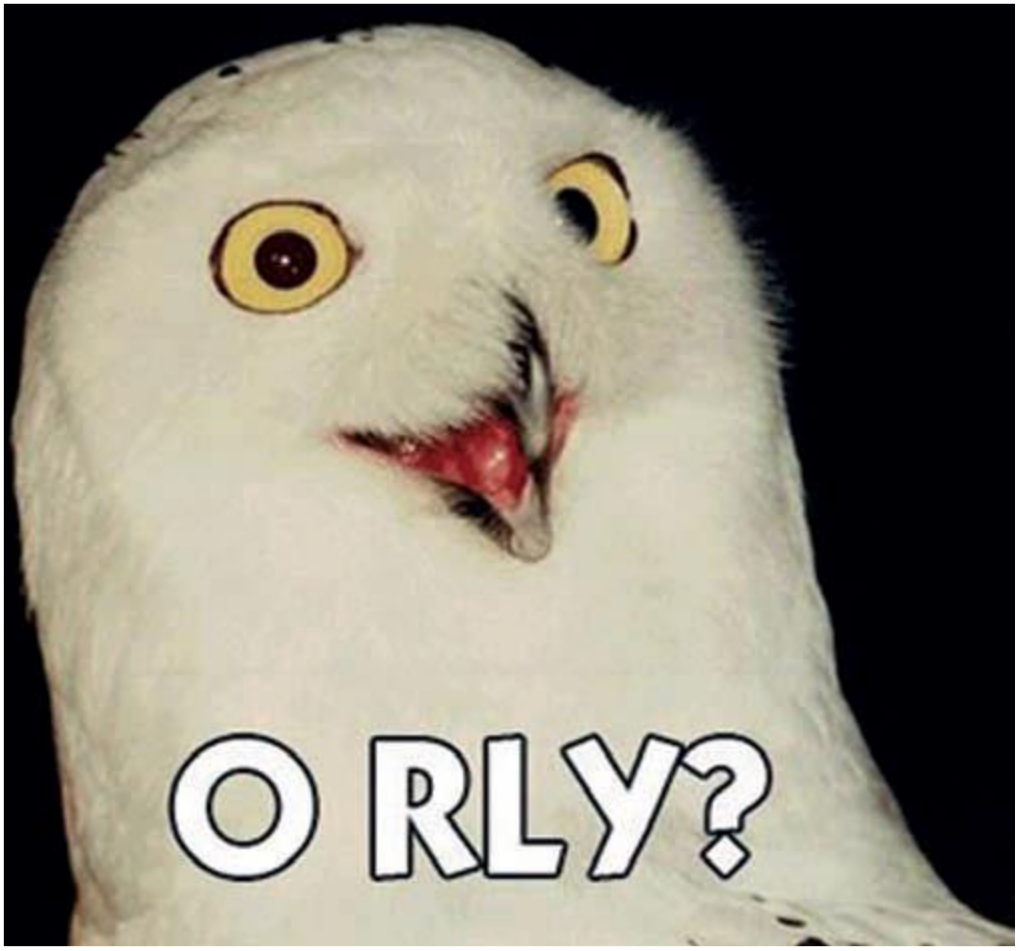


Figure 2.1: ‘O RLY?’, a ‘classic’ image macro meme (Rintel, 2013, p. 257).

Early instances of these memes, however, would not have been referred to as ‘image macros’, because the name came from the programming process behind their eventual standardization.

In computer programming, macro instructions are "scripts that save time and effort for a programmer by replacing a lengthy or repetitive task with a set of defined procedures" (Brideau & Berret, 2014, p. 309). The creation of 'meme generators' was the result of this format standardization. Prior to this, the text could be anywhere on the image and use any font selected by the creator, but after, a certain 'look' of image macro memes came to be, with a more rigid top text/bottom text structure, and a specific font – capitalized **IMPACT** – becoming known as the 'meme font'. The rigidity of the structure of image macro memes facilitated the recognizability of any of these memes of a particular genre as 'a meme'. The stricter text format, as well, could be applied to *any* image, priming viewers to process the piece of content as a meme. Within the high copy-fidelity of structure, meme creators were afforded the opportunity of creating with low copy-fidelity, nevertheless having their creations recognized and remixed within the wider meme genre of image macros.

Figure 2.2: A meme using the 'Most Interesting Man in the World' template, exemplifying the standardized 'look' of image macro memes (Rintel, 2013, p. 259).

Another high copy-fidelity structural element that allowed for low fidelity content creation within a recognizable genre, Douglas (2014) argues, is in the general aesthetic of memes, or more bluntly, how they are "supposed to look like shit" (p. 314). Image macros, described above, are one type of meme most recognizable as 'an internet meme' by mainstream media entities. The other most recognizable type is 'rage comics' and 'rage faces', which perfectly capture the "internet ugly" look that is unique to memes more generally. This important aesthetic characteristic of meme culture, Douglas (2014) explains, is a "celebration of the sloppy and the amateurish" (p. 314). Rage faces are "stick figures, crudely drawn faces, [and] cut-and-paste characters" (p. 317) that would be used by meme creators to share stories by assembling comic strips ('Rage comics') out of available elements. Whether the stories were real or not was irrelevant, what mattered was that they were funny, catchy, shareable, and their elements endlessly remixable, nevertheless maintaining their unique aesthetic.

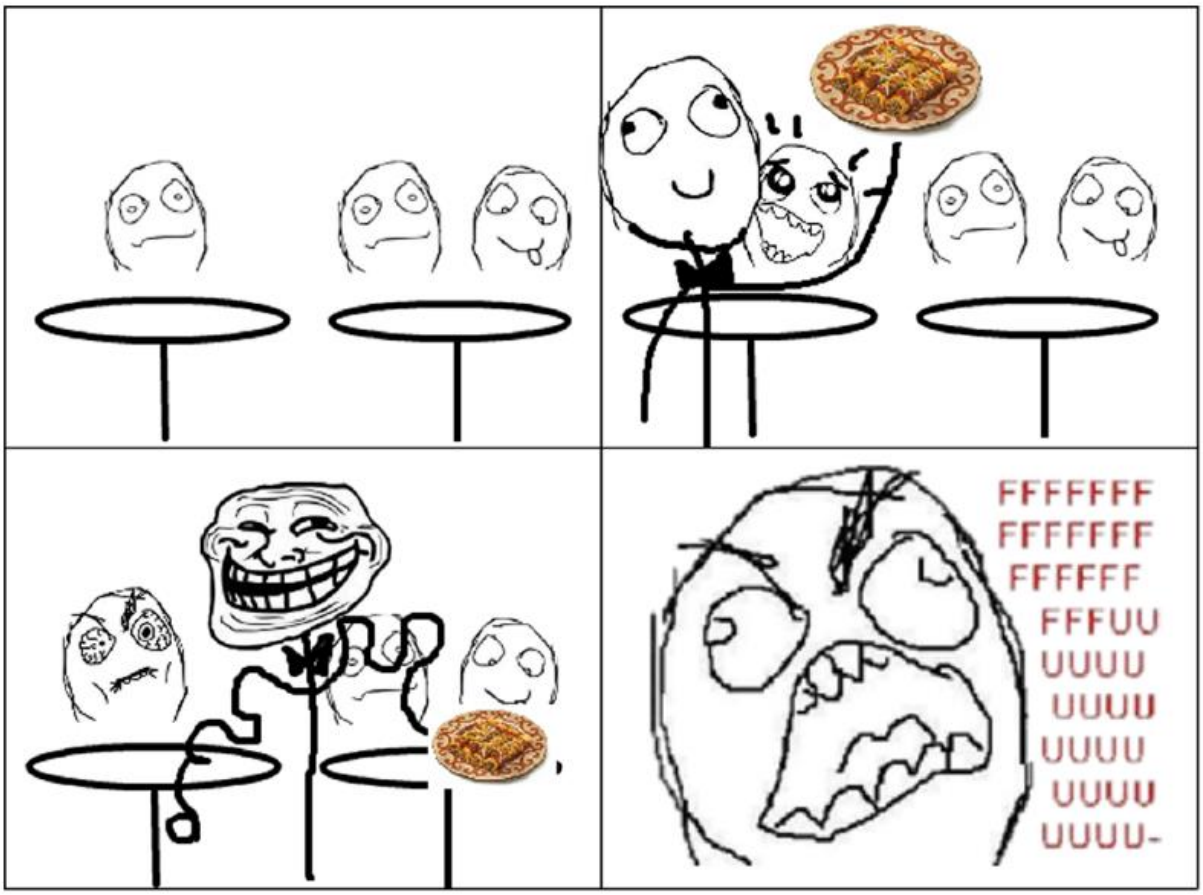


Figure 2.3: 'Waiter Rage', a classic rage comic strip featuring the 'original' rage face (bottom right), from which the genre of memes got its name (Douglas, 2014, p. 318).

The specific look of internet memes in general, as well as their remixability of content within a more rigid templatability of format is what allowed their popularity, proliferation, and growth as a communication medium of the participatory internet. Through their celebration of aesthetic ugliness, as well, memes could be a form of bottom-up folk art (Shifman, 2014, p. 15), framed as existing in opposition to the clean and polished wider popular culture. This opposition may range from a passive one through internet ugly aesthetics to an active one taking the form of culture jamming (Lievrouw, 2011), reappropriating popular culture to create memes that commentate on said culture through the "absurdist' humour" (Knobel & Lankshear, 2007, p. 217) inherent to them. More so than cultural memes, the intertextuality of internet memes is a significantly important characteristic of their creation, how they are understood and interpreted, how they may be used as social commentary (Rintel, 2013, p. 252), and how they enable "potentially transgressive or empowering modes of communication and participation" (Vickery, 2014, p. 302). The internet meme examples given above are some of the most widely recognized genres of memes, but they may fail to properly illustrate the current complexity of online meme culture. Shifman (2014) illustrates the differences of meme complexity and the ease of understanding and creating them in her chapter on *meme genres* (pp. 99-118) by introducing the concept of *literacy* as applied to memes: as she explains, "different meme genres involve different levels of literacy" (p. 100), with some possibly understood and created by "almost anyone", and others requiring "detailed knowledge about a digital meme subculture" (p. 100). In a sense, therefore, the study of internet memes would benefit from being taken beyond the structural limitations of 'classical' memetics, to study the complex referential webs of meanings and possible uses they carry. These memes exist in a state of constant self-referentiality, simultaneously reinforcing and critiquing their own intricacies, in a manner that could then be best described as *meta*-referential.

2.4 Memetic Communities, the Lulz, & Trolling

The meta-referential nature of internet memes implies a certain level of involvement to be able to fully 'understand' them. A certain user involvement in the communities where they proliferate would be necessary for that user to be properly literate in how to use memes. As Knobel (2006) had already explained, the online affinity spaces where memes proliferate are

imbued with a certain aura of *exclusivity.* Those not a part of the group would not 'get' the memes that are popular within that group. Even those in the group could be criticized as not understanding how to use a particular meme template the 'correct' way (Vickery, 2014). A predominant motivating logic behind these meta-memetic communities, in a way building upon the confrontational and fringe nature of these groups, is that of trolling and the 'lulz'.

The lulz are "a spirited but often malevolent brand of humour etymologically derived from lol[3]" (Coleman, 2014, p. 4). The lulz characterise online meme communities in a way that builds upon the rigid in-group/out-group nature that defines them. As explained previously, these affinity spaces have often framed themselves as in some sort of opposition to normalcy. Even as they have grown out of their fringe status of the early internet, the overarching confrontational identity has remained. As such, a certain gatekeeping elitism permeates these communities, and jokes are therefore seen as acceptable to be made at the expense of those not in the group. Taken to its extreme, the confrontational mentality of the lulz becomes what is known as 'trolling'.

While the lulz are characterized by a confrontational yet nonetheless *playful* humorous attitude towards certain people or groups outside of the community, trolling is significantly more mean-spirited, with 'trolls' deriving satisfaction or pleasure from actively offending and hurting others they consider 'lesser'. While there is no concrete way to accurately gauge the demographic percentages of these communities, Milner (2013, p. 70) found that they tended to be predominantly white and male, with whiteness and maleness assumed as the 'default' in the produced memetic content. The offensive content, as such, would tend to skew towards racist, sexist, and homophobic (to name a few) sentiments, directed at the 'other' implicitly assumed to not be part of the community. An important hub for meme emergence and trolling attacks towards other groups is 4chan, an anonymous imageboard forum founded in 2004. The forum is organized along different 'boards' that pertain to a particular kind of content. Its most popular board, however, is /b/ ('Random'), where 'anything goes' and which is often described, including by its members, as the "asshole of the internet" (Coleman, 2014, p. 4). On 4chan, any anonymous user may create a thread by posting an image and, optionally, adding some text next to it. Other users can then contribute to the thread with an image, text, or a combination of both. As a community, it is driven by a strong sense of ephemerality and anonymity. As Bernstein et al. (2011) showed,

[3] 'laugh out loud'

4chan users are not even hiding behind usernames: On 4chan, almost all users are 'anonymous', to the point where a single person could make a whole thread replying to themselves while pretending to be multiple users, and none would be the wiser. The anonymity of 4chan affords its users a sort of unhinged chaos which greatly intensifies its trolling campaigns, and offensive content is rampant on the platform, thanks to its other characteristic, that of ephemerality. On 4chan, nothing is archived; if a thread reaches its post limit or goes too long without a reply (being replaced on the page by active threads), it disappears. Interestingly, the internet ugly aesthetic Douglas (2014) explained can be seen as a side effect of the ephemerality of 4chan; if any user spends too long crafting a meme to respond to a thread, the thread might have been deleted by the time they come back. The average lifespan of a /b/ thread, Bernstein et al. (2011, p. 4) showed, was only 3.9 minutes. These characteristics of 4chan have therefore influenced the overall attitude of its members. Put simply, no individual member can be directly identified, the content is not saved, and the offensive content can be defended as being 'just jokes', so the underlying attitude is one of a 'free-for-all' of offensive content created at the expense of a nebulous 'other'.

Although other meme hub communities exist, the first and largest was /b/, where "memes are violently forged deep within the antagonisms of 4chan's /b/" (Milner, 2013, p. 87). The overarching attitude expressed in most memes that fit the trolling, antagonistic pattern of /b/'s memes, then, would subtly reinforce and replicate the underlying ideological attitudes behind /b/. One such other community is reddit, which despite not benefitting from 4chan's anonymity and ephemerality, nonetheless can reproduce its harmful behaviours. Massanari (2017), examining the site in the wake of the sexist hate/trolling campaigns that emerged out of it, appropriately coined the phrase "toxic technocultures" (p. 329) to describe these hubs, in which harmful and anti-feminist masculinities form and are self-reinforced through ideological group cohesion (Ging, 2019). Reddit users are 'traceable' in a way that 4chan users aren't, but there are also no strictly enforced limits to how many usernames a Reddit user can create, and any user on the platform may create their own self-contained community ('subreddit'), where themes can be anything that isn't explicitly illegal. Milner (2013, p. 88) gives the example of r/Creepshots, a (now banned) subreddit where users would post sexualizing pictures taken of women without their consent. Following the deletion of r/Creepshots for illegal and offensive content, its users created r/CandidFashionPolice, where the content was the exact same, but framed as critiquing the fashion choices of women – in sexualizing pictures taken without their consent.

2.5 Misinformation

Ideological group cohesion was described above as an underlying element of meta-memetic communities, but it is important to note that it is not exclusive to these communities, online. Kata (2012) showed for example that under the participatory paradigm of Web 2.0, there has been growth of the anti-vaccination 'movement', whose members are together ideologically driven in their scientific scepticism, spreading scientific misinformation which can have a very real negative effect on the public's trust in healthcare science. She argues that as the internet grew in popularity and became open to more and more individuals contributing to the wider online discourse, an overarching attitude of everybody being an "expert" (p. 3779) arose. The "connective power of the Internet" (p. 3779) allowed individuals whose beliefs would previously be considered "fringe" (conspiracy theorists, for example) to find each other and form affinity spaces (Knobel, 2006) where their misinformed opinions (cultural memes) would spread. As such, the participatory internet has led to a sort of informational overload, where the "infinite personalized truths" each become "portrayed as legitimate", therefore "weakening messages from qualified experts" (Kata, 2012, p. 3779) by being contrasted to them.

The anti-vaccination movement is the example that Kata (2012) uses, but the information overload inherent to the participatory internet can also be seen as having contributed to the proliferation of fake news, a process in which internet memes can play a part as well. Here, the internet memes described are not necessarily humorous online cultural artefacts of the kind extensively explained previously, but they disseminate information in just the same way. Keeping with the example of the scientific scepticism expressed in anti-vaccination affinity spaces and shared through their memes, an average online user would not necessarily have their opinions of healthcare science changed from a single shared meme, but that single meme may cast *enough* of a doubt in their mind to lead them to search for more information on vaccines, a process which would lead them to anti-vaccination communities - where the ideological strength of a *community* might convince them – before being led to scientific data. This also plays into the conspiratory mindset of believers in the alien abduction memeplex that Blackmore (1999) described: the cultural meme (the information) would not necessarily spread because it is true, but because it is catchy, and (mis)information using 'common-sense language' to discredit experts would be catchier than the dry,

academically presented research of said experts. The communities of healthcare science sceptics, as well, may imbue members with a feeling of being 'in the know' in a way that most people wouldn't, exclusivity also being 'catchy', in a sense. Even if reading of an anti-vaccine website would not motivate someone to join the community, its influence can still have an impact: Research has indeed shown a mere 5 to 10 minutes of viewing such websites measurably increased the perceived risks attached to vaccines in respondents (Kata, 2012, p. 3780). Online misinformation, in a sense, does not work by changing people's mind in a single memetic instance, but works by normalizing the misinformation, conveying it in common sense language, presenting it in a catchy manner, and doing so over a long enough time period that people's minds might be changed with the misinformation becoming the 'norm' of what they see. Furthermore, under Web 2.0's subtle assumption of "everybody [being] an expert" (Kata, 2012, p. 3779), misinformation that is not believed may still lead individuals to be sceptical and critical of *any* information, no matter the legitimacy of its source.

Expanding on what Kata's (2012) research showed, Alvermann (2017) more generally explained how the participatory paradigm normalized by social media made it particularly easy to "surround yourself with others who think like you" (p. 335), an environment in which sharing content in line with those like-minded people "creates the opportunity for an instantaneous positive feedback loop that can perpetuate poor decision making" (p. 335). In this online media environment, she argues, has culminated in our current "post-factual" context, where debate and discussion has become framed by "appeals to emotion" (Alvermann, 2017, p. 336) of like-minded people, rather than revolving around factual information. It is in this environment that fake news proliferates.

2.6 Disinformation, Propaganda, & Radicalization

The previous subsection illustrated the post-factual status of today's online media environment. In such an environment, it has become much easier for members of meta-memetic communities to spread misinformation and offensive content through memes 'for the lulz', but under the more aggressive attitude of trolling, as combined with a growing lack of critical examination of information brought on by an attitude of distrust towards experts and a reinforcement of conspiratory mindsets through like-minded communities, trolls have weaponized spreading misinformation, therefore turning it into *dis*information. Although

both play a role in the spread of fake news in our post-factual era, misinformation is more passive in its un-truth: To refer back to Kata's (2012) example of anti-vaccine communities, members would spread information that is indeed *wrong*, but that they would nevertheless genuinely believe themselves, framing it as common-sense knowledge because to them, it very well may be. Remember Blackmore's (1999) example of fortune-telling, astrology, and alternative medicine (pp. 182-186), explained previously: the misinformation was disseminated *passively* by actors who may have been unaware of the memetic process, and simply have been doing 'what works.' Disinformation, in contrast, is more active in its falseness, and akin to propaganda in its dissemination process: For the case of 4chan trolls spreading fake news, for example, they would do so with the full awareness of the falseness of the information, seeking only to sow chaos for the sake of comedy or 'for the lulz'.

As was previously explained, internet memes may be offensive in nature, and reflect the harmful ideologies of the communities that birthed them, but would remain interpretable as 'just jokes' and satirical to a certain extent. Echoing Blackmore's (1999) point of memetic understanding being greatly influenced by the individual's knowledge and past experiences, however, Johnson, del Rio and Kemmitt (2010) showed that satire can lead to a widely complex range of responses in audiences. They explained that satirical texts – and irony more generally – can be polysemic in its understanding and polyvalent in its interpretation (p. 396). Polysemy "refers to the possibility that a text will have multiple meanings" while polyvalence "to how a text might be understood by viewers in a similar way but evaluated differently" (Johnson, del Rio & Kemmitt, 2010, p. 396). What the authors found was that satire – of the kind that could be applied to offensive internet memes – could act as an ideological reinforcement for those "in the know" (p. 397), even if they missed the satirical intention of the text. In simpler terms, with this idea applied to offensive memes, it would be possible for an offensive meme to have been intended as a joke, but a viewer could 'miss' the satirical intent because they hold the harmful belief the meme satirizes, so the meme would act as ideological reinforcement, and the community reinforcement would have a radicalizing effect of intensifying harmful beliefs.

When writing about cultural memes, Conte (2001, p. 102) illustrated that while memetic agents act in a self-interested manner to achieve their own goals, they will also adapt and adopt the goals of those in their group to achieve said goal, even if they do not believe in the group's goals. This aspect of cultural memetic transmission can be applied when considering the assumed 'default' user of meta-memetic communities. Milner (2013) and

Massanari (2017) together showed that the 'average' user of meme hub communities would be male, straight, and white, expressing prejudice in the offensive content as directed to the most likely 'other' of a straight, white, male audience; women and ethnic minorities. A reasonable assumption to be made, however, is that not all users fit the above default stereotype, though they would most likely not express their 'otherness' within the community unless they feel the need. One must ask themselves, then, how they can be exposed to the offensive content seemingly directed at them without internalizing the negative information in a way that affects their sense of self-worth. There are multiple possibilities for this, but within the context of the present research, the one that will be focused on is the internalized considerations of these offensive memes being 'just jokes', no matter how hurtful.

In contrast to *priming*, another social psychology research field that will be useful here is that of *inoculation* theory. Inoculation can be conceptualized as a kind of memetic (in the cultural, *idea* sense) immunization, almost like a vaccine for the mind. The offensive meme content that 'othered' users would be exposed to, through their active expectation of it being meant as hurtful, will inoculate them against being affected by ideologically similar content in other formats. Roozenbeek et al. (2020) studied how one can be inoculated against being convinced by fake news content, and one method, they found, was for research participants to be familiar with the mechanism and processes through which fake news spread, rather than being familiar with factually correct information to counter the misinformation. This 'prebunking' was indeed effective, but a better method of protecting against fake news, they found, was for one to know how fake news take form and spread. To apply this concept of inoculation to internet memes more generally, remember Shifman's (2014) point on literacy (p. 100): different meme genres/formats will be understood and used by users who are appropriately literate in their usage. If mis/disinformation is presented in a particular meme format, therefore, individuals equipped with the appropriate meme literacy would be inoculated against the offensive/harmful information expressed in the meme. Banas and Rains (2010), in their meta-analysis of research on inoculation theory, found that people with ideologically opposing beliefs to those of the information presented were more likely to not be influenced. This ideological inoculation, when applied to offensive memes, can further illustrate how marginalized (in regards to the content) community members may defend against the harmful messages, but also how those outside the community, lacking literacy to understand the content's satire, may be further radicalized in their beliefs.

The manner through which online users who are more 'in the know', of the tactics employed by trolls in spreading harmful information, can more efficiently reject the harmful message and not let them influence them through a normalization effect, can be useful in the formulation of this research's second hypothesis, aiming to answer its second research question[4]. Similarly to hypothesis 1, the concept of *perceived normality* was split as *higher perceived credibility of information* and *lower perceived harmfulness of information.* Hypothesis 2 can therefore be expressed thusly:

> **<u>Hypothesis 2A</u>: Respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were inoculated against perceiving it as more credible, and therefore considered it less credible, when compared to other groups.**
>
> **<u>Hypothesis 2B</u>: Respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were inoculated against perceiving it as less harmful, and therefore considered it more harmful, when compared to other groups.**

Offensive memes shared by trolls as part of disinformation campaigns could appeal to those who legitimately hold the harmful beliefs, and the communal influence of fringe ideologies found online could then act as a self-reinforcement mechanism of the radicalizing effect of the harmful ideology. Meta-memetic communities like 4chan are not necessarily driven by a particular ideology, but instead by the drive to sow chaos (Merrin, 2019, p. 203) in existing social and informational systems, an attitude evocative of the main motivation behind trolling, as well. Although not politically engaged in a traditional manner other than whichever ideology will be most disruptive in the context of disinformation and trolling campaigns, the political alignment of 4chan trolls has, in recent years, been that of right-wing extremism, reframed as 'alternative right' or 'alt-right', named so after the small subset of right-wing politicians espousing white nationalist rhetoric disguised in 'common sense' and 'intellectualizing' language in order to appeal to a mainstream audience. The alt-right community has in recent years taken advantage of the United States' resurgence of pro-white

[4] To what extent does high internet meme literacy protect viewers from accepting a harmful idea as normal, if initially exposed to it through internet memes?

rhetoric (Hartzell, 2018, p. 6), which seems to have aligned with these meta-memetic communities' 'default' racial identity of 'whiteness', as well as with the harmful racist ideologies expressed in some of their offensive memes. As such, Hartzell (2018) explained, the community of nihilistic trolls of the alt-right has been folded into existing structures of white nationalism and white supremacy. Within these structures, some university circles of like-minded white nationalist/white supremacist individuals had been trying to spread their ideology through linguistic masking of their more extreme views, but they may not have been as efficient in normalizing this kind of language and content as alt-right trolls have been in their disinformation push of offensive content, especially racist content, which has also seemingly acted as ideological reinforcement for isolated like-minded individuals, finding strength in previously fringe groups.

3. METHODS

3.1 Introduction

Based on the examined literature, we can see how, while offensive memes would play a comparatively small role in the normalizing effect of harmful ideologies, they can play *enough* of a role that it becomes important to protect users from that effect. As such, this research aimed to verify whether this normalizing effect would be significant enough to be measurable. The main idea is that respondents exposed to racist memes first were *primed* to perceive later text as 'more normal'. To reiterate, the main research question of this project is as follows:

> **To what extent does the exposure to a harmful idea through internet memes influence the later acceptance of that same idea as normal, in different formats?**

To answer the first research question, two related hypotheses were devised, each pertaining to one of the two variables together deemed representative of *perceived normality* in the context of an experimental survey:

> **H1A: Respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were primed to perceive it as more credible, when compared to other groups.**
>
> **H1B: Respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were primed to perceive it as less harmful, when compared to other groups.**

As what demonstrated from the literature review, however, the extent to which a person is familiar with the structure and 'language' of internet memes would protect them from the normalizing ideological effect of offensive memes. A high meme literacy would therefore *inoculate* respondents against the harmful ideologically normalizing effects the first research question addresses. To test this, the second research question of this project is as follows:

To what extent does high internet meme literacy protect viewers from accepting a harmful idea as normal, if initially exposed to it through internet memes?

To answer it, two related hypotheses were devised, each, again, pertaining to one of the two variables conceptualized as measuring *perceived normality*:

H2A: Respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were inoculated against perceiving it as more credible, and therefore considered it less credible, when compared to other groups.

H2B: Respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were inoculated against perceiving it as less harmful, and therefore considered it more harmful, when compared to other groups.

As was explained in the theory section of this report, plenty of research has been conducted into the radicalizing and propaganda effect that online meta-memetic communities have had. Most of the research, however, has been of a qualitative research design. The focus, as well, has tended to be on the communities, rather than the memes themselves. Although there has been some quantitative research conducted, the general design would seem less prevalent than qualitative in the study context of internet memes and their communities. Huntington's (2019) research explored the effect that memes could have in regards to information processing, in a research design similar to what was chosen here, but her research was specific to political communication, mis/disinformation in fake news memes, and how this information would be processed when compared to non-meme fake news content. Instead of this focus on political content, my research instead approaches the issue of internet memes affecting perception of information through a broader lens, by examining this effect in a wider, ideological context.

To test hypotheses and answer the research questions they relate to, an online experimental survey design was deemed *most* appropriate. There are, however, still limitations to such a design in replicating the conditions under which participants would experience offensive memes and their effects online, but they will be addressed in detail later. This quantitative design was selected for being the "primary tool for studying causal relationships" (Babbie, 2014, p. 255), where manipulating conditions would be possible to

reduce the risk of confounding effects (Neuman, 2014, p. 319). The main experimental goal of the survey had to be initially hidden from participants – although they were debriefed after – to not affect results, so a between-subjects design was chosen; different participants were randomly assigned to different experimental condition groups, and their results were then compared using statistical analysis program SPSS. The survey was distributed online in order to – as much as possible – replicate the conditions under which participants would encounter memes (offensive or otherwise), as it was deemed the context of in-person testing at a university lab would affect results by priming participants to be more mindful of their responses.

3.2 Sampling & Data

As an increasing number of people are beginning to be involved in the participatory internet through the affordances of social media, they are also being potentially exposed to more and more misinformation in the form of memes. The target population of such a study, then, would be anyone on the internet who has some level of active involvement in social media groups.

Within the context of the small exploratory study this research is, however, it was unfortunately not possible to obtain a representative sample of such a population. The experimental survey of this study, hosted on survey platform Qualtrics, was distributed through two posts – made three days apart – on the r/SampleSize subreddit, where users share their research surveys in search for participants. The survey link was also posted to a selection of five Facebook 'survey sharing' groups, where mostly students share their research projects in looking for participants.

Both avenues of posting allowed me to collect data from a wider ideologically driven variety of respondents than if the survey was shared to more specifically left or right-leaning platforms, for instance, but the platforms and groups meant that most respondents were nonetheless of a particular 'type' in a way that meant the sample was not accurately representative of the population. Reddit being a meme hub, respondents obtained there could be reasonably assumed to know more about memes than the average member of the target population. Similarly, the selected Facebook group members tended to be students, who could also be assumed to know more about memes than the average person. As was explained from Milner's (2013), Massanari's (2017), and Ging's (2019) research, the reddit userbase

would tend to be mostly male, white, and straight, so the sample obtained from the site would tend to skew towards those demographics, although the survey being posted on an 'ideologically neutral' subreddit like r/SampleSize would at least minimize this bias *to a certain extent.* The sample obtained from Facebook would be more demographically varied than from reddit, but it could still be assumed that respondents obtained there would be younger than the studied population, considering the selected groups being mostly aimed at students. The sampling process, in conclusion, was to an extent a convenience sample, obtained from online environments most likely to bring in more respondents quicker, at the expense of – potentially- obtaining a representative sample.

Following data collection, a total of 134 complete responses were obtained. More partial responses were available from respondents who quit the survey after the testing phase. These responses could have been used in the statistical analyses, but it was decided not to do that. The consent form told participants that they may withdraw their consent at any time by quitting the survey. It could be reasonably assumed that respondents who quit at 97 or 98%, when being asked about what they thought the study purpose was, could have thought they had reached the end of the survey. Just to be safe from potential ethics concerns, however, these responses were not used.

As allowed by Qualtrics in the settings of survey creation, a 'force response' option was enabled for all sections except for the demographic one. For privacy concerns, it was deemed better to not force respondents to give this sort of information, even when 'Prefer not to answer' was available as an option. Considering the ethics requirement of participants being older than 18, however, respondents who did not give their age were excluded, and their responses deleted from the working dataset. For further analysis, the final sample (N= 130) constituted the usable dataset.

In the final sample (N= 130), 41.5% (54 participants) identified as male, 50.8% (66) identified as female, and 7.7% (10) selected another option when asked about their gender identity. A majority of respondents identified as white (77.7%, 101 participants). The median age of respondents was 23.5 years old. Most respondents (53.1%, 69 participants) self-reported as having *some* undergraduate (Bachelor's Degree) university education (37 participants) or as having *completed* undergraduate university education (32). The next most represented education level was having *completed* postgraduate (Master's Degree and beyond) university education (20.8%, 27 participants). Most respondents identified as left-

wing, politically (69.2%, 90 participants). 16.9% of respondents identified as right-wing (22 participants), 5.4% (7) as neither left nor right leaning (centre), and 8.5% (11) did not answer the political ideology self-identification question.

The final sample was mostly representative of the studied population regarding gender identity, but otherwise, the demographic distribution tended to skew towards a white racial identity, a relatively young age, and at least some university education. The final sample, as such, was unfortunately not representative of the studied population. One way this could be countered, would this research project be repeated, is by attaining participants from a wider and more varied list of online platforms, and by running the data collection phase for longer.

3.3 Structure & Manipulation

In order to obtain the valid results most sought after in quantitative experimental research, the structure of the survey had to be made quite simple, to minimize the risk of confounding variables affecting the results. It was necessary, as well, to give a "cover story" (Geuens & De Pelsmacker, 2017, p. 88) as to the purpose of the research, since an awareness of its experimental nature would influence the results. This also meant that the design had to be a between-subjects one, as a within-subjects design would invalidate the potential effect studied with respondents being aware of the purpose in the second phase of testing. Ethical concerns of respondents being 'tricked' were raised, but participants were debriefed as to the *actual* research purpose when finishing the survey. Geuens and De Pelsmacker (2017) explain that in experimental survey research, it is of particular importance that participants be debriefed as to the true study intention, especially if they have been given "false information" (p. 88). In the context of the present research, a debriefing was also necessary to minimize the risk of participants becoming distressed through exposure to offensive content.

In section (1), respondents were given the cover story, that the survey was meant to study the perception that people of different backgrounds have of offensive internet memes, then were asked to accept the consent form, which emphasized the potential psychological distress that the content shown could cause them. Section (2) contained a standard set of demographic questions; age, highest attained educational degree, country of residence, gender, and racial identity. Section (3) contained a set of questions to gauge the respondent's ideological positioning; first, they were asked to self-identify their political leaning on a scale

from 1 (‘extreme left’) to 11 (‘extreme right’), with 6 representing the centre, based on Kroh’s (2007) conclusion as to the importance of mid-points in political ideology scales. Section (3) also featured a selection of five statements, taken from Manganelli Rattazzi et al.’s (2006) shortened version of the right-wing authoritarianism (RWA) scale, to which respondents had to indicate their level of agreement on a 7-point Likert scale (‘Strongly Agree’ to ‘Strongly Disagree’). Although it would have been interesting to see how a combination of left-right and authoritarian-libertarian respondent placement could have influenced results, time constraints for the survey length meant that it was not possible to use the full scale, or even a full subscale. The statements, then, became part of the cover story given to respondents.

Section (4) was when responses began mattering for the context of the study. To measure respondent internet meme literacy, a selection of four image macro meme images (text-less) was made: Confession Bear, Socially Awkward Penguin, Foul Bachelor Frog, and Overly Attached Girlfriend were selected. Vickery (2014) explained the meaning and intricacy of use of the first three in her research. The fourth was selected due to its large popularity at the time. Three questions were asked about each; the first was about the emotion or attitude that would be best expressed with the template, with four possible answers, only one of which being ‘fully’ incorrect; the second was the likeliness of the respondent finding a meme using that template funny, answered on a 5-point scale (‘Extremely Likely’ to ‘Extremely Unlikely’, with a sixth ‘n/a’ answer if they did not know the meme); and the third was the ease through which they would be able to create a meme using the template the ‘correct’ way, answered on a 5-point scale (‘Extremely Easy’ to ‘Extremely Difficult’, again with a sixth ‘n/a’ option).

Following section (4), a randomizer was applied in Qualtrics, with respondents sorted into four different condition groups for section (5). In condition (1), respondents were shown a selection of five offensive memes (one homophobic, one racist and transphobic, three racist), then a short paragraph of offensive text taken from an ideologically white nationalist article (Spraguer, 2018). The non-racist offensive memes were added as filler items to contribute to the cover story of the research being about *offensive* memes generally, not specifically racist memes. In condition (2), the order was reversed, and respondents saw the text first, then the memes. Those were the testing conditions. Condition (3) respondents were shown only the memes, and in condition (4), only the text. Those were the control conditions. All stimuli in section (5) had the same two questions; how *credible* respondents thought the

information presented was (5-point response scale from ‘Extremely Credible’ to ‘Extremely not Credible’ with a sixth ‘n/a’ option for non-comprehension); and how *harmless/harmful* they thought the content was (on a scale from 1 ‘harmless’ to 11 ‘harmful’).

Section (6) was common across all participants, and three (of the five in the stimuli) memes were shown; these three memes were the racist ones. For each, respondents were asked if they understood the conveyed information, and if they answered no, an explanation of the offensiveness was shown. Respondents were then asked a single question, about what they thought the intent behind that specific memes creation was, with five response options (‘Satire’, ‘Offense for the sake of comedy’, Offense for the sake of hurting others’, ‘Expression of genuine personal beliefs’, and ‘Not sure’). In section (7), the final section, respondents were first asked if they had ‘figured out’ the study purpose, with an optional open text box, then were debriefed by being told the actual experimental purpose of the study, then linked to the websites of two NGOs which aim to combat racism (European Network Against Racism: enar-eu.org) and transphobia (Transgender Europe: tgeu.org), in order to dispel potential harmful information they may have internalized after participating in the study. To maximize response rate and to keep in line with survey length requirements of such a study done in the context of academic student research, the survey was kept to an average length of 10 minutes to be completed. Across the respondent pool, the median time taken to complete it was 9.1 minutes. Although the survey was described here, it can be found in its entirety, as it was presented to respondents, in Appendix A.

3.4 Stimuli

To reflect the literature, it was important to select internet memes that could reasonably be perceived as having been intended to be jokes. As explained in the previous section, the ideological undertones of offensive memes would tend to be reflective of the ideological undertones of the communities out of which they came. Although Geuens and De Pelsmacker (2017, pp. 85-86) explain the value of experimental researchers creating their own stimuli in order to minimize confounding effects, I eventually decided against it. Their work dealt with advertising research, but their insights can still be useful: As they explain, conducting research sometimes benefits from conceptualizing non-existent brands. If real brands are used, study participants may have existing associations with, or feelings towards them, invalidating results and study conclusions. This idea could have been useful here. However, I

eventually decided against it, as creating offensive memes could mean that my own biases and ideology could be subconsciously implemented in them. In a future study featuring an extended version of the study design and a solid pilot testing phase, offensive memes created strictly for the purpose of experimental research could be useful: Future research could, for example, create memes with very specific types of prejudices, as well as different intensities of expression and clarity of said prejudices in the memes. For now, though, offensive memes were sought out on the online platforms where it would be more likely to find them; reddit, 4chan's /pol/ ('Politically incorrect') board, and a few far-right-leaning 'Patriot' Facebook pages, where racist memes tend to be quite present. A large amount of these memes was saved before going through a selection procedure that eventually resulted in the five used in the experiment being selected. Unfortunately, sourcing and referencing were not properly conducted at the time of collection, so there is no way to report from where the memes came specifically. In the context of internet meme research, however, this would not be a problem, as memes proliferate in a manner that would invalidate the need for viewers of them to go to a specific place, online, to be exposed to them. It is indeed true that someone who would not frequent the more offense and dark humour-driven subcommunities of reddit and 4chan would have a *lower* chance of seeing these types of memes, but there cannot be a guarantee that they would *never* see them. It therefore proved useful to not create memes for the purpose of the experiment, and to instead use those 'found in the wild', to reflect what the average online user *could* encounter.

Although the research focused specifically on a single type of prejudice and offensiveness of content – racism – in order to avoid potential confounding effects of the study pertaining to multiple kinds of prejudice, some offensive memes other than racist had to be chosen to contribute to the cover story of the study being about offensive memes generally. The final selection of memes can be found in the survey in appendix A, but two of them can here help in illustrating the *type* of meme content which was sought after to use as stimuli. Figure 3.1 below can help illustrate how offensive humour in memes can easily be perceived as a harmful insult, or a harmless joke. The information conveyed in the meme is quite clear; young people of today, according to the meme creator, abandon religion and embrace homosexuality. This could be interpreted as a homophobic insult, *but* it could also be interpreted as some kind of satirical statement. This meme does well to illustrate how people of different backgrounds and life experiences may receive and interpret it differently. The context in which this meme is found, as well, could prime viewers to interpret it

differently: If found in the comments section of an article on homosexuality being sinful, on a far-right religious blog, for example, the perceived intent behind its creation and the viewer's interpretation of it would be quite different than if found in the comments of an article about Gay Pride and the negative, albeit slowly diminishing protest against it as a social movement. In one instance, this meme would be an insult, in another, a joke that could even be interpreted as positive reinforcement of LGBTQ+ group identity, a once demonized minority group.

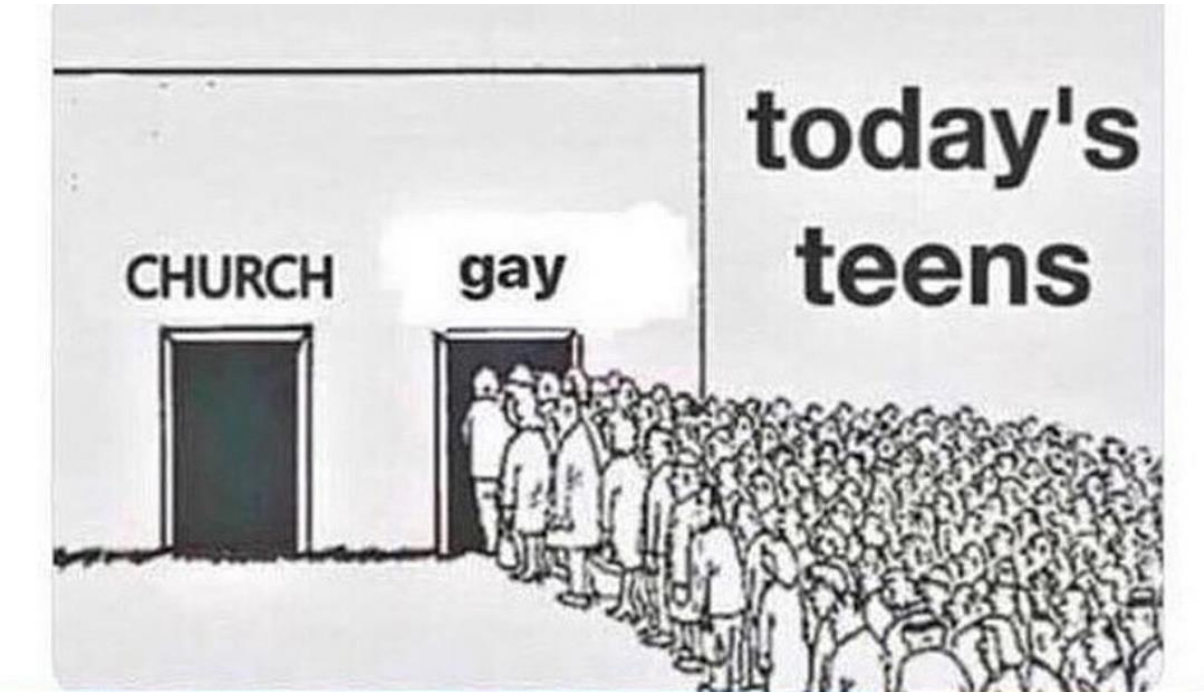


Figure 3.1: A homophobic meme.

The two examples of possible interpretations, expressed above, are quite extreme, but their aim it to illustrate a point explained by Blackmore (1999) about cultural memes, as applied to internet memes, and reinforced by priming theory: understanding and interpretation are highly dependent on framing, the viewer's ideological and intellectual baggage, and cultural contextual elements. The second example, seen in Figure 3.2, can also illustrate how differing interpretations can take place depending on the viewer's knowledge and existing ideology. Here, the comic/meme is shown in its original horizontal format, but in the survey, the panels were rearranged vertically to better fit the screen of respondents participating through their phones. Upon rearranging, as well, it was deemed appropriate to erase the website url, as the creator of *Stonetoss* draws extremely offensive and problematic

comics, expressive of a bigoted extreme-right ideology. Wanting to protect respondents who may have suffered significant psychological and emotional distress visiting that website, I considered it acceptable to erase the url.

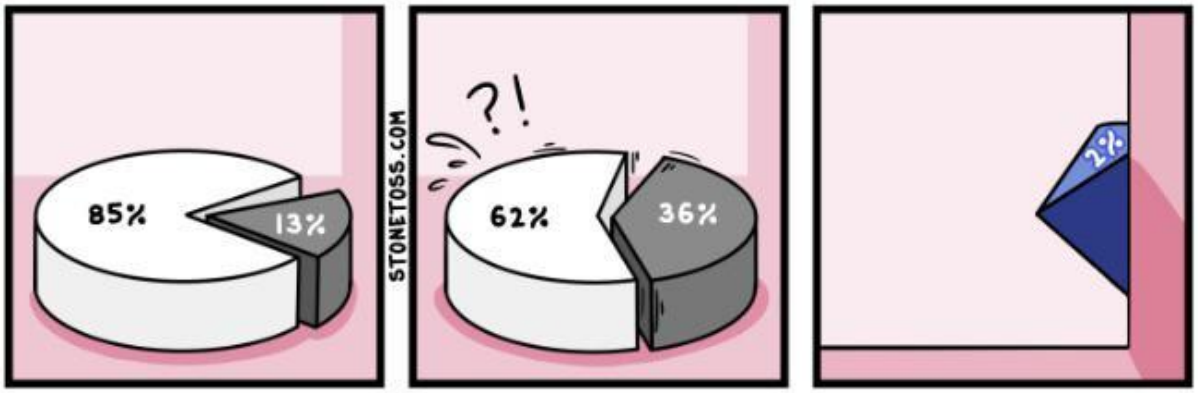


Figure 3.2: A racist and anti-Semitic meme.

The information expressed in this meme is significantly more extreme in its offensiveness than the meme shown in figure 3.1. Here, what is represented is the extreme-right conspiracy theory of 'white genocide', which can be summarised thus: As more people of non-white ethnicities 'mix in' with a predominantly 'white' population, 'whiteness' (an abstract notion of it) becomes reduced, up to the point of disappearance. An element of fascist ideology, 'white genocide' essentially means that as more non-white people reproduce and birth children with white people, whiteness will disappear. Whiteness, therefore, must be 'preserved' and 'maintained' through aggressive anti-immigration policy, for believers of this conspiracy theory. Further adding to the intensity of hate speech expressed, the last panel links to another fascist conspiracy, the 'Jewish conspiracy', which implies that Jewish people secretly control the world from 'behind the scenes'. All in all, this comic conveys that white genocide is orchestrated by the Jewish people who control the world, in order to eradicate whiteness.

This meme is quite extreme, but also quite subtle in the expression of its message, something also seen in the survey results, where 41.5% of respondents (54 participants) selected the n/a option ('I do not understand this meme') when asked about the credibility of information conveyed. This is unfortunate, and indicative of the need for extensive pilot testing, would this research be repeated, but this meme can still help us understand how they can be used for hate speech dissemination understandable only by those 'in the know' of the

extremist communities where this meme could be found. As seen from the majority of respondents not understanding it, the average web user might not understand the information conveyed, but the meme, through its spread, could nonetheless contribute to the ideological group cohesion of the extremist users 'in the know', who would understand it. The hate speech expressed, however, could also be considered as so extreme as to have actually been intended as satire. Following the testing phase, although a majority of my survey respondents, when asked about what they thought the intent behind this meme was (after being told what the conveyed information was), thought it to be 'expression of genuine personal beliefs' (46.1%, 60 participants) or 'offense for the sake of hurting others' (23.1%, 30 participants), some still saw it as satire (8.5%, 11 participants). For this meme specifically, the satirical potential would indeed be difficult to see, but it nonetheless, when considered along with Figure 3.1's potential satire, can be illustrative of the typical defence of 'it's just jokes' that would be used by members of hateful meme communities when they come under criticism. Put simply, hate speech can only be policed as such when it undeniably *is* hate speech, not when doubts can be cast as to the intent behind it.

These two above memes are good illustrators of how offensive memes can be used to disseminate harmful ideologies concisely and efficiently, how they therefore can contribute to online radicalization by acting as group reinforcement for those 'in the know', and how they cannot be easily policed, due to their muddled intent and how it is highly dependent on context. Along with the other three similarly 'typed' memes which can be found in Appendix A, they made up the 'meme' stimuli of this experimental survey. The purpose of this study was specifically related to offensive racist memes affecting processing of hateful text shown later. As such, in order to avoid potential confounding effects from the order of the stimuli memes, they were presented in the same order for every experimental condition.

Originally, when initially conceptualizing the format of this research, the intention was to select a few pieces of text from more mainstream (yet still far-right and expressive of racist ideologies) sources. Under constraints of keeping the survey under a certain length, however, only one piece of text was used. In the end, it was deemed too difficult to find a piece of text that would simultaneously express a racist ideology and be subtle enough to be 'accepted' (in the sense of 'not *actively* rejected') by a more mainstream audience, from a mainstream source. After finding and reading Hartzell's (2018) paper on the 'intellectualizing' of white nationalism and of the kind of information that would eventually be shared as memes by far-right trolls in the 'alt-right', it was deemed appropriate for a short

piece of text to be selected from that platform, alternativeright.com. The source of the text was not given in the survey, but the source itself was considered appropriately 'niche' and 'fringe' for the text (and therefore, source) to not be recognized by respondents, something that could have potentially led to confounding effects and an invalidation of results. The short text paragraph that was selected can be found below. It also expresses, at its core, a belief in the 'white genocide' conspiracy theory, but does so more subtly, masking the prejudice behind an aura of intellectualism and 'common sense' language, in order to appeal to a more mainstream audience and to enact a normalizing effect of this type of ideology, to later make more extreme and extremely-expressed forms of prejudice more 'palatable' and easily accepted. In a sense, what my research attempts to demonstrate with offensive memes priming respondents to be accepting of textual stimuli such as this text, this text seems to attempt to do as well, priming readers to be more accepting of more extreme and radicalizing forms of prejudice.

> 'One of the chief forms of racial discrimination, for example, is exclusion. However, a total lack of exclusion/separation is an invitation to miscegenation, to cultural disintegration via borrowing and mixing, to the rewriting and rereading of one's history, to the complete loss of identity, etc. Without any barriers to entry into a group, the group simply can not protect itself even marginally (or protect its ways, customs, or genes), let alone survive as a distinct group. It will assimilate the surrounding world or be assimilated into it. In this way, open borders combined with comprehensive anti-discrimination laws, is not merely bad policy, it is implicitly/effectively genocidal.' (Spraguer, 2018)

In the experiment, each stimuli were shown on their own (a single meme at a time) with all three questions appearing on the page at the same time as the stimuli, so respondents could consider all simultaneously, and not have their opinion (and answer) changed by the appearance of other questions. It was for this reason, as well, that the questions remained the same for *all* stimuli, whether memes or text. This consistency would minimize potential confusion on the part of the respondents, something that could have happened if each type of stimuli had different questions, which could have muddled the results.

3.5 Operationalization, Reliability, & Validity

This research project is very exploratory in nature, so the conceptual scales and measures would greatly benefit from being further developed, tested in a pilot study, and validated

across a larger respondent pool. Although issues arose in this project, the design itself could be useful, if further developed. The meme stimuli that were chosen, having been found online and therefore already existing, could also be encountered by respondents in their day-to-day web activities, so in the regards to external validity, the memes chosen could replicate a respondent's experience in a non-experimental context. The timeframe of such a study, however, meant that exposure to offensive content – over a long period of time in non-experimental contexts – had to be quite compressed. The effect studied would not necessarily happen in the same way over a long time where subtle normalization effects might take place as it would over a compressed experimental context, so a more developed expansion of this study could for instance study this effect over multiple experimental sessions spread out over a longer time. The external validity of this study, therefore, would unfortunately be low, although the design itself, if expanded, could yield high validity results.

3.5.1 Meta-Memetic Literacy

The measure of meta-memetic literacy of respondents was particularly difficult to conceptualize. There are no existing scales measuring one's literacy in the language of internet memes. I therefore had to create one for this project, which could potentially measure how knowledgeable respondents were in internet memes, but that would remain short enough as to not inflate the length of the survey. Under the restrictive conditions of this research, some issues arose in the development of such a scale. The first issue encountered was the selection of memes. The four image macros chosen were especially old by today's internet meme standards, but they were nevertheless chosen for a specific reason. More recent memes can be highly versatile in their use, and do not adhere to the rigid top-text/bottom-text format of older image macro memes with established templates, meaning the selected memes, as well, could be relatively recognizable by the general public as 'memes'. Still, though, the response format was such that the correct response seemed – in most cases – too obvious to extract a meaningful measure of meme literacy from the 4-choice question regarding meaning. In most cases, simply knowing the name of the template would tell respondents what the correct answer is. In addition, upon performing reliability testing on these measures in SPSS, the Cronbach's alpha check indicated a low reliability score of .413. As such, this item was abandoned as a measure of meta-memetic literacy.

The second issue was that, to measure meta-memetic literacy, it was quickly realized that perceived humour would also not be a good measure, and it was not even tested for reliability. Respondents could, after all, be completely familiar with a particular meme template, but considering the subjectivity of humour, they could also not find it funny at all. This confounding effect was deemed too unclear to be used as basis for analysis. The ease of use item, however, seemed to be a valid indicator of one's literacy in the usage of meme templates. As Vickery (2014) has nevertheless extensively explained, even well-established templates can see some disagreement over their 'correct' usage, in meta-memetic communities, but out of all the survey measures, this one was deemed the *most* appropriate for measuring literacy within the exploratory context of this research. The scale had a relatively high level of internal consistency, as determined by a Cronbach's alpha of .745.

3.5.2 Perceived Credibility & Harmfulness

Another difficulty encountered during the building of the survey was coming up with questions that could be asked about both memes and text, and which could somewhat accurately measure the extent to which one considered the text paragraph information 'normal', without seeming nonsensical when asked about the memes. The way that was deemed most appropriate under the circumstances was to ask about credibility of information and about harmfulness. A measure of credibility could be best applied to text, while a measure of harmfulness could be best applied to memes, with both being usable for all stimuli. Those measures, as well, build on the literature.

Cultural memes, as explained by Dawkins (1976) and Blackmore (1999), tend to spread not because they are true, but because they are catchy. The same could be said for internet memes, with 'catchiness' being reasonably substituted with 'credibility'; if a meme is spreading some kind of information, the more credibly presented this information is, the catchier the meme could seen as being. As explored by the many internet meme researchers whose work was explored above, somewhat offensive memes tend to be perceived as the 'norm' within meta-memetic communities, but those in the know can also understand them to be some form of satire, or 'simply jokes'. As such, this measure can nicely tie in with the meta-memetic literacy measure. Perceived information credibility could relate to the manner in which misinformation and disinformation spreads in memes, as well: Those 'in the know' (meta-memetically literate) could see a meme which expresses some kind of false

information, but these users' knowledge of the underlying ideology expressed in memes would allow them to recognize the disinformation trolling, while a non-literate web user might not perceive the 'humour' intended. Asking about credibility, therefore, seemed valuable as one of the two questions asked about each stimuli. Perceived harmfulness is similarly valuable as being somewhat representative of perceived normality, in the context of this study. Those 'in the know' could see an offensive meme, one that conveys extreme rhetorical prejudice, but they would be familiar with the underlying ideology of the community where the meme came from, so they could be more likely to see the 'joke' as equally harmless as most other jokes in memes of that kind. Those not in the know, in contrast, would not have the experience of the general offense and dark humor of these meme communities, and therefore, the information expressed would seem as more extreme or more offensive, and therefore more harmful.

The overall length of the survey had to be kept quite short to reduce respondent dropout rate, and the overall design was already relatively intricate for such a short experiment, so concessions had to be made. Number of items to measure literacy was reduced, as well as number of stimuli and amount of questions regarding each. As such, it was not possible to have multi-item scales for measuring both perceived credibility and perceived harmfulness. In a longer experimental context, these measures would ideally be expanded. Overall, however, perceived credibility of information and perceived harmfulness were considerate representative enough of the concept of perceived normality, in such a context, and were therefore used.

4. RESULTS

4.1 Tests, Results, & Statistical Significance

Following data cleaning, new variables were computed to be used in statistical analyses. Instead of four different variables indicating to which condition each respondent was assigned, a single new 'condition' variable was created by using the 'recode into same variables' SPSS option, with four values labelled (1) to (4). Across all conditions, perceived credibility and perceived harmfulness scores were summed into single variables, for each meme stimuli, and for the text stimuli. The 'n/a' options values were excluded as to not skew the new variables' values.

H1A: Respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were primed to perceive it as more credible, when compared to other groups.

To test hypothesis 1A, a one-way ANOVA test was conducted to determine if *perceived text credibility* scores varied between condition groups (1) (shown meme stimuli *then* text stimuli), (2) (shown text stimuli *then* meme stimuli), and (4) (shown *only* text stimuli). Results were not statistically significant (F (2, 94) = 0.99, p = .373). In spite of this, a consideration was made that non-significance could have been due to small sample sizes (n = 36 in group (1); n = 29 in group (2); n = 32 in group (4)), and mean credibility scores were informally considered. There was, however, more of a difference between groups (2) (M = 3.72, SD = 1.13) and (4) (M = 3.37, SD = 1.18) than between groups (1) (M = 3.78, SD = 1.40) and (2) (M = 3.72, SD = 1.13), expected to have the biggest difference in scores due to stimuli being shown in opposite orders. As such, these results were determined to most likely be due to random chance.

As another way to test h1a, condition groups (2) and (4) were artificially combined, since both were first exposed to the text stimuli *first*. With only two groups being now compared, an independent samples t-test was selected instead of an ANOVA, comparing *perceived text credibility* scores between groups (1) (n = 36) and (2+4) (n = 61). Results were, again, not statistically significant, with p = .090 for Levene's Test for Equality of Variances and t (95) = 0.90, p = .371 for 2-tailed t-test significance, but an informal

examination of mean scores showed a slightly more pronounced difference between groups (1) (M = 3.78, SD = 1.40) and (2+4) (M = 3.54, SD = 1.16), with respondents having been exposed to text first perceiving it as more credible ('1'= high credibility; '5'= low credibility). Nevertheless, in the context of this research, hypothesis 1A was rejected, meaning that **respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were <u>not</u> primed to perceive it as more credible, when compared to other groups.**

<u>H1B</u>: Respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were primed to perceive it as less harmful, when compared to other groups.

A similar process as for hypothesis 1A was conducted to test hypothesis 1B, instead looking at differences in *perceived text harmfulness* between groups. A one way ANOVA was conducted and results were not statistically significant (F (2, 85) = 0.78, p = .462), although informal examination of the scores showed than respondents who were shown meme stimuli first (group (1); M = 7.55, SD = 3.39) perceived the text as less harmful ('1' = harmless; '11' = harmful) than those shown text first (group (2), M = 8.31, SD = 2.31), but the text-only control group's score (group (4); M = 7.41, SD = 2.61) again tells us the differences in score were likely due to random chance.

Groups (2) and (4) were again artificially combined and an independent samples t-test performed, and while Levene's Test for Equality of Variances showed statistical significance (p = .010), the t-test did not (t (52.811) = -0.43, p = .670, 2-tailed), with differences in scores being minimal (group (1), M = 7.55, SD = 3.39; group (2), M = 7.84, SD = 2.49) and likely due to chance. Hypothesis 1B, therefore, can also be rejected, meaning that **respondents who were exposed to ideologically racist text *after* being exposed to ideologically racist internet memes were <u>not</u> primed to perceive it as less harmful, when compared to other groups.**

<u>H2A</u>: Respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist memes

were inoculated against perceiving it as more credible, and therefore considered it less credible, when compared to other groups.

To test hypothesis 2, a Univariate Analysis of Variance was performed, looking at interaction effects of two independent variables – *condition* and *meme_literacy* – on a single dependent variable. For hypothesis 2A, that dependent variable was *perceived credibility of information*. The purpose of this test was to see whether different levels of internet meme literacy could influence the influencing effect that would arise differently under each experimental condition, based on the order of presented stimuli. Looking at the output for Tests of Between-Subjects Effects, the *p* value of the interaction of *literacy* and *condition* meant the results did not have statistical significance (F (2, 87) = 0.95, p = .390), although the value was still lower than that of the *condition* variable (F (2, 87) = 0.62, p = .541) and the *literacy* variable (F (1, 87) = 0.004, p = .952).

As an alternative manner to test hypothesis 2A, a median split was performed on the *meme_literacy* variable, to obtain two groups (*high* literacy and *low* literacy), and a Univariate Analysis of Variance was again performed to test for the same interaction effect. The obtained *p* values indicated – again – no statistical significance. The interaction effect between *condition* and *meme literacy category* on *text credibility* (F (2, 87) = 0.64, p = .527) strangely increased following the median split, with the new categorical variable (F (1, 87) = 2.06, p = .286) strangely having a much lower *p* value than its continuous instance above. The *p* value of the *condition* variable remained similar as to previously (F (2, 87) = 0.69, p = .590). Hypothesis 2A can therefore be rejected, meaning that **respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist memes were <u>not</u> inoculated against perceiving it as more credible, and therefore considered it less credible, when compared to other groups.**

<u>H2B</u>: Respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist memes were inoculated against perceiving it as less harmful, and therefore considered it more harmful, when compared to other groups.

Testing hypothesis 2B required a similar process as for testing 2A, albeit with a different dependent variable; here *perceived harmfulness of information* in the text stimuli. Results were – again, not statistically significant when looking at the interaction effects between

literacy and *condition* (F (2, 78) = 0.15, p = .856) on influencing *perceived harmfulness*, with *literacy* (F (2, 78) = 0.001, p = .980) and *condition* (F (2, 78) = 0.02, p = .975) showing similarly non-significant values. Results of the same test following a median split were similarly nont statistically significant (F (2, 78) = 0.03, p = .969 for the interaction effect. This therefore means that hypothesis 2B can be rejected as well, meaning that **respondents with high internet meme literacy who were exposed to ideologically racist text *after* being exposed to ideologically racist memes were <u>not</u> inoculated against perceiving it as less harmful, and therefore considered it more harmful, when compared to other groups.**

4.2 <u>Discussion</u>

Hypothesis 1 testing did not show statistical significance in both of its iterations, each pertaining to one of the two concepts deemed representative of *perceived normality* in the context of this experiment. It could be that there genuinely would not be any difference between experimental groups who were each shown stimuli in different orders, but based on the literature and study design, a reasonable assumption could be made that there would be statistically significant results under different circumstances. It could be that the final sample was too small for there to be any statistical significance in performed tests: Indeed, the full sample was quite sample to have been representative of a population within this study's exploratory design. Additionally, condition groups each could only be a quarter (approximately, as per the Qualtrics randomizer) of that already small sample. Unfortunately, the study design and its main focus on stimuli *order* necessitated that there would also be two control groups instead of one, greatly reducing the possibilities for performing tests that would yield significant results. Would there have been a large respondent pool with – still – no statistical significance, another possible explanation comes to mind: It could be that the single piece of text used as part of the stimuli might have been simply 'too extreme' for any normalizing effect to have been measurable on such a compressed experimental scale. If future research reproduces/adapts the present design, an important addition to potentially yield good results would be to expand the number of stimuli greatly. This way, if one text paragraph is too extreme, this would not invalidate the whole study, as there would be more stimuli to measure processing influence on. Alternatively, if only a single textual stimuli

would still be used, selection could be made of something more nuanced and subtle in its expressed prejudices.

Hypothesis 2 testing also did not show statistical significance and was thus rejected. In a similar way as for concerning hypothesis 1, the lack of statistical significance could be assumed to not be a genuine lack of effects, as based on theory, but to be the small sample, the compressed nature of the survey, or previously not considered confounding effects. For the case of the second research question and second hypothesis more specifically, as well, it could have been that the conceptualized measure of meta-memetic literacy was not a valid or appropriate one, further illustrating the benefits of repeating this study design on a larger timeframe with more extensively researched and validated scales for measuring internet meme literacy.

5. CONCLUSION

5.1 Implications for Theory

As none of the tests performed in the context of analysing this research's data were statistically significant, it would prove difficult to concretely illustrate and explain implications for the theory. Quite a few unpredicted issues arose when this project entered its analysis phase. Based on the extensiveness of the theoretical background, it was thought that there would at least be *some* validity in results. One limitation that came to be realized will be explained here, as it relates to one of the research papers used in the theory section. Out of all the existing academic research collected and used in this project, Huntington's (2019) seemed the most similar in her research design to mine.

Her research yielded results, something that surely is also due to the larger scale of her study, but one problem I had was wanting to be *too* broad in my exploration, in a way that in retrospect seems the lack the specificity required of quantitative research. I considered – at the time of conceptualizing this research – that a focus on racism, a *specific* kind of prejudice, would combat the issues caused by a more general focus on ideology, but that wasn't the case. Huntington's (2019) research, in contrast, was extremely specific in its focus on internet memes. Unlike me and my unreasonably broad focus on offensive and racist memes (something that can take a myriad forms in that kind of format), Huntington's (2019) was on political memes and fake news *specifically*. A similarly fine-tuned focus could have yielded better results, although they may have still not achieved statistical significance due to my small sample.

5.2 Answering Research Questions

To what extent does the exposure to a harmful idea through internet memes influence the later acceptance of that same idea as normal, in different formats?

Within the small scale and exploratory context of this research, offensive internet memes do not enact a normalizing effect on the later perception and processing of offensive text.

To what extent does high internet meme literacy protect viewers from accepting a harmful idea as normal, if initially exposed to it through internet memes?

Within the small scale and exploratory nature of this research, high internet meme literacy was not found to protect viewers from the harmful normalizing effect explored in the previous research question.

5.3 Limitations

The biggest limitations of this project were two-fold: The highly exploratory nature of the research, combined with the limited scope afforded by its student-research context, meant that whatever results were found could not be applicable to any sort of wider population. The small size of the sample, and the even smaller experimental groups, was a major difficulty in the obtaining of valid results. The utility of this project, ultimately, was to conceptualize a study design that future research could expand and build upon. In previous sections, limitations of this research were mentioned and explained as they became topical, but the one that seems to have been the biggest is the compression of such a complex cultural influence as ideological normalization/radicalization through memes into such a small, 10-minute long experimental survey. Qualitative research on radicalization through the internet, of which offensive internet memes only played a small part, would not be able to 'prevent' the damage done, but with the benefit of hindsight, it could at least consider and study a multitude of influences that radicalized individuals may have been subjected to.

Here, the research intention ultimately proved to be too broad, attempting to reducing what would be months or years of exposure to mis/disinformation and harmful ideologies, online, of which offensive memes would only play a small part. As such, although the usage as stimuli of memes found 'in the wild' and the experimental survey taking place online could have to a certain extent replicated the conditions of encountering these stimuli in day to day internet activity, the short allotted length meant it was not possible to have filler items of the non-offensive kind, something that may have helped in masking the true purpose of the study. Although respondents wouldn't have necessarily guessed the true purpose to a *full* extent, the cover story given may have still been too close to the real research goal to act as *enough* of a cover story. Knowing they were being tested for perception of offensive memes could have influenced respondents to be more mindful and considerate of their answers in a

manner that could have also invalidated results, had there been statistical significance. By knowing the survey dealt with prejudice, respondents could have subconsciously prepared to answer questions on prejudice, not necessarily by specifically guessing the testing was about racism, but statistically significant results may have still been negatively affected by the confounding effect of the cover story and the true purpose being so close to one another.

5.4 Suggestions for Future Research

There are a few things that could be done prior to potentially replicating this project, the most important one being the development and validation of a proper scale of measurement of meta-memetic literacy. This would require multiple stages of pilot testing. Potentially, there would need to be a few different scales measuring literacy in certain categories or genres of memes, since they can be so intricate in their language. With such a scale development done, there could also be further research into what type of meme literacy can have an effect over what other type of content. The scale items used here were, as mentioned, quite old by the standards of internet memes, but current meme culture can be so complex to almost require the development of multiple literacy scales for different types of memes, or different communities out of which they emerge. Such a process might need to start at a qualitative level, with interview or focus groups with respondents who self-identify as 'knowing a lot' about internet memes. Over time, comparaisons and links could be drawn between responses, eventually transitioning development to a quantitative design for testing and validation.

Future studies with my design would benefit a lot from having a complete (even if shortened) scales to measure ideological positioning, both on horizontal left-right and vertical authoritarian-libertarian axes. This way, a respondent's ideological positioning would be more accurately devised, from existing validated scales that would remove the negative effects that could arise out of a fully self-identification placements along a single axis. As was already explained, the stimuli could benefit from being devised specifically for this research, although that process would need to be a collaborative one to minimize the risk of researcher subconsciously implementing their own internalized biases. This way, stimuli could be organized along different intensities of the offensive information expressed, and different levels of subtlety, to see how these different levels could potentially prime/inoculate respondents to consider later stimuli differently.

The most important suggestion that could be given as to replicate this research in the future, ultimately, would be to greatly expend the design. Individuals who have been radicalized through memes have not done so through a single 10-minute survey. As such, a more ideal version of this research would take place over multiple months, with enough time between experimental sessions for respondents to (possibly) internalize information they've been given. This method would give more nuanced results than the compressed design of this research would have, had results been statistically significant. Concerning this, as well, an important suggestion for future research would be to have many more respondents than I did, enough as to obtain a sample that would be more representative of the studied population.

REFERENCES

Albrecht, S., Fielitz, M. & Thurston, N. (2019). Introduction. In M. Fielitz & N. Thurston (Eds) (2019). *Post-digital cultures of the far right: Online actions and offline consequences in Europe and the US,* pp. 7-22. Bielefeld, Germany: Transcript Verlag.

Alvermann, D. E. (2017). Social media texts and critical inquiry in a post-factual era. *Journal of Adolescent & Adult Literacy, 61*(3), 335-338.

Banas, J. A. & Rains, S. A. (2010). A meta-analysis of research on inoculation theory. *Communication Monographs, 77*(3), 281-311.

Bernstein, M. S., Monroy-Hernández, A., Harry, D., André, P., Panovich, K. & Vargas, G. (2011). 4chan and /b/: An analysis of anonymity and ephemerality in a large online community. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 50-57.

Babbie, E. (2014). *The basics of social research, 6th edition.* Boston, MA: Wadsworth CENGAGE Learning.

Benkler, Y., Faris, R. & Roberts, H. (2018). The propaganda pipeline: Hacking the core from the periphery. In *Network propaganda: Manipulation, disinformation, and radicalization in American politics,* pp. 225-233. Oxford, UK: Oxford University Press.

Bertazolli, A. (2019). Russian resistance and propaganda through memes in the 2010s. In *Internet memes and society: Social, cultural, and political contexts,* pp. 54-184. New York, USA / London, UK: Routledge.

Blackmore, S. (1999). *The meme machine.* Oxford, UK: Oxford University Press.

Brideau, K. & Berret, C. (2014). A brief introduction to Impact: 'The meme font'. *Journal of Visual Culture, 13*(3), 307-313.

Christensen, K., Spahiu, A., Wilson, B. & Duval, R. D. (2015). Mobilizing on the fringe: Domestic extremists and antisocial networking. In P. A. Simpson & H. Druxes (Eds) (2015). *Digital media strategies of the far-right in Europe and the United States,* pp. 55-83. Lanham, MD, USA: Lexington Books.

Coleman, G. (2014). *Hacker, hoaxer, whistleblower, spy. The many faces of Anonymous.* New York, USA / London, UK: Verso.

Conte, R. (2001). Memes through (social) minds. In R. Aunger (Ed.). *Darwinizing culture: The status of memetics as a social science,* pp. 83-120. Oxford, UK: Oxford University Press.

Dawkins, R. (1976, 2006). Memes: The new replicators. In R. Dawkins, *The selfish gene, 3rd Ed.* pp. 189-201. Oxford, UK: Oxford University Press.

Dawkins, R. (1999). Foreword. In S. Blackmore (1999). *The meme machine,* pp. vii-xvii. Oxford: Oxford University Press.

Dillman Carpentier, F. R., Roskos-Ewoldsen, D. R. & Roskos-Ewoldsen, B. B. (2008). A test of the network models of political priming. *Media Psychology, 11*(2), 186-206.

Douglas, N. (2014). It's supposed to look like shit: The internet ugly aesthetic. *Journal of Visual Culture, 13*(3), 314-339.

Drakett, J., Rickett, B., Day, K. & Milnes, K. (2018). Old jokes, new media – Online sexism and constructions of gender in Internet memes. *Feminism & Psychology, 28*(1), 109-127.

Gal, N., Shifman, L. & Kampf, Z. (2016). "It Gets Better": Internet memes and the construction of collective identity. *new media & society, 18*(8), 1698-1714.

Gerend, M. A. & Sias, T. (2009). Message framing and color priming: How subtle threat cues affect persuasion. *Journal of Experimental Social Psychology, 45,* 999-1002.

Geuens, M. & De Pelsmacker, P. (2017). Planning and conducting experimental advertising research and questionnaire design. *Journal of Advertising, 46*(1), 83-100.

Ging, D. (2019). Alphas, betas, and incels: Theorizing the masculinities of the manosphere. *Men and Masculinities, 22*(4), 638-657.

Hartzell, S. L. (2018). Alt-white: Conceptualizing the "Alt-Right" as a rhetorical bridge between white nationalism and mainstream public discourse. *Journal of Contemporary Rhetoric, 8*(1/2), 6-25.

Huntington, H. E. (2019). Partisan cues and internet memes: Early evidence for motivated skepticism in audience message processing of spreadable political media. *Atlantic Journal of Communication,* 1-15.

Johnson, A., del Rio, E. & Kemmitt, A. (2010). Missing the joke: A reception analysis of satirical texts. *Communication, Culture & Critique, 3,* 396-415.

Kata, A. (2012). Anti-vaccine activists, Web 2.0, and the postmodern paradigm – An overview of tactics and tropes used online by the anti-vaccination movement. *Vaccine, 30,* 3778-3789.

Knobel, M. (2006). Memes and affinity spaces: Some implications for policy and digital divides in education. *E-Learning, 3*(3), 411-427.

Knobel, M. & Lankshear, C. (2007). Online memes, affinities, and cultural production. In M. Knobel & C. Lankshear (Eds.) (2007). *A new literacies sampler,* pp. 199-228. New York, USA: Peter Lang Publishing.

Kroh, M. (2007). Measuring left-right political orientation: The choice of response format. *Public Opinion Quarterly, 71*(2), 204-220.

Lievrouw, L. A. (2011). Chapter 3: Monkeywrenching the media machine – Culture jamming. In *Alternative and activist new media,* pp. 72-97. Malden, MA, USA: Polity Press.

Manganelli Rattazzi, A. M., Bobbio, A. & Canova, L. (2007). A short version of the Right-Wing Authoritarianism (RWA) scale. *Personality and Individual Differences, 43,* 1223-1234.

Massanari, A. (2017). #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *new media & society, 19*(3), 329-346.

Merrin, W. (2019). Chapter 13: President troll: Trump, 4chan and memetic warfare. In C. Happer, A. Hoskins & W. Merrin (Eds.) (2019). *Trump's media war,* pp. 201-226. London: Palgrave Macmillan.

Milner, R. M. (2013). Hacking the social: Internet memes, identity antagonism, and the logic of lulz. *The Fibreculture Journal, 22,* 62-92.

Moreno-Almeida, C. (2020). Memes as snapshots of participation: The roles of digital amateur activists in authoritarian regimes. *new media & society, 00*(0), 1-22.

Neuman, W. L. (2014). *Social research methods: Qualitative and quantitative approaches, 7th edition.* Harlow, UK: Pearson Education Limited.

Rintel, S. (2013). Crisis memes: The importance of templatability to internet culture and freedom of expression. *Australasian Journal of Popular Culture, 2*(2), 253-271.

Roozenbeek, J., van der Linden, S. & Nygren, T. (2020). Prebunking interventions based on "inoculation" theory can reduce susceptibility to misinformation across cultures. *The Harvard Kennedy School Misinformation Review, 1*(2), 1-23.

Shifman, L. (2014). *Memes in digital culture.* Cambridge, MA, USA: The MIT Press.

Smith, C. A. (2019). Weaponized iconoclasm in Internet memes featuring the expression 'Fake News'. *Discourse & Communication, 13*(3), 303-319.

Spraguer, E. (2018, April 28th). Anti-racism is immoral. *AltRight.com*. Retrieved from https://altright.com/2018/04/28/anti-racism-is-immoral/

Srull, T. K. & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37,* 1660-1672.

Vickery, J. R. (2014). The curious case of Confession Bear: The reappropriation of online macro-image memes. *Information, Communication & Society, 17*(3), 301-325.

Wiggins, B. E. & Bowers, G. B. (2015). Memes as genre: A structurational analysis of the memescape. *New media and society, 17*(11), 1886-1906.

Williams, A., Oliver, C., Aumer, K. & Meyers, C. (2016). Racial microaggressions and perceptions of Internet memes. *Computers in Human Behavior, 63,* 424-432.

APPENDIX A – Qualtrics survey

Start of Block: 1) Info, consent

info1 Welcome!

Thank you for taking the time to fill out this survey. Completing it should take you no more than 10 minutes.

The purpose of my research is to examine how people from different backgrounds perceive and interpret potentially offensive, 'edgy' internet memes.

Please be aware that you may be shown content of a shocking nature.

This survey *does not* contain shocking ***imagery***, but the *textual* content may be considered offensive.

Page Break

consent CONSENT REQUEST FOR PARTICIPATION IN RESEARCH

For questions about the study, contact: Nicolas Koutonias, 430273nk@student.eur.nl

Description: I am a student from the Erasmus University Rotterdam, in The Netherlands, conducting research for my Master's Degree thesis.

The purpose of my study is to examine how people from different backgrounds perceive and interpret potentially offensive and 'politically incorrect' internet meme images. As part of the research, you may be exposed to content that could offend and shock you. This content could be of a racist, sexist, homophobic, or transphobic nature. There is, therefore, a risk of some psychological distress.

By agreeing to participate in this study, you consent to being shown such content.

Time involvement: Your participation in this study will take approximately 10 minutes. You may interrupt your participation at any time.

Payments: There will be no monetary compensation for your participation.

Participants' rights: If you have decided to participate in this project, please understand your participation is voluntary and you have the right to withdraw your consent by discontinuing participation at any time without penalty.The answers you provide will be used in analysis for the purpose of the research, but no private information of any kind will be collected.Your individual privacy will be preserved in all published and written data resulting from this study.

Signing the consent form: By clicking 'Yes, I consent' below, you agree to the terms of participation

of this survey, and agree that your anonymous answers will be recorded and used in my research. You must be at least 18 years old to participate.

- ○ Yes, I consent (Continue survey) (1)
- ○ No, I do not consent (Quit survey) (2)

Skip To: End of Survey If CONSENT REQUEST FOR PARTICIPATION IN RESEARCH For questions about the study, contact: Nicolas Kou... = No, I do not consent (Quit survey)

End of Block: 1) Info, consent

Start of Block: 2) Demographics

info2 In the following section, you will be asked a few short questions about your background and identity.

Page Break

*

age How old are you? Please enter the *year of your birth*.

_____

edu Please indicate your level of schooling/education.

- ○ Some secondary education (high school) (1)
- ○ Completed secondary education (2)
- ○ Trade/technical/vocational training (3)
- ○ Some undergraduate education (college/university) (4)
- ○ Completed undergraduate education (Bachelor's degree) (5)
- ○ Some postgraduate education (6)
- ○ Completed postgraduate education (Master's degree and beyond) (7)
- ○ Other (Please specify) (8) _____

Page Break

residence In which country do you currently reside?

▼ Afghanistan (1) ... Zimbabwe (196)

home [Optional] Is there a country *other than* your country of residence that you would consider your *home*?

▼ Afghanistan (1) ... Zimbabwe (196)

Page Break

gender To which gender identity do you *most identify*?

- ○ Male (1)
- ○ Female (2)
- ○ Non-Binary (3)
- ○ Prefer not to answer (4)
- ○ Other (Please specify) (5) _____

race To which racial or ethnic identity do you *most identify*?

- ○ Arab (1)
- ○ Asian (2)
- ○ Black (3)
- ○ Hispanic (4)
- ○ Indian (5)
- ○ Pacific Islander (6)
- ○ White (7)
- ○ Prefer not to answer (8)
- ○ Other (Please specify) (9) _____

End of Block: 2) Demographics

Start of Block: 3) Ideology

info3a In the following section, you will be asked a few short questions about your beliefs. There are no right or wrong answers.

Page Break

left-right In politics, people sometimes talk of 'left' and 'right'. Where would you place yourself on a scale from 1 to 11 where 1 means *extreme left* and 11 means *extreme right*?

Center

1 2 3 4 5 6 7 8 9 10 11

1 ()	

Page Break

info3b Read the following statements. For each of them, please indicate your level of agreement/disagreement.

ideo1 'The fact on crime, sexual immorality and the recent public disorders all show that we have to crack down harder on deviant groups and troublemakers, if we are going to save our moral standards and preserve law and order.'

- Strongly agree (1)
- Agree (2)
- Somewhat agree (3)
- Neither agree nor disagree (4)
- Somewhat disagree (5)
- Disagree (6)
- Strongly disagree (7)

ideo2 'Homosexuals and feminists should be praised for being brave enough to defy "traditional family values."'

- Strongly agree (1)
- Agree (2)
- Somewhat agree (3)
- Neither agree nor disagree (4)
- Somewhat disagree (5)
- Disagree (6)
- Strongly disagree (7)

Page Break

ideo3 'Everyone should have their own lifestyle, religious beliefs, and sexual preferences, even if it makes them different from everyone else.'

- Strongly agree (1)
- Agree (2)
- Somewhat agree (3)
- Neither agree nor disagree (4)
- Somewhat disagree (5)
- Disagree (6)
- Strongly disagree (7)

ideo4 'It is good that nowadays young people have greater freedom "to make their own rules" and to protest against things they don't like.'

- Strongly agree (1)
- Agree (2)
- Somewhat agree (3)
- Neither agree nor disagree (4)
- Somewhat disagree (5)
- Disagree (6)
- Strongly disagree (7)

---

ideo5 'We should treat protestors and radicals with open arms and open minds, since new ideas are the lifeblood of progressive change'

- Strongly agree (1)
- Agree (2)
- Somewhat agree (3)
- Neither agree nor disagree (4)
- Somewhat disagree (5)
- Disagree (6)
- Strongly disagree (7)

End of Block: 3) Ideology

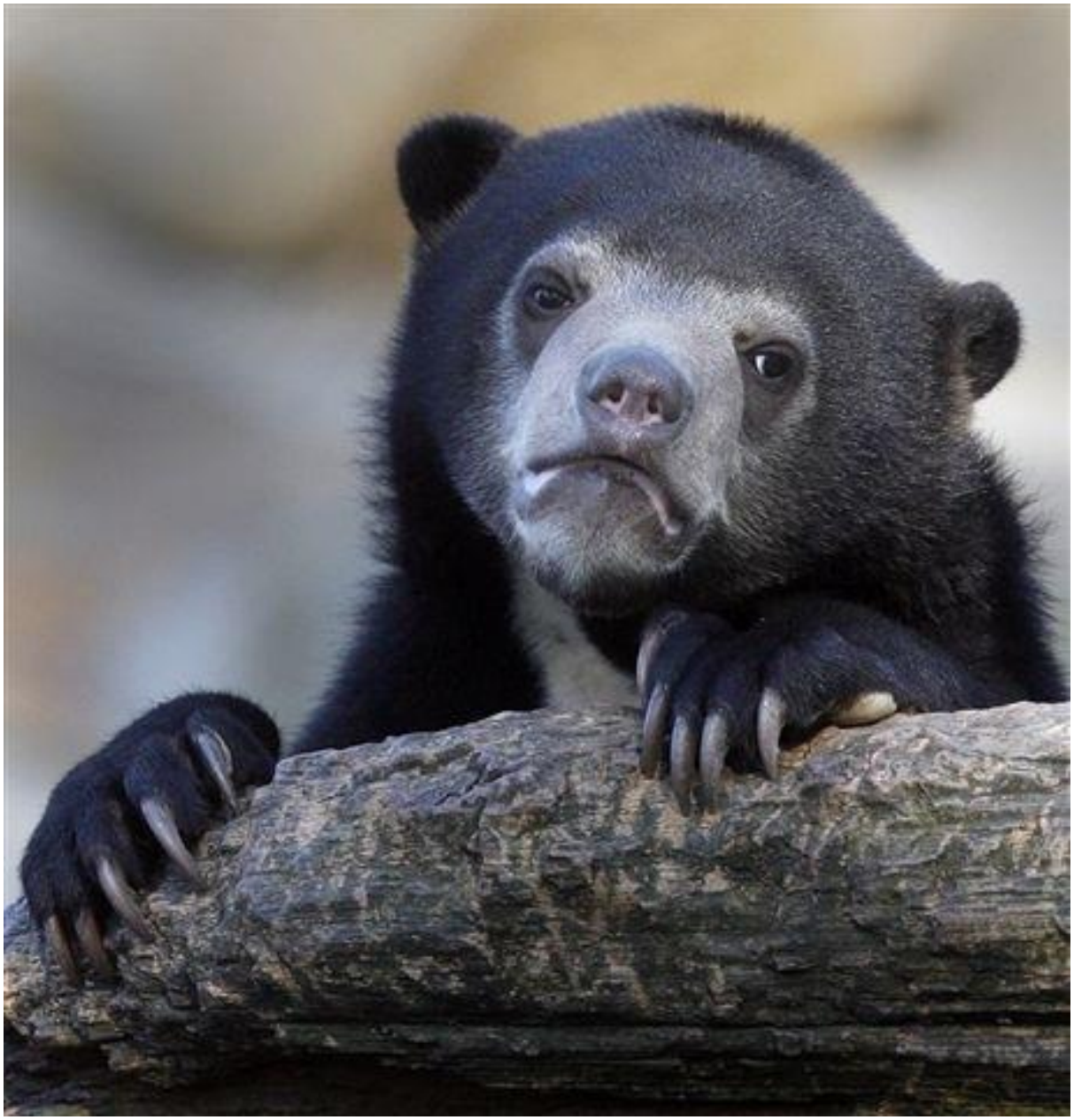
Start of Block: 4) Literacy

info4 In the following section, you will be shown a few text-less internet meme images (or *meme templates*). For each of them, please answer the questions below.

If you do not recognize or know the meme, please select the 'n/a' option.

Page Break

meme1



m1q1 In your opinion, which emotion/attitude/state could be *best* conveyed when using this template?

- ○ Confession (1)
- ○ Shame (2)
- ○ Toxic beliefs (3)
- ○ Annoyance (4)

---

m1q2 How likely are you to find humor in a meme made using this template?

- ○ Extremely likely (1)
- ○ Somewhat likely (2)
- ○ Neither likely nor unlikely (3)
- ○ Somewhat unlikely (4)
- ○ Extremely unlikely (5)
- ○ n/a (6)

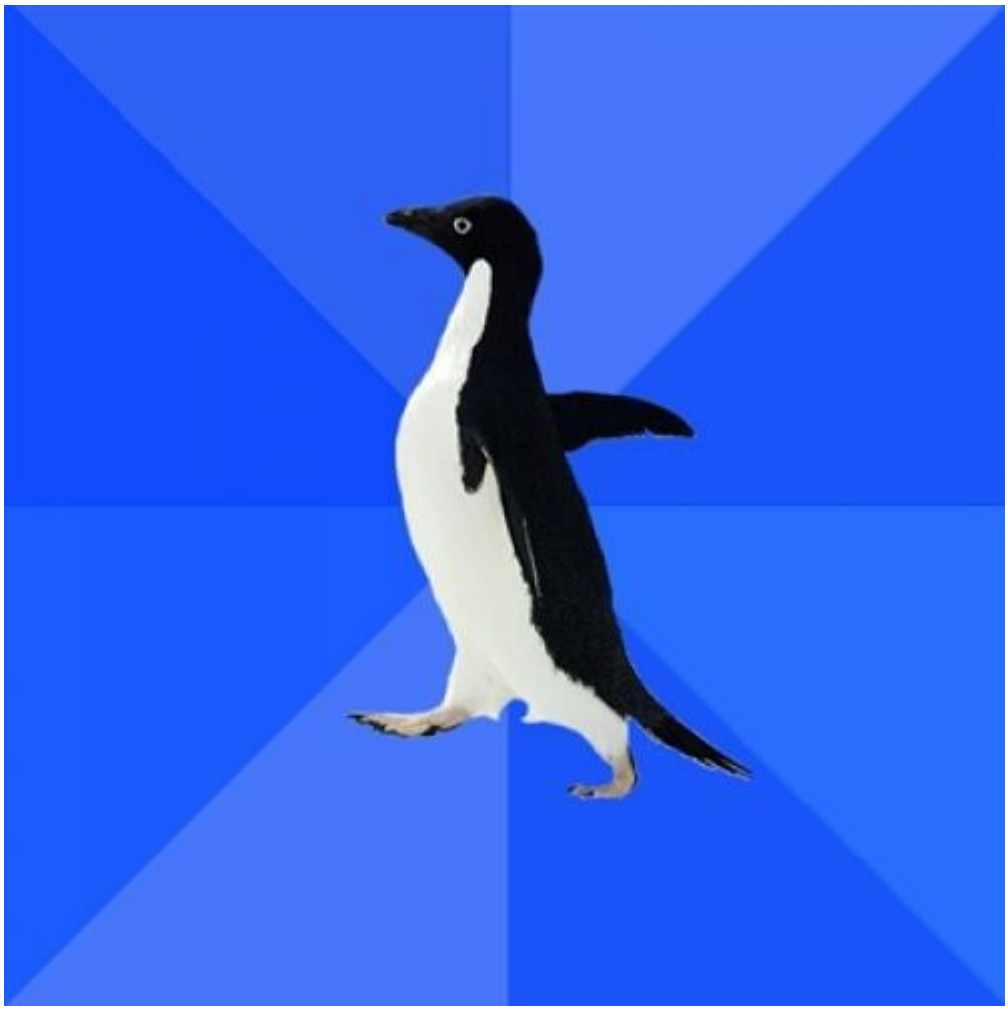
---

m1q3 How easy/difficult would you find using this meme the *'correct' way*?

- ○ Extremely easy (1)
- ○ Somewhat easy (2)
- ○ Neither easy nor difficult (3)
- ○ Somewhat difficult (4)
- ○ Extremely difficult (5)
- ○ n/a (6)

Page Break

meme2



m2q1 In your opinion, which emotion/attitude/state could be *best* conveyed when using this template?

- Social awkwardness (1)
- Willful ignorance (2)
- Social isolation (3)
- Social confidence (4)

m2q2 How likely are you to find humor in a meme made using this template?

- ○ Extremely likely (1)
- ○ Somewhat likely (2)
- ○ Neither likely nor unlikely (3)
- ○ Somewhat unlikely (4)
- ○ Extremely unlikely (5)
- ○ n/a (6)

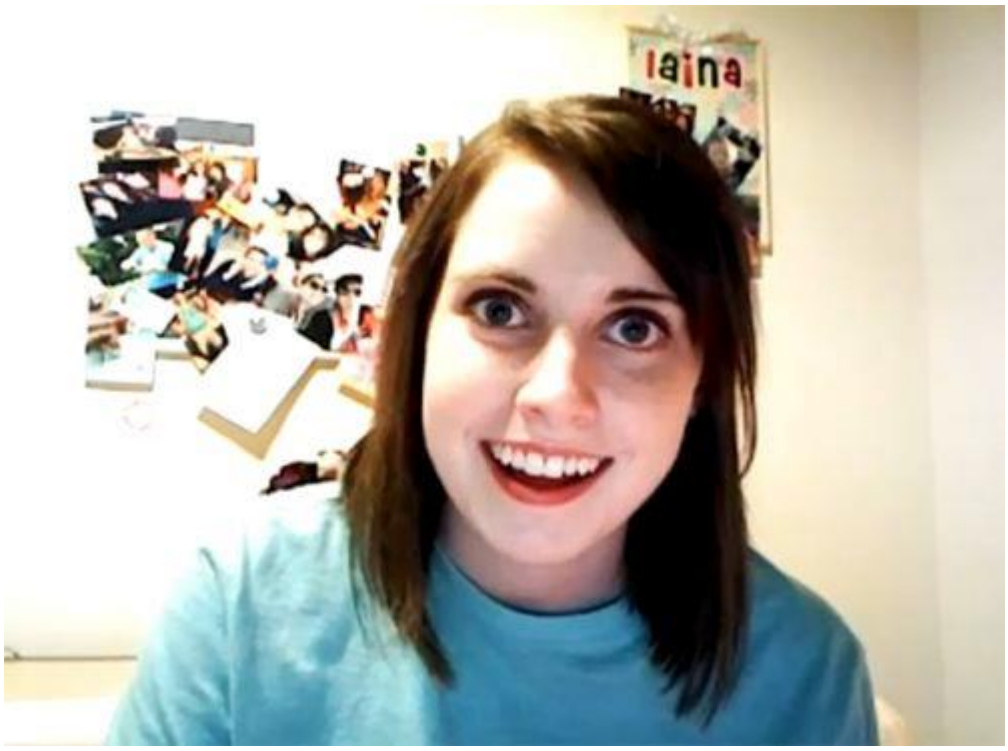
---

m2q3 How easy/difficult would you find using this meme the *'correct' way*?

- ○ Extremely easy (1)
- ○ Somewhat easy (2)
- ○ Neither easy nor difficult (3)
- ○ Somewhat difficult (4)
- ○ Extremely difficult (5)
- ○ n/a (6)

Page Break

meme3



m3q1 In your opinion, which emotion/attitude/state could be *best* conveyed when using this template?

- Excessive attachment (1)
- Social manipulation (2)
- Emotional manipulation (3)
- Emotional abuse (4)

m3q2 How likely are you to find humor in a meme made using this template?

- ○ Extremely likely (1)
- ○ Somewhat likely (2)
- ○ Neither likely nor unlikely (3)
- ○ Somewhat unlikely (4)
- ○ Extremely unlikely (5)
- ○ n/a (6)

---

m3q3 How easy/difficult would you find using this meme the *'correct' way*?

- ○ Extremely easy (1)
- ○ Somewhat easy (2)
- ○ Neither easy nor difficult (3)
- ○ Somewhat difficult (4)
- ○ Extremely difficult (5)
- ○ n/a (6)

Page Break

meme4

m4q1 In your opinion, which emotion/attitude/state could be *best* conveyed when using this template?

- Uncleanliness (1)
- Laziness (2)
- Sexual frustration (3)
- Promiscuity (4)

m4q2 How likely are you to find humor in a meme made using this template?

- Extremely likely (1)
- Somewhat likely (2)
- Neither likely nor unlikely (3)
- Somewhat unlikely (4)
- Extremely unlikely (5)
- n/a (6)

m4q3 How easy/difficult would you find using this meme the *'correct' way*?

- Extremely easy (1)
- Somewhat easy (2)
- Neither easy nor difficult (3)
- Somewhat difficult (4)
- Extremely difficult (5)
- n/a (6)

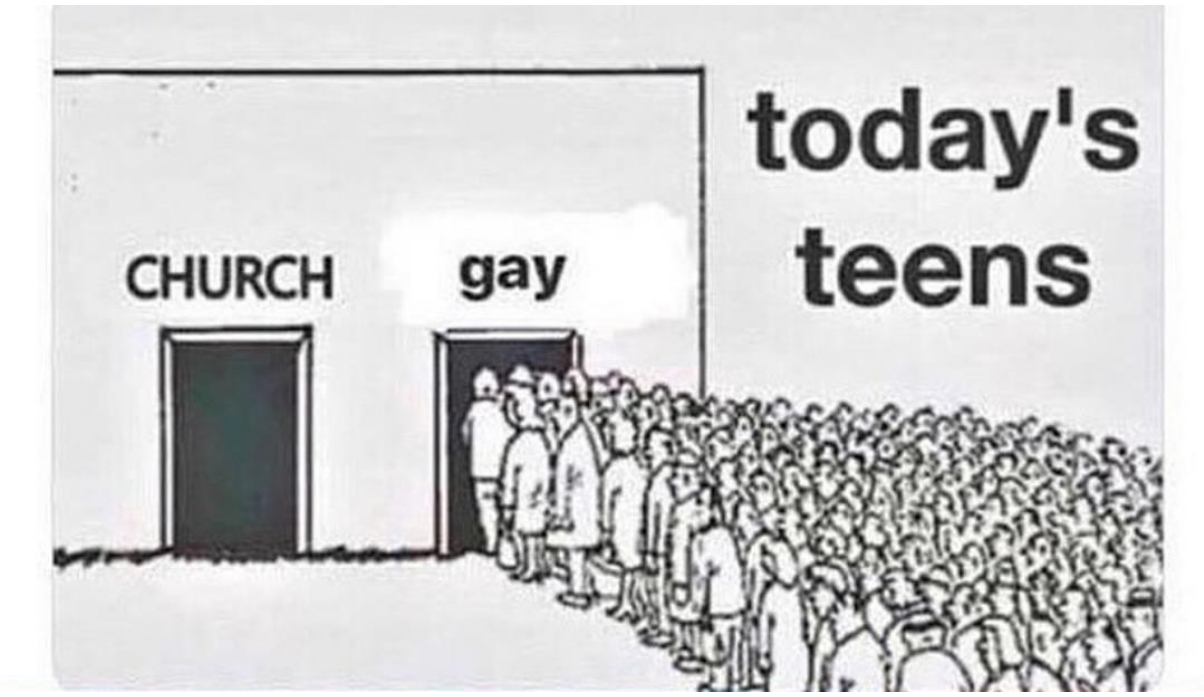
End of Block: 4) Literacy

Start of Block: 5A) Condition1

info5a In the following section, you will be shown a series of internet memes and a short text paragraph. Please read and consider each, then answer the questions below.

Page Break

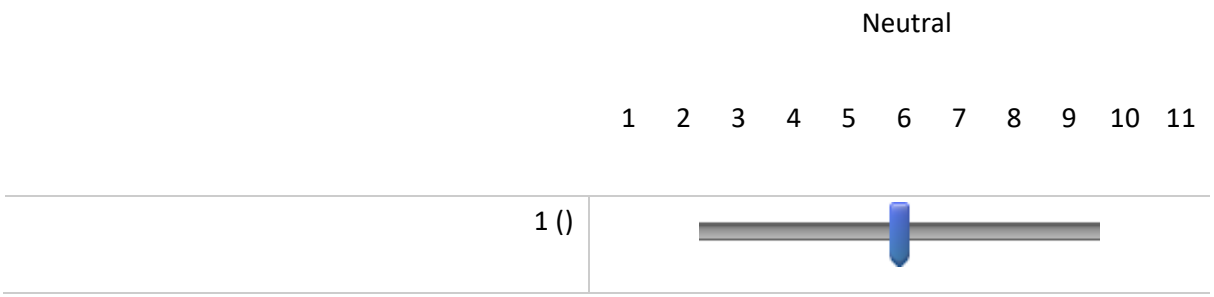
cnd1m1



cnd1m1q1 The above meme presents some kind of information. How *credible* do you think this information is?

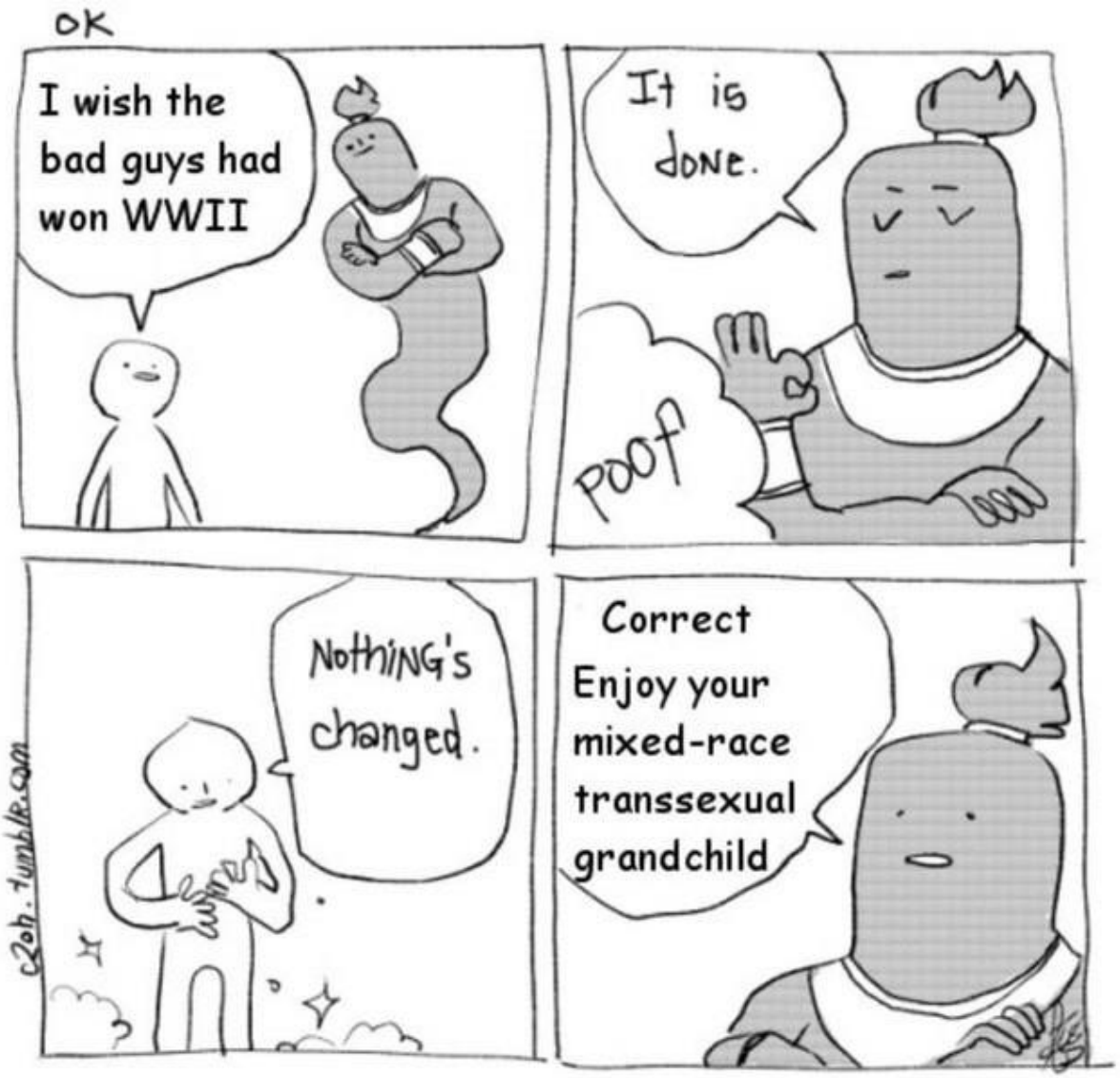
- Extremely credible (1)
- Somewhat credible (2)
- Neutral (3)
- Somewhat not credible (4)
- Extremely not credible (5)
- n/a (I do not understand this meme) (6)

cnd1m1q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

cnd1m2


OK
I wish the bad guys had won WWII
It is done.
Poof
Nothing's changed.
c2oh.tumblr.com
Correct
Enjoy your mixed-race transsexual grandchild

cnd1m2q1 The above meme presents some kind of information. How *credible* do you think this information is?

- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd1m2q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?

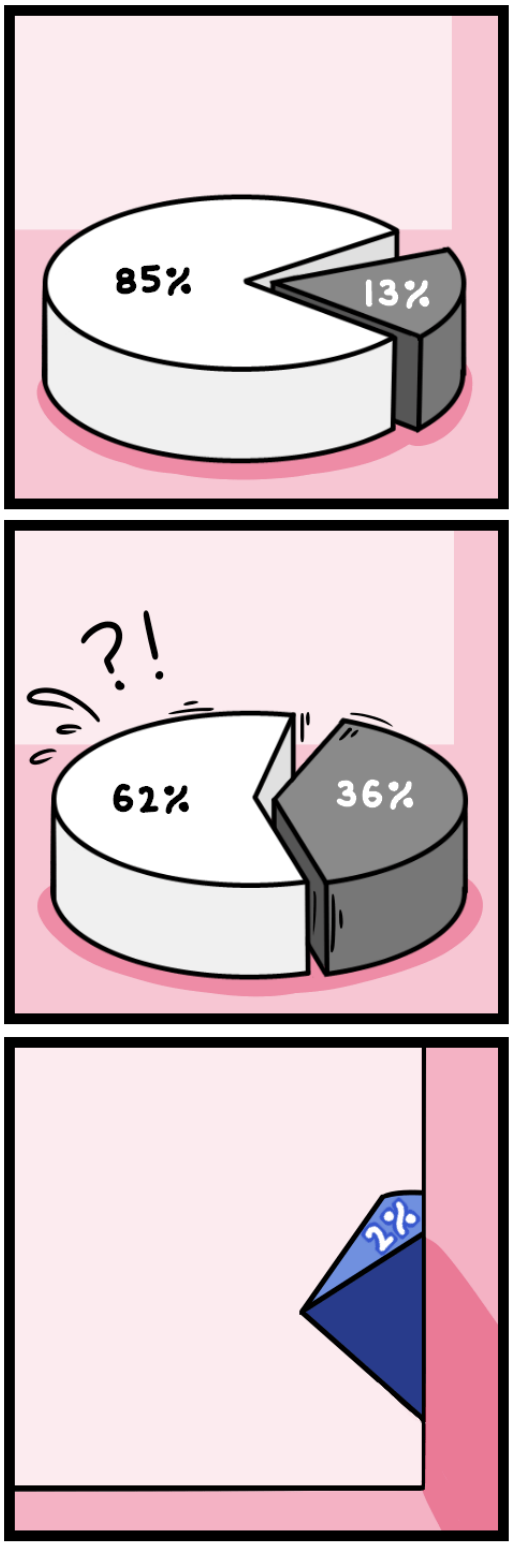
Neutral

1 2 3 4 5 6 7 8 9 10 11

1 ()

Page Break

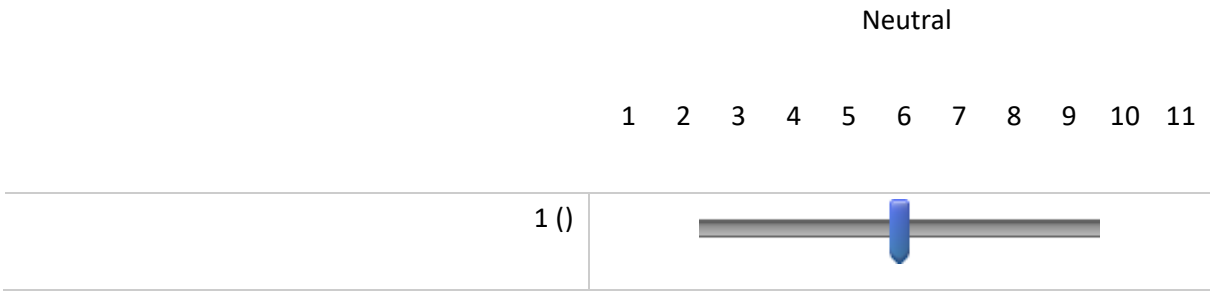
cnd1m3



cnd1m3q1 The above meme presents some kind of information. How *credible* do you think this information is?

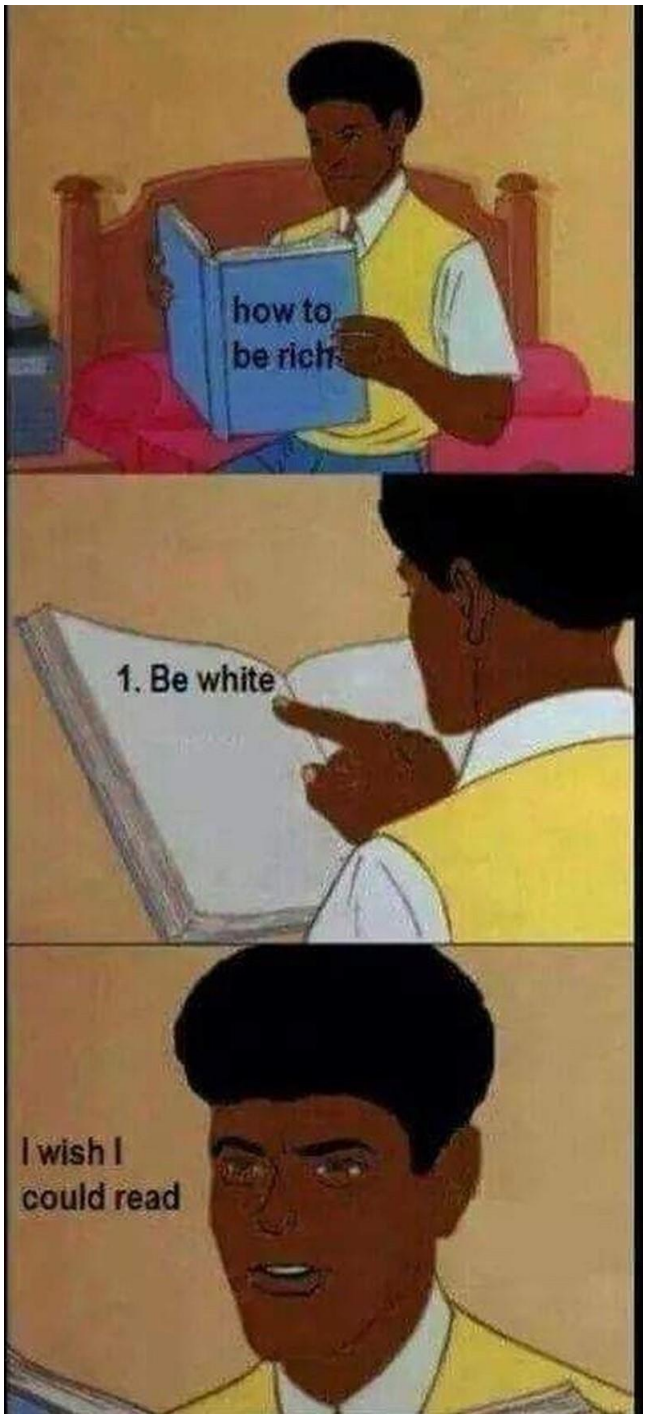
- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd1m3q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

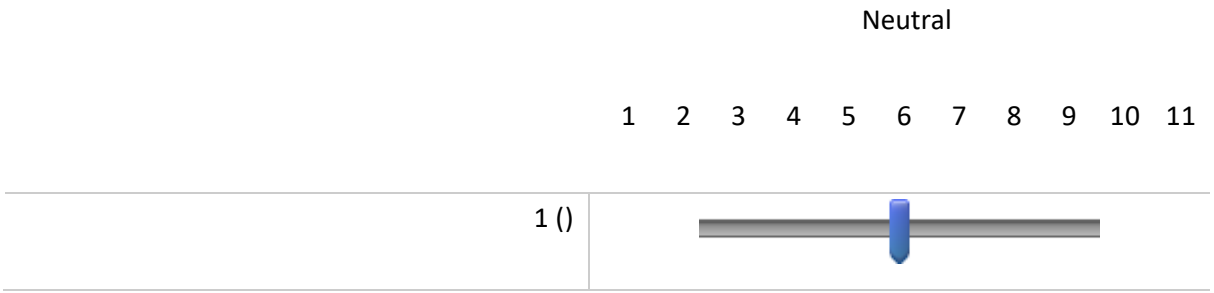
cnd1m4



cnd1m4q1 The above meme presents some kind of information. How *credible* do you think this information is?

- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd1m4q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



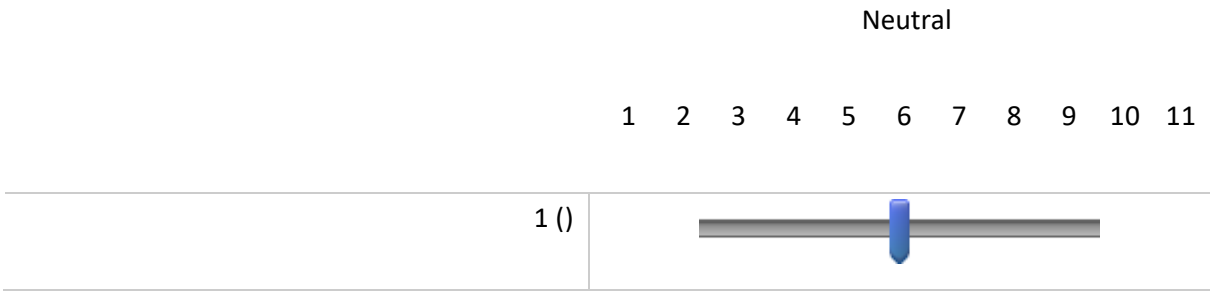
Page Break

cnd1m5

cnd1m5q1 The above meme presents some kind of information. How *credible* do you think this information is?

- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd1m5q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



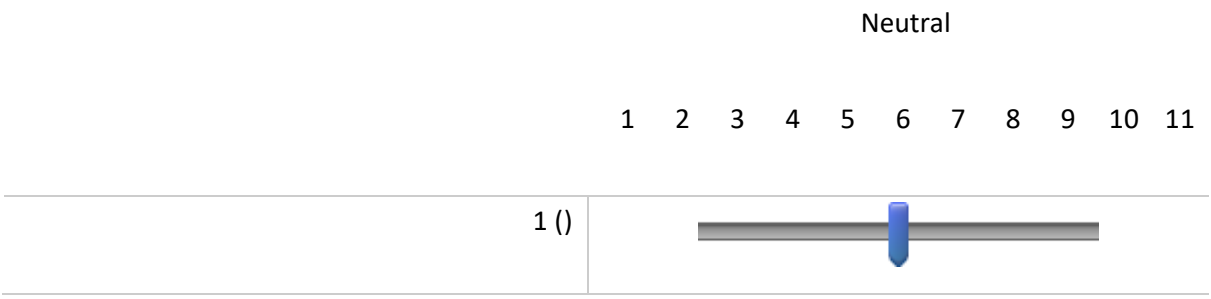
Page Break

cnd1t 'One of the chief forms of racial discrimination, for example, is exclusion. However, a total lack of exclusion/separation is an invitation to miscegenation, to cultural disintegration via borrowing and mixing, to the rewriting and rereading of one's history, to the complete loss of identity, etc. Without any barriers to entry into a group, the group simply cannot protect itself even marginally (or protect its ways, customs, or genes), let alone survive as a distinct group. It will assimilate the surrounding world or be assimilated into it. In this way, open borders combined with comprehensive anti-discrimination laws, is not simply bad policy, it is implicitly/effectively genocidal.'

cnd1tq1 The above text presents some kind of information. How *credible* do you think this information is?

- Extremely credible (1)
- Somewhat credible (2)
- Neutral (3)
- Somewhat not credible (4)
- Extremely not credible (5)
- n/a (I do not understand this text) (6)

cnd1tq2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above text could *best be placed*?



End of Block: 5A) Condition1

Start of Block: 5B) Condition2

info5b In the following section, you will be shown a short text paragraph and a series of internet memes. Please read and consider each, then answer the questions below.

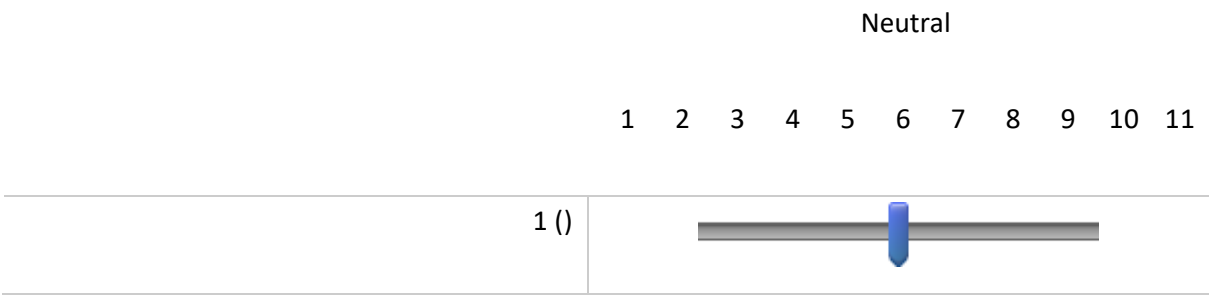
Page Break

cnd2t 'One of the chief forms of racial discrimination, for example, is exclusion. However, a total lack of exclusion/separation is an invitation to miscegenation, to cultural disintegration via borrowing and mixing, to the rewriting and rereading of one's history, to the complete loss of identity, etc. Without any barriers to entry into a group, the group simply cannot protect itself even marginally (or protect its ways, customs, or genes), let alone survive as a distinct group. It will assimilate the surrounding world or be assimilated into it. In this way, open borders combined with comprehensive anti-discrimination laws, is not simply bad policy, it is implicitly/effectively genocidal.'

cnd2tq1 The above text presents some kind of information. How *credible* do you think this information is?

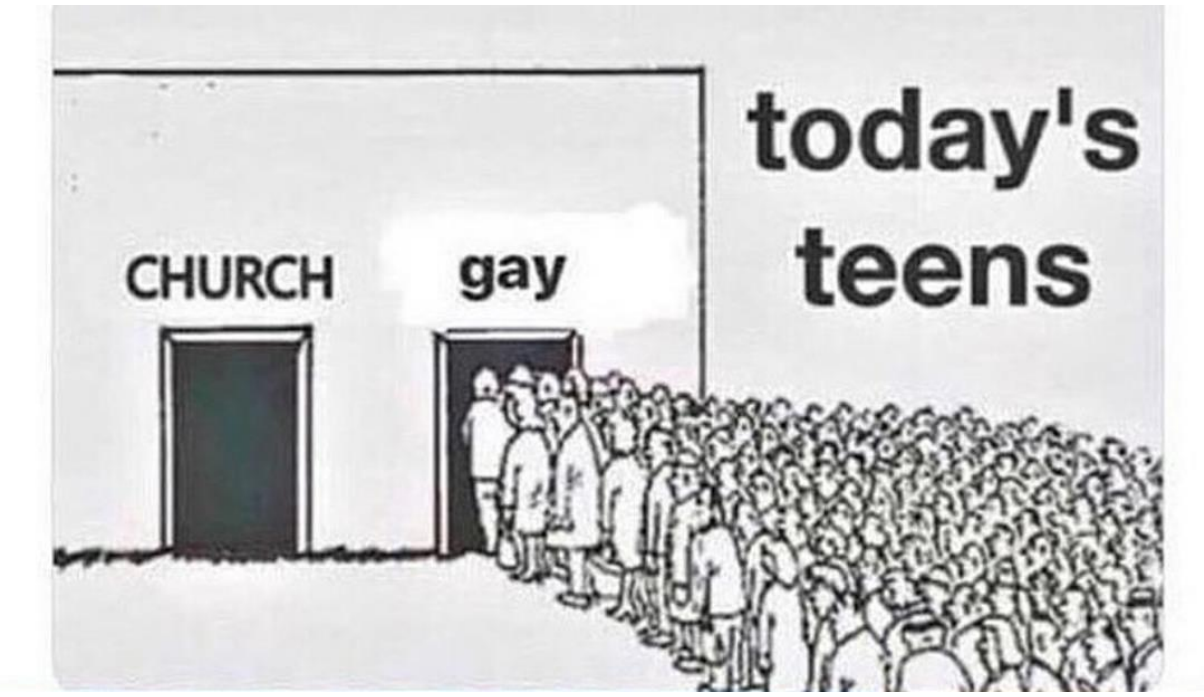
- Extremely credible (1)
- Somewhat credible (2)
- Neutral (3)
- Somewhat not credible (4)
- Extremely not credible (5)
- n/a (I do not understand this text) (6)

cnd2tq2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above text could *best be placed*?



Page Break

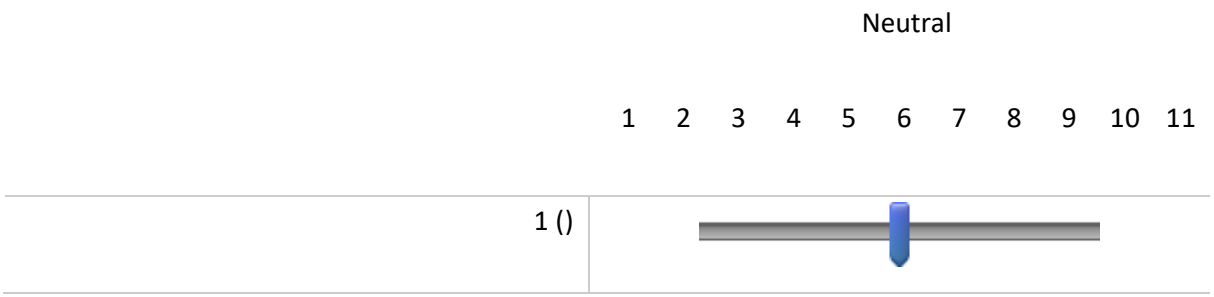
cnd2m1



cnd2m1q1 The above meme presents some kind of information. How *credible* do you think this information is?

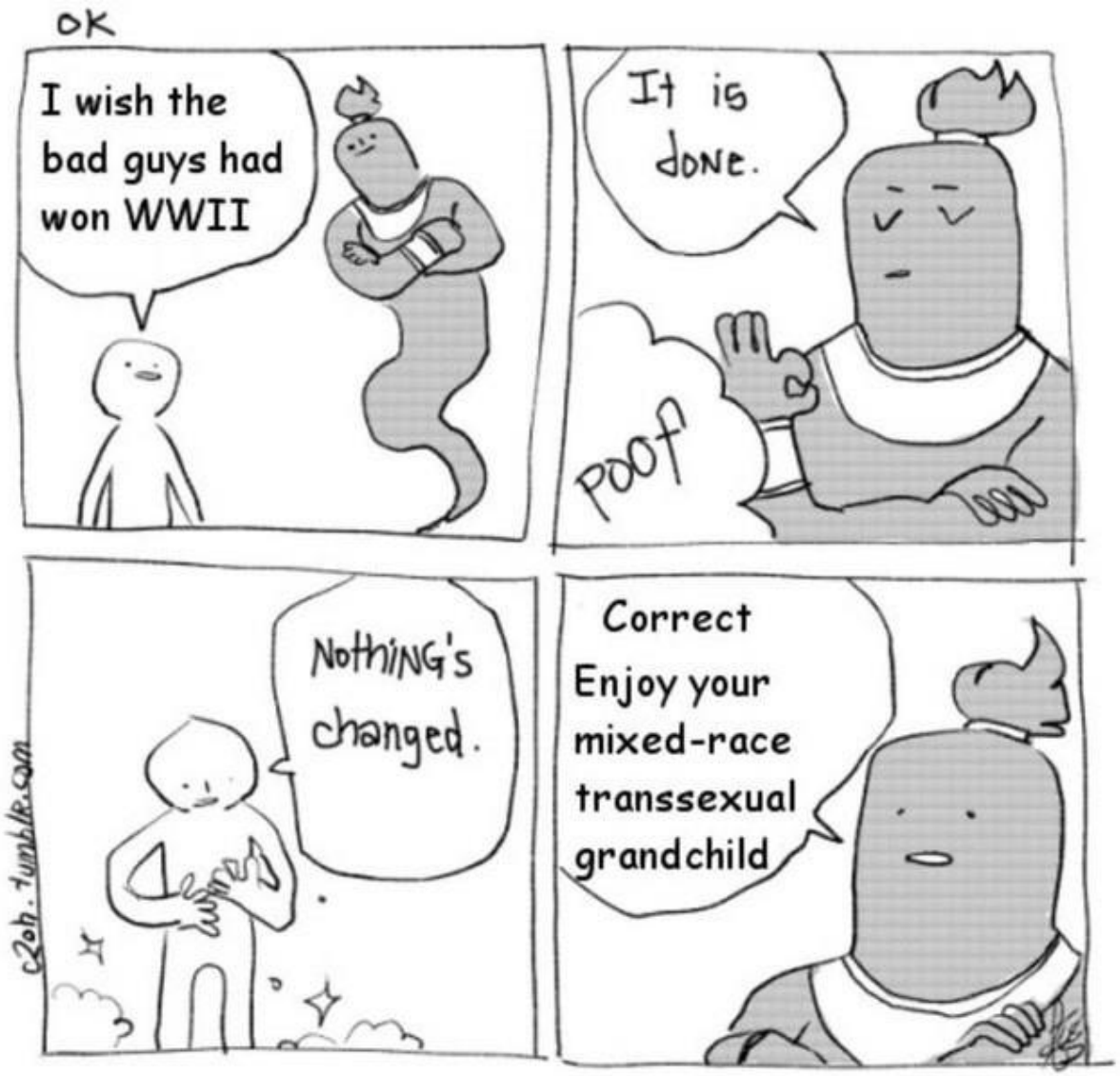
- Extremely credible (1)
- Somewhat credible (2)
- Neutral (3)
- Somewhat not credible (4)
- Extremely not credible (5)
- n/a (I do not understand this meme) (6)

cnd2m1q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

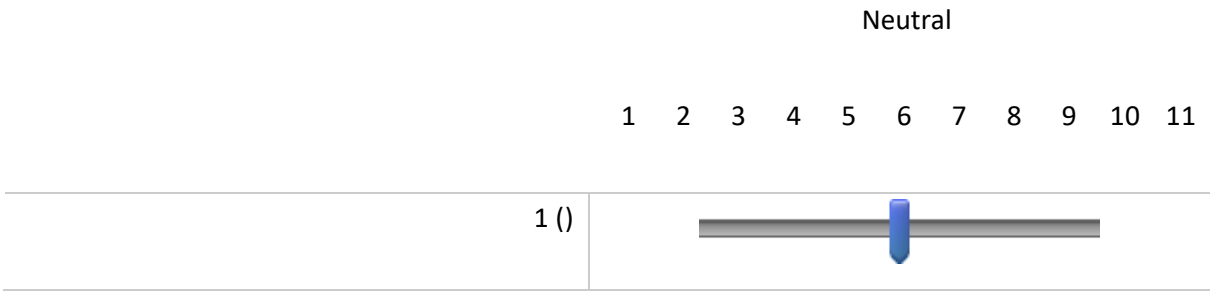
cnd2m2


OK
I wish the bad guys had won WWII
It is done.
poof
Nothing's changed.
c2oh.tumblr.com
Correct
Enjoy your mixed-race transsexual grandchild

cnd2m2q1 The above meme presents some kind of information. How *credible* do you think this information is?

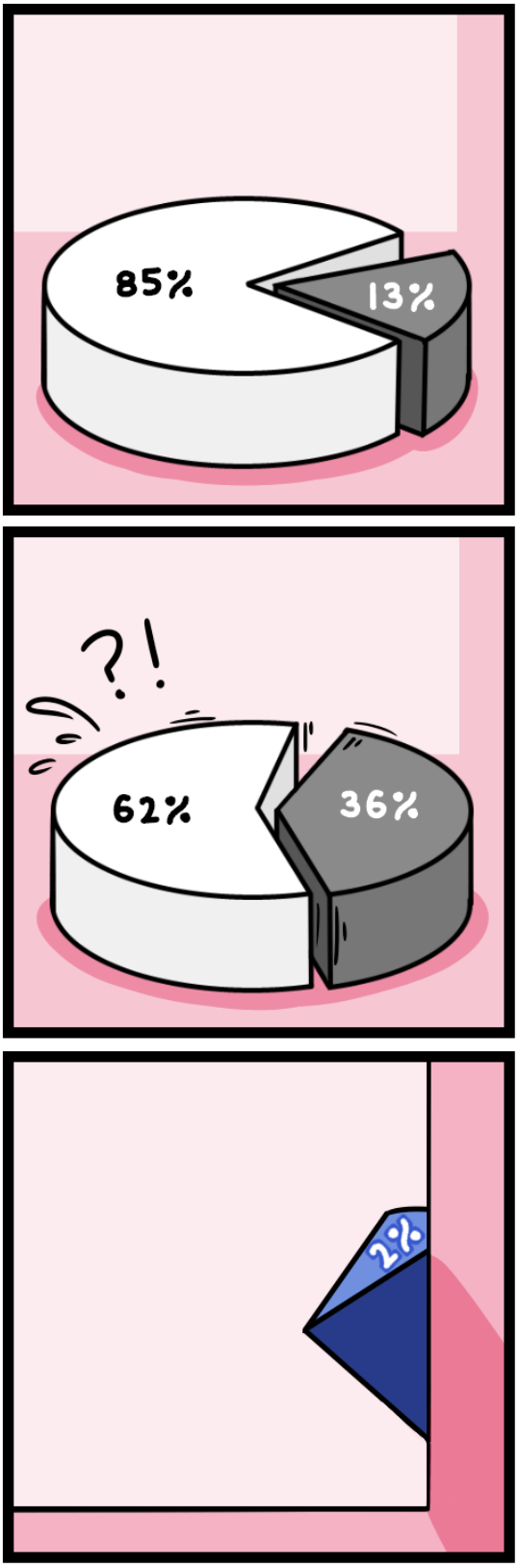
- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd2m2q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

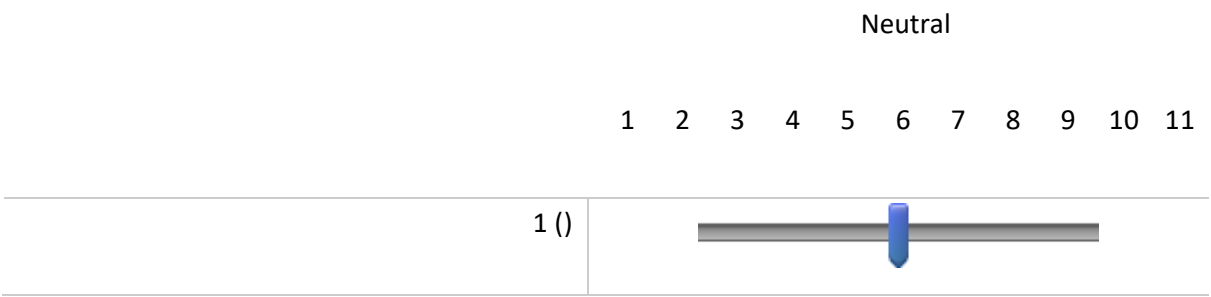
cnd2m3



cnd2m3q1 The above meme presents some kind of information. How *credible* do you think this information is?

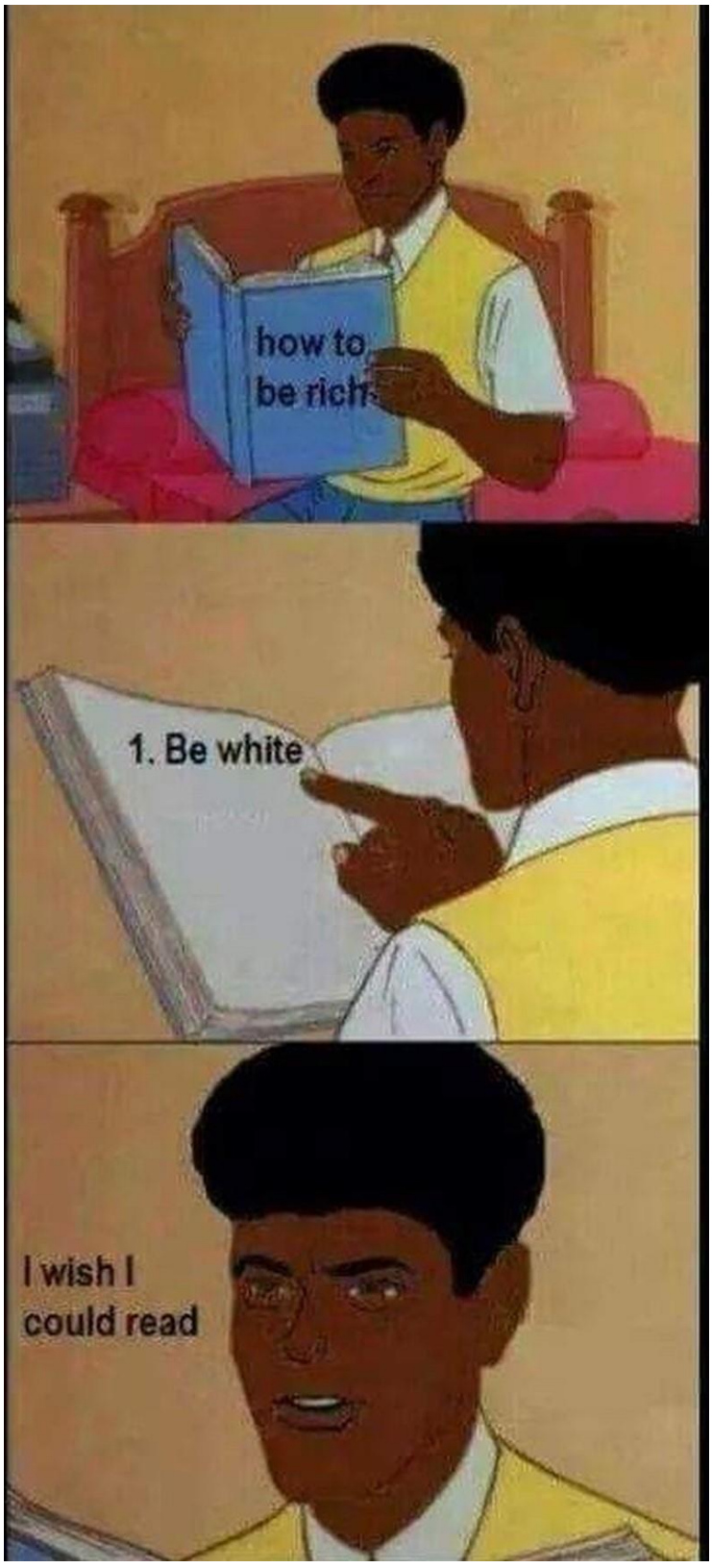
- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd2m3q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

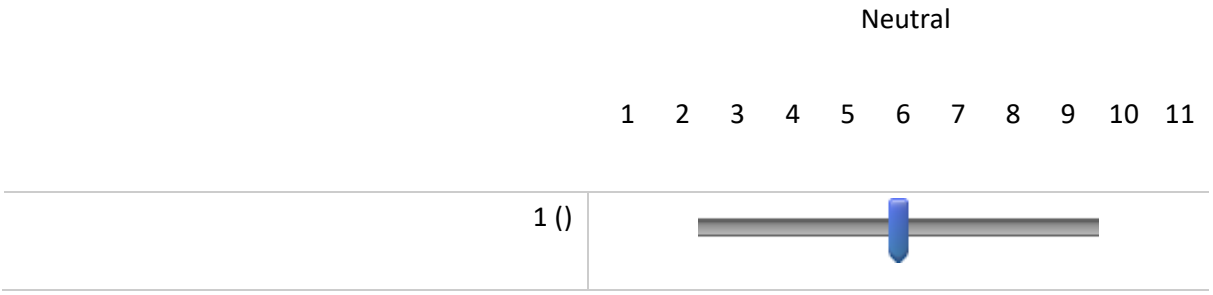
cnd2m4



cnd2m4q1 The above meme presents some kind of information. How credible do you think this information is?

- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd2m4q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



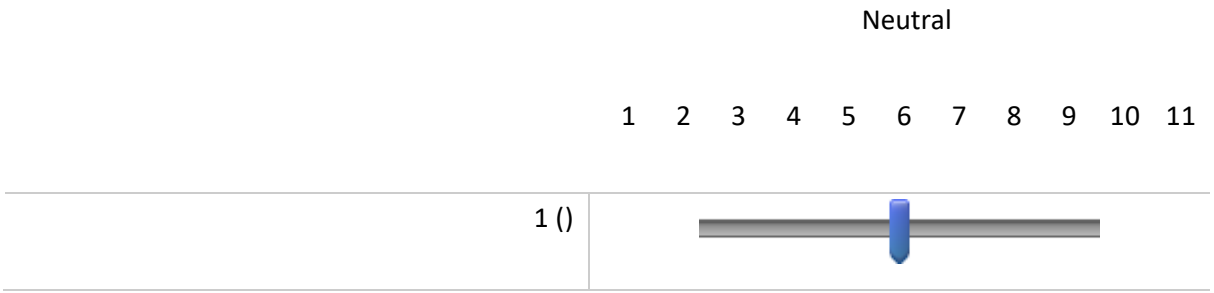
Page Break

cnd2m5

cnd2m5q1 The above meme presents some kind of information. How credible do you think this information is?

- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd2m5q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



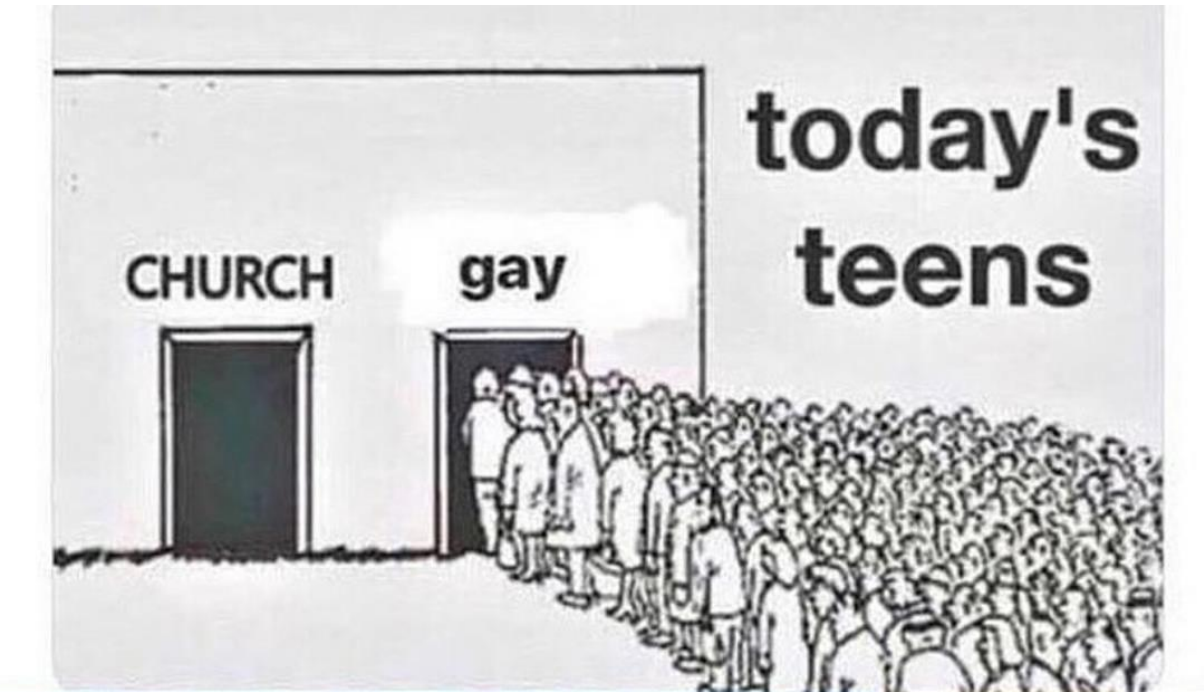
End of Block: 5B) Condition2

Start of Block: 5C) Condition3

info5c In the following section, you will be shown a series of internet memes. Please read and consider each, then answer the questions below.

Page Break

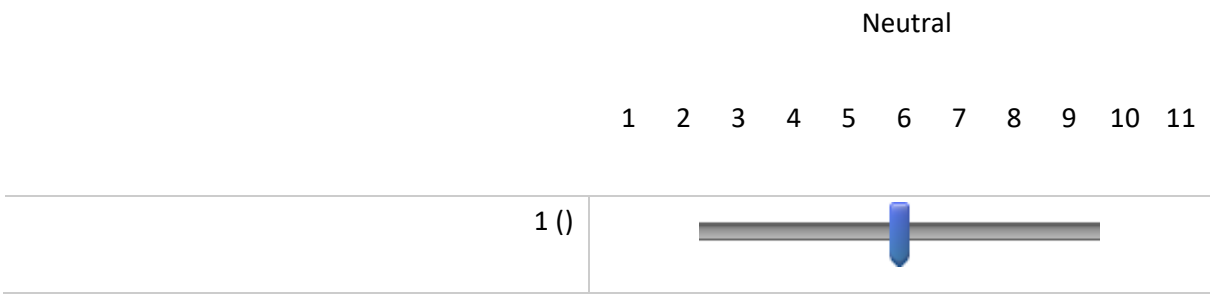
cnd3m1



cnd3m1q1 The above meme presents some kind of information. How *credible* do you think this information is?

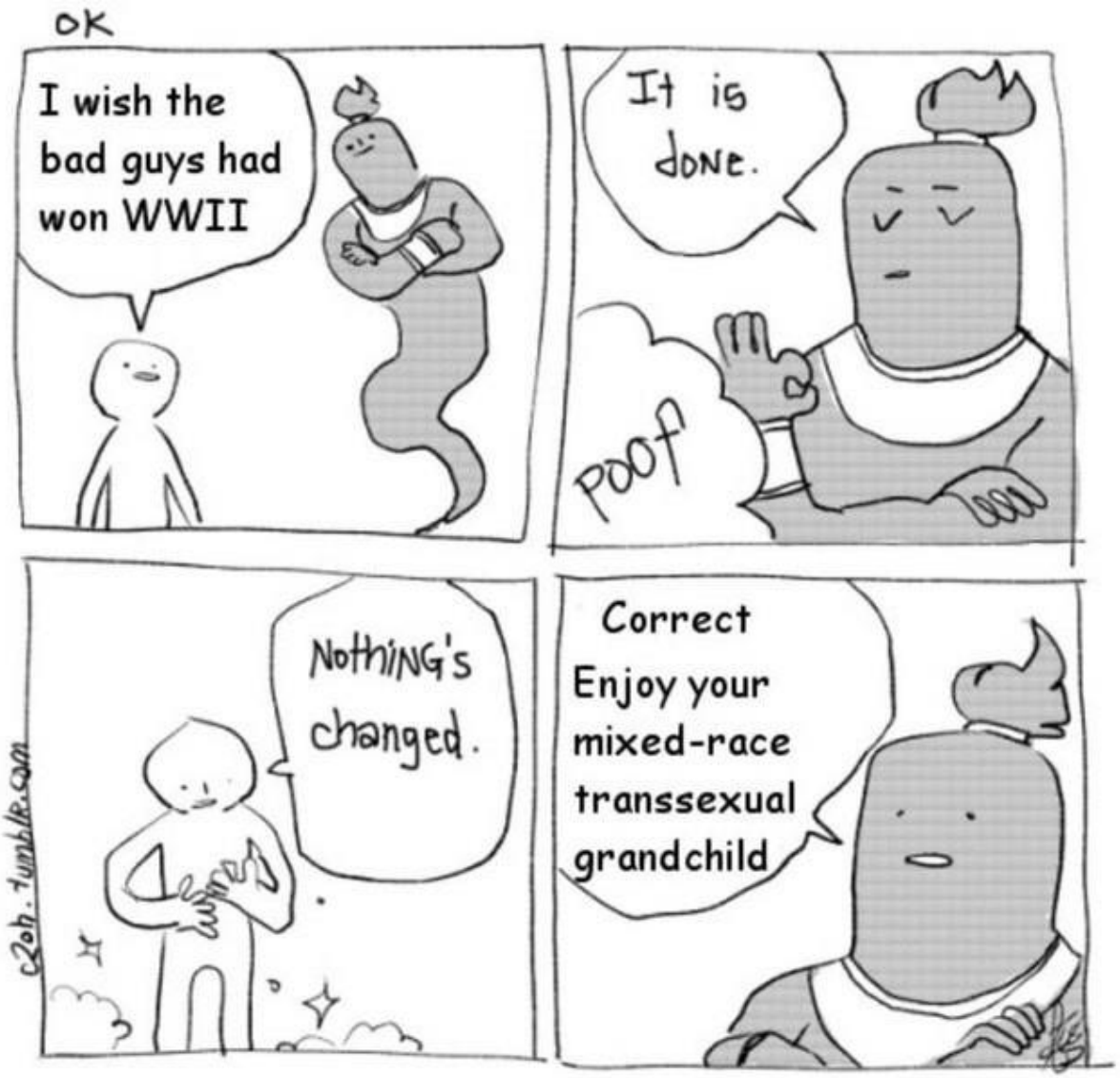
- Extremely credible (1)
- Somewhat credible (2)
- Neutral (3)
- Somewhat not credible (4)
- Extremely not credible (5)
- n/a (I do not understand this meme) (6)

cnd3m1q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

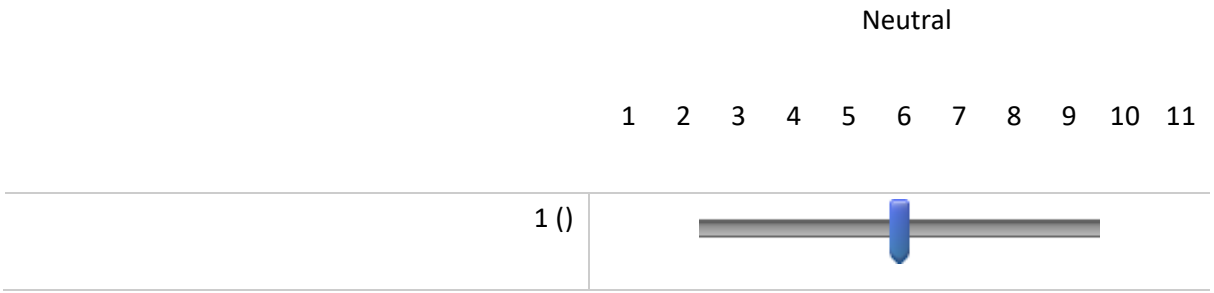
cnd3m2


OK
I wish the bad guys had won WWII
It is done.
poof
Nothing's changed.
c2oh.tumblr.com
Correct
Enjoy your mixed-race transsexual grandchild

cnd3m2q1 The above meme presents some kind of information. How *credible* do you think this information is?

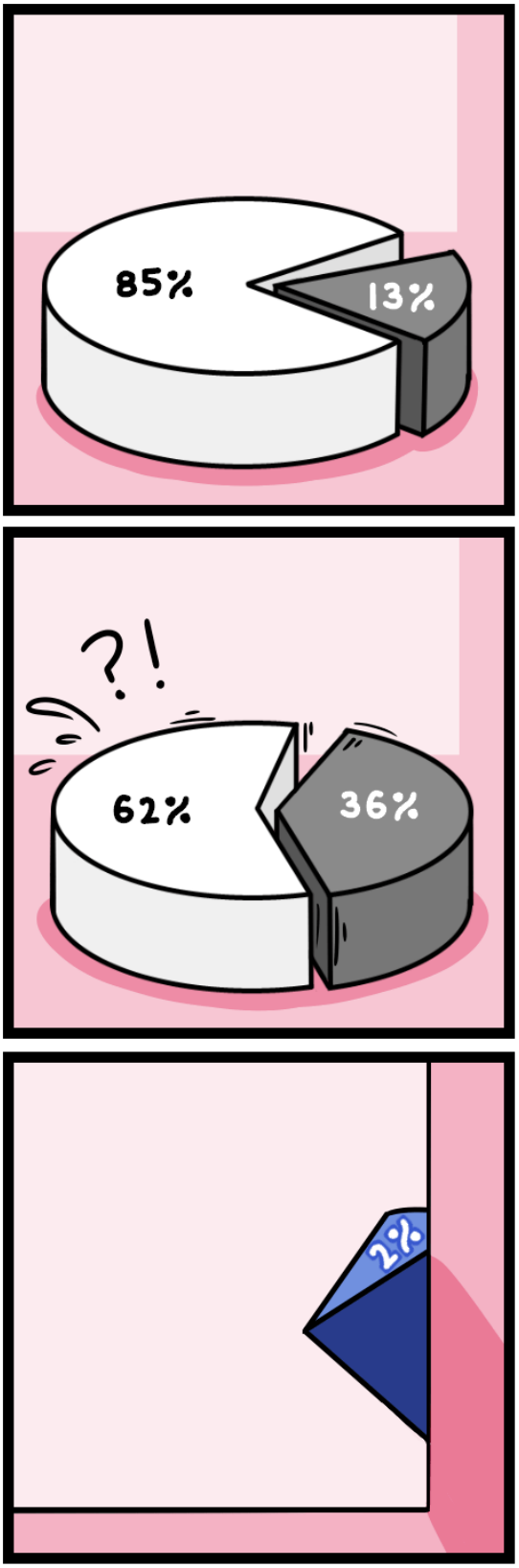
- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd3m2q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

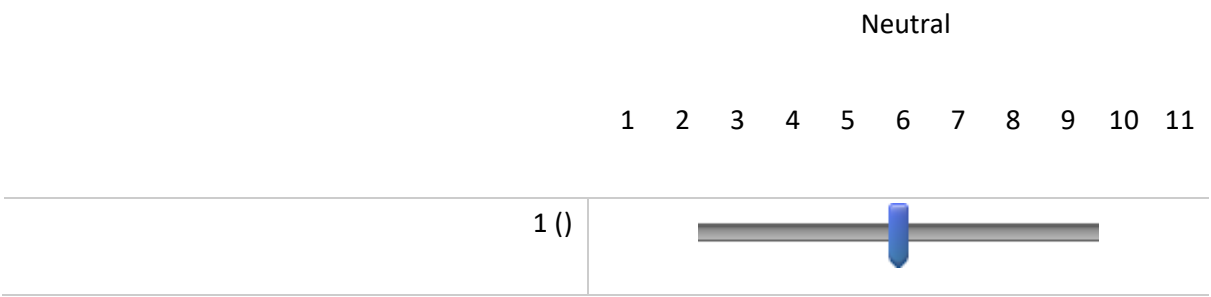
cnd3m3



cnd3m3q1 The above meme presents some kind of information. How *credible* do you think this information is?

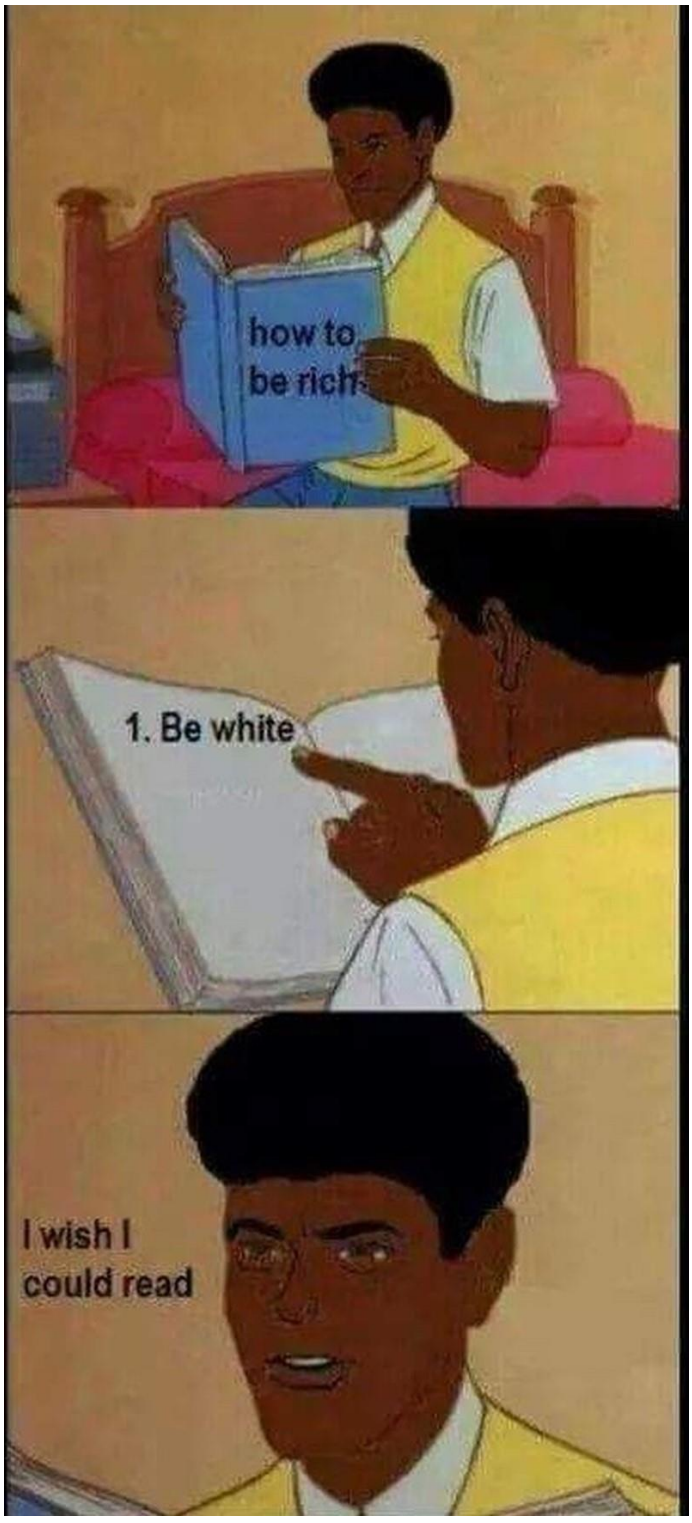
- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd3m3q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



Page Break

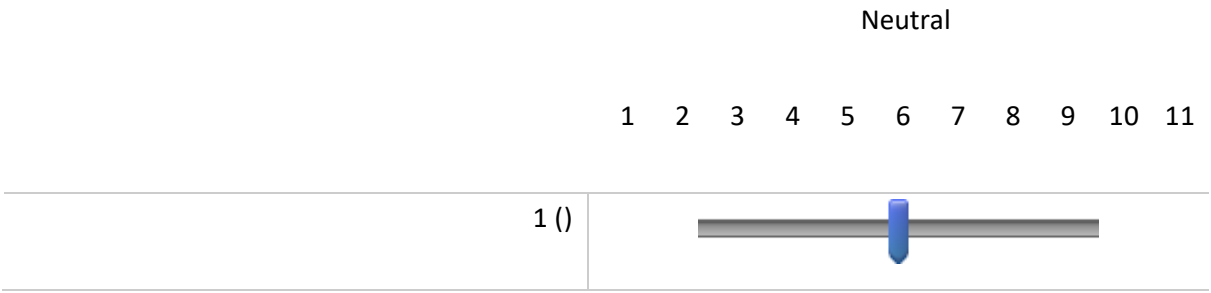
cnd3m4



cnd3m4q1 The above meme presents some kind of information. How credible do you think this information is?

- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd3m4q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



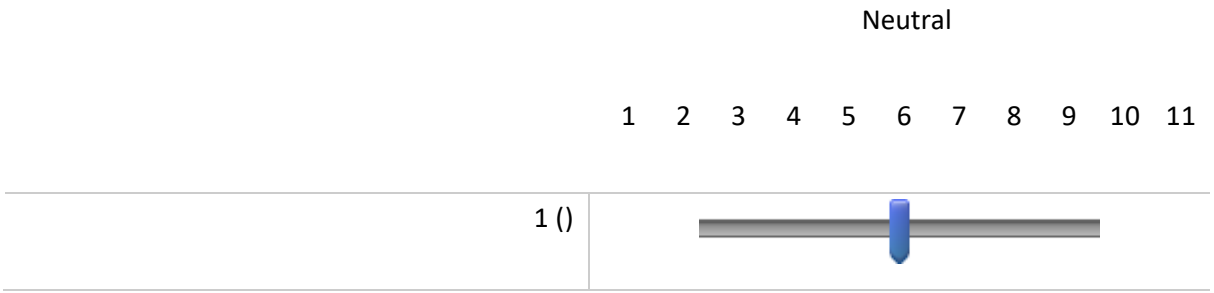
Page Break

cnd3m5

cnd3m5q1 The above meme presents some kind of information. How credible do you think this information is?

- ○ Extremely credible (1)
- ○ Somewhat credible (2)
- ○ Neutral (3)
- ○ Somewhat not credible (4)
- ○ Extremely not credible (5)
- ○ n/a (I do not understand this meme) (6)

cnd3m5q2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above meme could *best be placed*?



End of Block: 5C) Condition3

Start of Block: 5D) Condition4

info5d In the following section, you will be shown a short text paragraph. Please read and consider it, then answer the questions below.

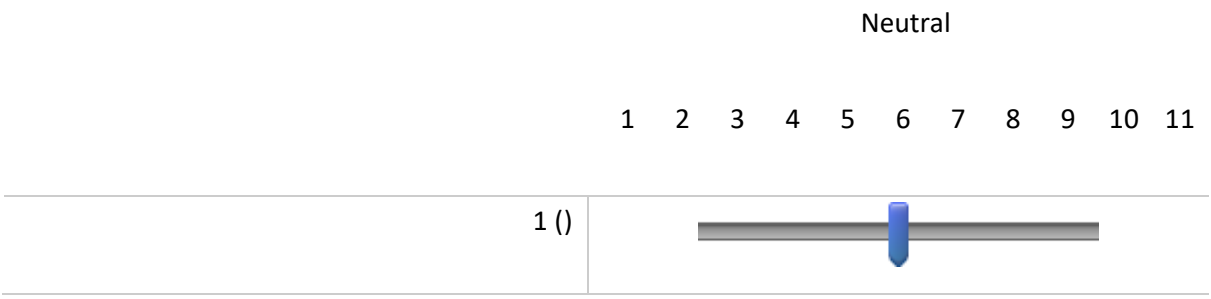
Page Break

cnd4t 'One of the chief forms of racial discrimination, for example, is exclusion. However, a total lack of exclusion/separation is an invitation to miscegenation, to cultural disintegration via borrowing and mixing, to the rewriting and rereading of one's history, to the complete loss of identity, etc. Without any barriers to entry into a group, the group simply cannot protect itself even marginally (or protect its ways, customs, or genes), let alone survive as a distinct group. It will assimilate the surrounding world or be assimilated into it. In this way, open borders combined with comprehensive anti-discrimination laws, is not simply bad policy, it is implicitly/effectively genocidal.'

cnd4tq1 The above text presents some kind of information. How *credible* do you think this information is?

- Extremely credible (1)
- Somewhat credible (2)
- Neutral (3)
- Somewhat not credible (4)
- Extremely not credible (5)
- n/a (I do not understand this text) (6)

cnd4tq2 On a scale of 1 to 11, where 1 means *harmless* and 11 means *harmful*, where do you think the above text could *best be placed*?



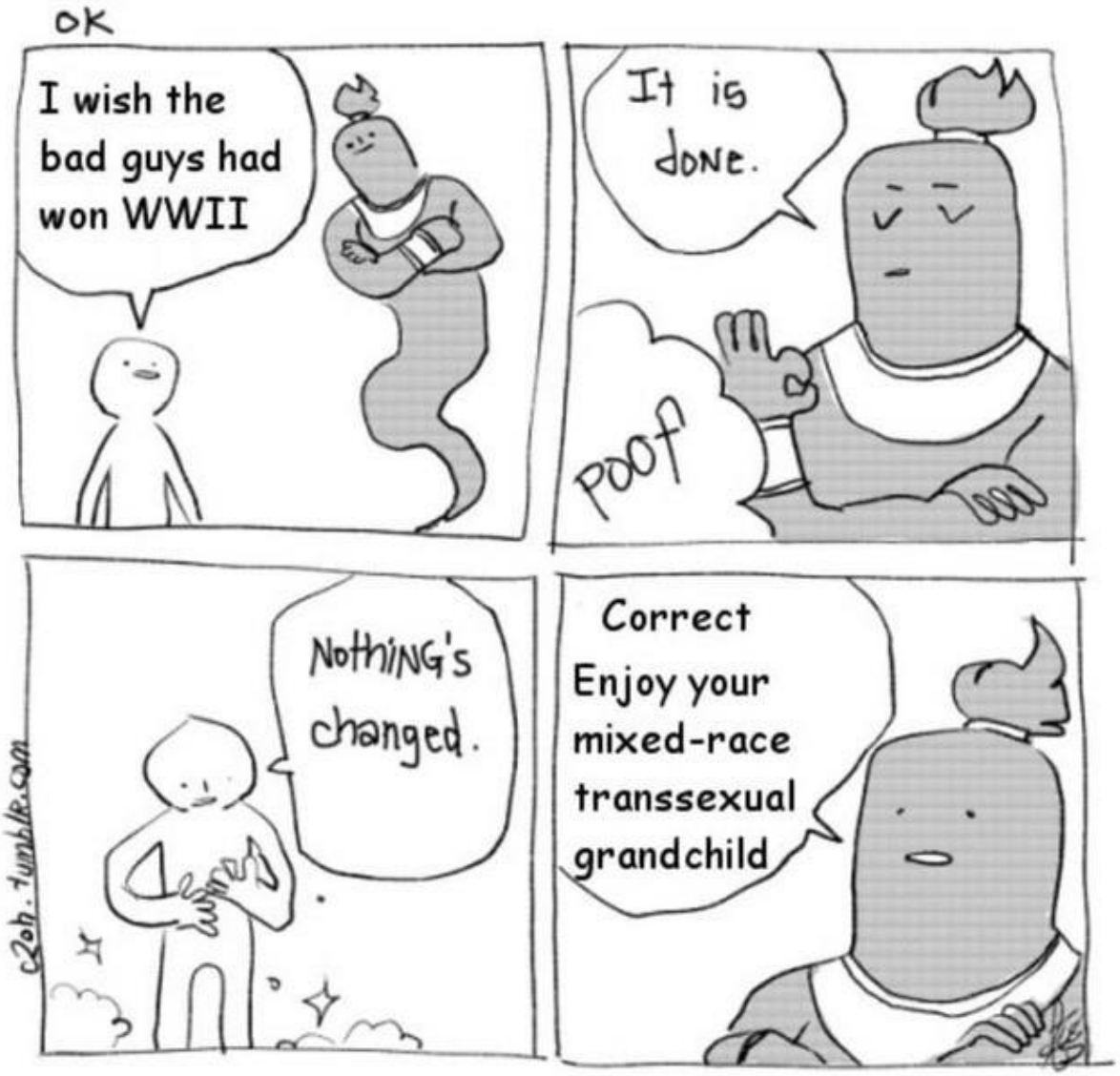
End of Block: 5D) Condition4

Start of Block: 7) Perceived intent

info7 In the last section of this survey, you will be shown some offensive memes. Consider each, then please indicate what you think the intent behind its creation was.

Page Break

intm1a



intm1und Did you understand what information this meme is meant to convey?

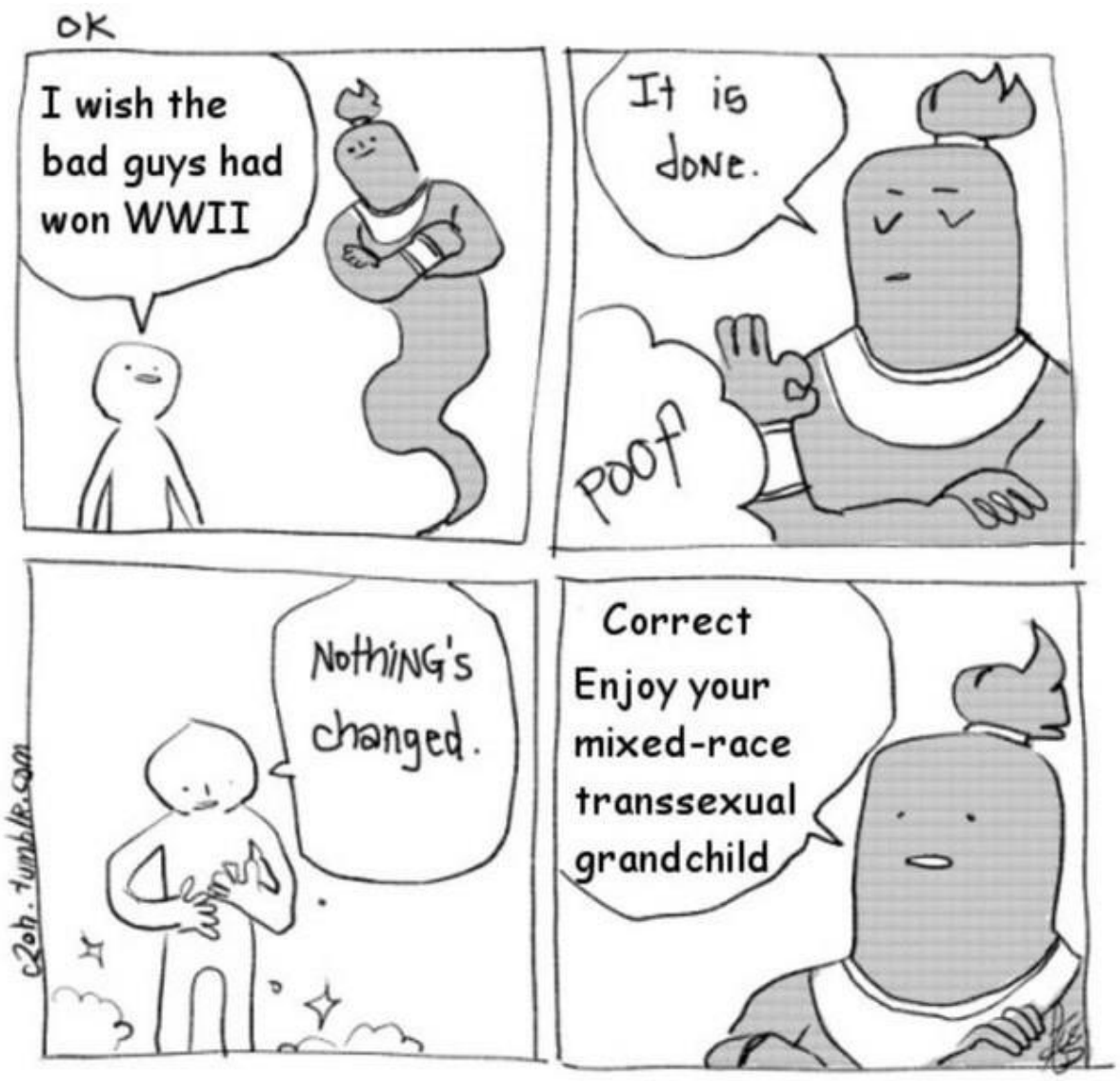
- Yes (1)
- No (2)

Display This Question:

If Did you understand what information this meme is meant to convey? = No

info7b This meme is racist and transphobic.
It implies that the Nazis were 'correct' (and therefore the 'good guys') in their persecution of people of non-white ethnic backgrounds or with a trangender personal identity.

intm1b

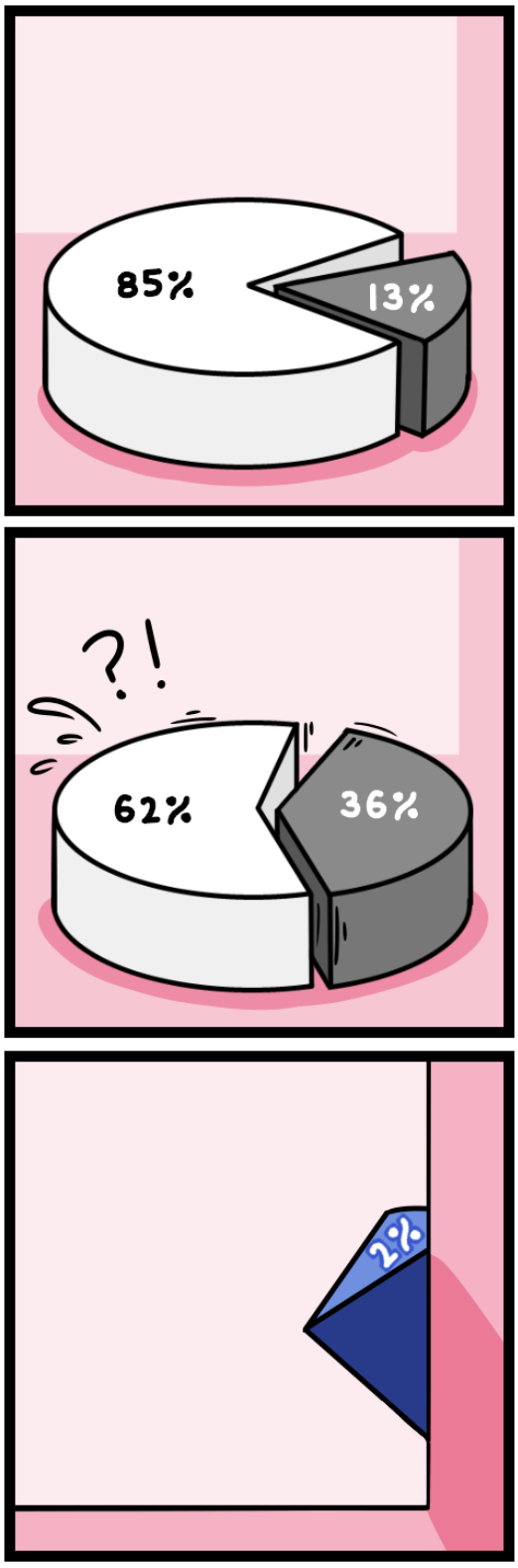


intm1q In your opinion, what was the *intent* behind the creation of this meme?

- ○ Satire (1)
- ○ Offense for the sake of comedy (2)
- ○ Offense for the sake of hurting others (3)
- ○ Expression of genuine personal beliefs (4)
- ○ Not sure (5)

Page Break

intm2a



intm2und Did you understand what information this meme is meant to convey?

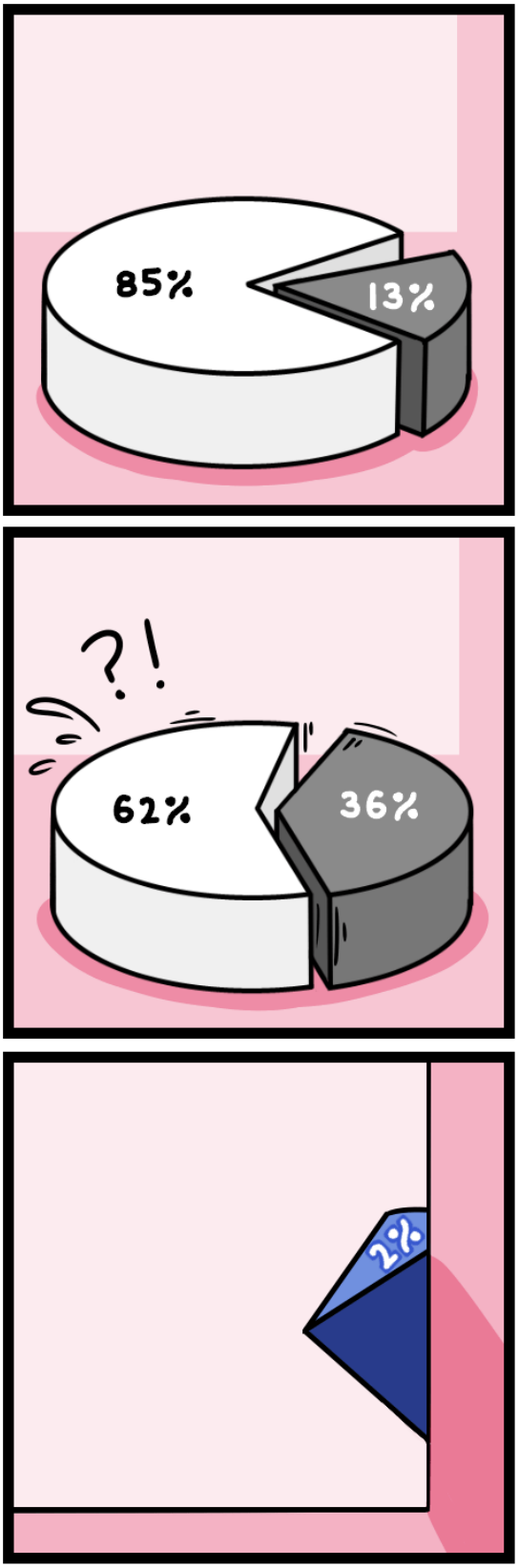
- ○ Yes (1)
- ○ No (2)

Display This Question:

If Did you understand what information this meme is meant to convey? = No

info7c This meme is racist.
It implies that 'white genocide' (a conspiracy theory which states that 'race-mixing' will lead to a disappearance of the 'white race') is orchestrated by the Jewish people (who are, based on another conspiracy theory, secretively controlling the world).

intm2b

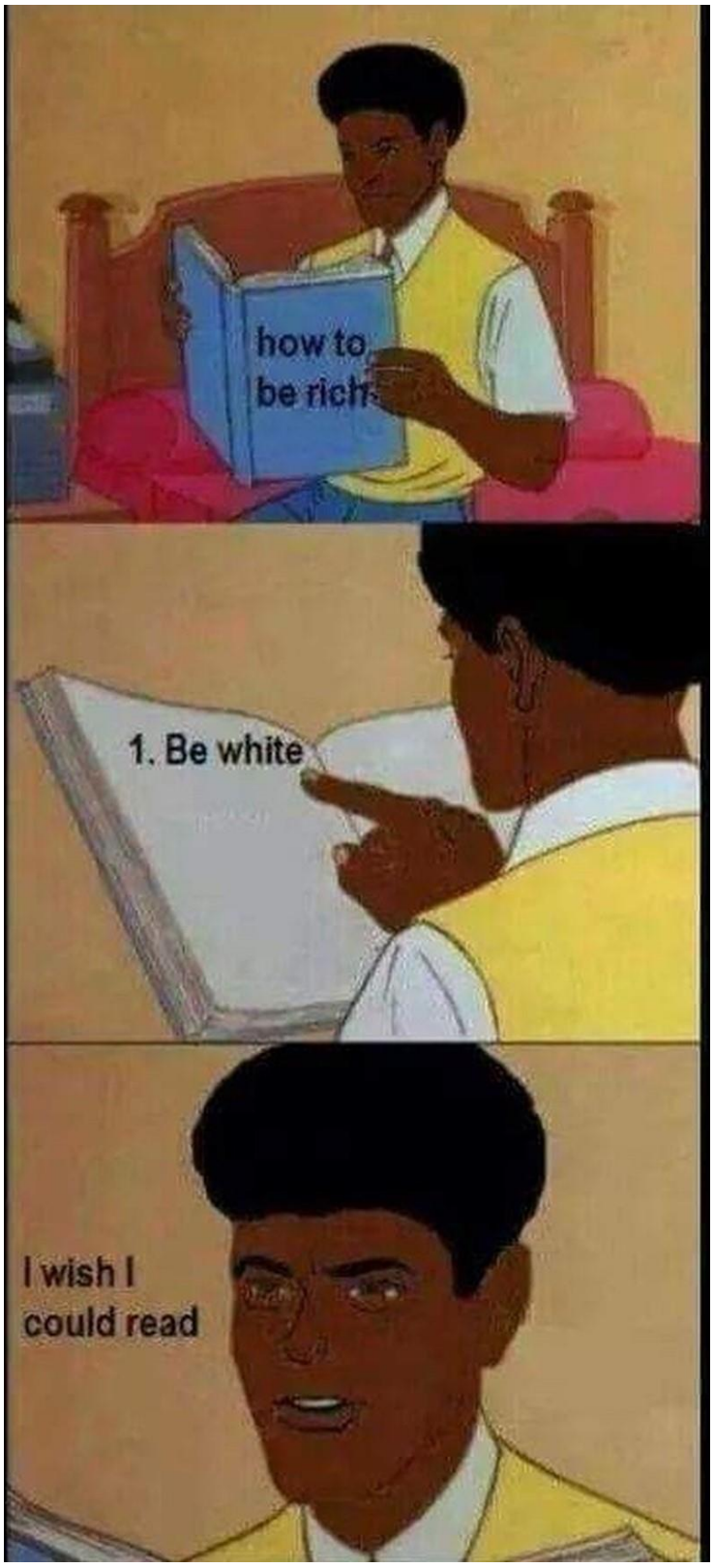


intm2q In your opinion, what was the *intent* behind the creation of this meme?

- ○ Satire (1)
- ○ Offense for the sake of comedy (2)
- ○ Offense for the sake of hurting others (3)
- ○ Expression of genuine personal beliefs (4)
- ○ Not sure (5)

Page Break

intm3a



intm3und Did you understand what information this meme is meant to convey?

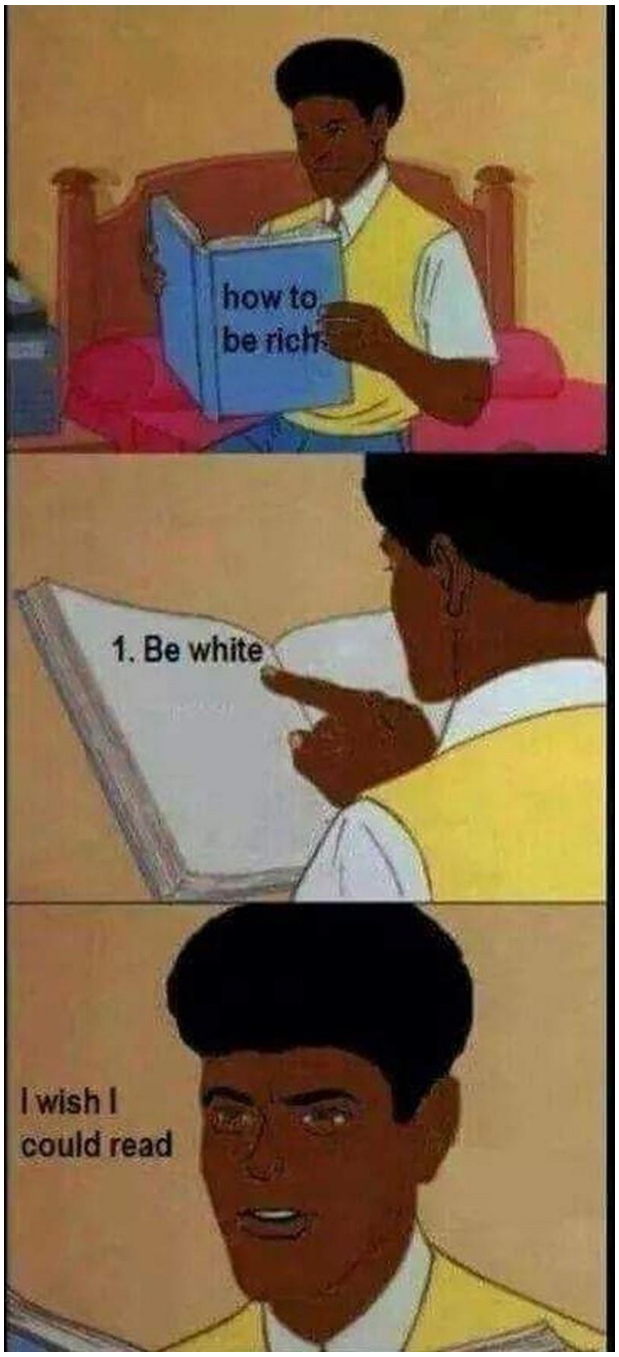
- ○ Yes (1)
- ○ No (2)

Display This Question:

If Did you understand what information this meme is meant to convey? = No

info7d This meme is racist.
It implies that Black people cannot read, and that only White people can achieve financial success.

intm3b



intm3q In your opinion, what was the intent behind the creation of this meme?

- ○ Satire (1)
- ○ Offense for the sake of comedy (2)
- ○ Offense for the sake of hurting others (3)
- ○ Expression of genuine personal beliefs (4)
- ○ Not sure (5)

End of Block: 7) Perceived intent

Start of Block: 8) Debriefing

purposeyn You have now reached the end of the survey. Do you think you know what this research was *actually about*?

- ○ Yes (1)
- ○ No (2)

Display This Question:

If You have now reached the end of the survey. Do you think you know what this research was actually... = Yes

*

purpose Please write a few words about what you think the research purpose of this survey was (max. 280 characters).

_____

Page Break

info8a

info8b Thank you for your participation in my research!

The *actual* purpose of this survey was **not** to examine how people from different backgrounds perceive and interpret offensive and 'edgy' internet memes, but instead to experiment if perceptions could be slightly altered if an idea is first introduced in an offensive meme, then shown again in a different format. If you did not see such memes **prior** to being asked about intent, you were part of the control group.

In order to dispel negative information you *may* have unwittingly internalized after participating in this experiment, feel free to visit the websites of the following NGOs, which work to combat widespread harmful perceptions:

European Network Against Racism:
enar-eu.org

Transgender Europe:
tgeu.org

If you would like to share this survey with someone, please make sure you do not accidentally reveal to them the true research purpose.

Thank you!

End of Block: 8) Debriefing