Giving Formidability a Face

Stig Gulbrandsen

Master Thesis at the Department of Psychology

University of Oslo

14.10.2015

Giving Formidability a Face

2015

Giving Formidability a Face

Stig Gulbrandsen

http://www.duo.uio.no/

Trykk: Reprosentralen, Universitetet i Oslo

Abstract

Author: Stig Gulbrandsen

Title: Giving Formidability a Face

Supervisor: Thomas W. Schubert

Background: Previous research has shown that holding a lethal weapon make people look taller and physically stronger than they actually are, which suggests that height and strength are part of a summary representation of the body (Fessler, Holbrook, & Snyder, 2012). My goal was to extend upon this finding by investigating whether the same is true for the face, focusing on holding a lethal weapon.

Method: In all studies, participants completed a survey containing a two-image forced choice (2IFC). In a 2IFC task, participants complete multiple trials in which they are required to select from two blurry facial images the one that most closely resembles the target category. These data are then processed with a reverse correlation technique. The result is a classification image (CI) that captures the internal mental representation of what the target category looks like. The target category in the pre-studies was the winner or loser of a hypothetical fistfight while the target category in both main studies was the person holding the lethal or non-lethal object. Both main studies included separate independent rating studies in which new participants rated the CIs from the 2IFC tasks in terms of formidability.

Results: The pre-studies confirmed that the classification image of the winner of a hypothetical fistfight looks more formidable than the loser classification image does and the two main studies confirmed that the lethal weapon classification images look more formidable than the non-lethal weapon classification images does.

Conclusion: The results show that the facial features that signal formidability also form a mental summary representation of what facial formidability looks like. The results also show holding a lethal weapon activates the mental summary representation of facial formidability. The practical implication of the mental summary representation of facial formidability is that it may work in such a manner that it is projected onto the face of person holding a weapon, making him look more formidable than he actually is. Future research should investigate if evaluations of facial formidability are based on static or dynamic facial features.

Preface

I would like to extend my sincere gratitude to my supervisor, Thomas W. Schubert, for his support, guidance, and devotion to this project. It has been an exciting, interesting and educational journey.

Contents

1 Introduction: History of violence

The history of humankind is conflict-ridden. Our legacy as a species attests that human beings have a proclivity towards resolving to violence and use of force when settling disputes and conflicts. In order to choose whether to engage in or avoid a potentially agonistic conflict entails an evaluation of the relative formidability of a prospective foe, relative to oneself. Physical size and strength are key determinants of the resource holding potential that constitute formidability (Fessler, Holbrook, and Snyder, 2012). Being able to evaluate the potential interpersonal danger posed by other individuals has thus been crucial to survival. Social by nature, most human beings interact with a multitude of other people on a daily basis in an increasingly population dense society. While the rate and frequency of agonistic conflict vary greatly in different societies, it is much less prevalent than it used to be (Pinker, 2011). In modern day society, being able to recognize formidability is more so crucial to successful functioning and well-being in everyday life. Although being able to recognize and detect formidability is less crucial to survival in modern day society, the adaptive mechanisms designed to detect formidability are still ever so present.

Seldom do we stop and consciously reflect on how and why we perceive someone as formidable. We just intuitively know. Research investigating face perception suggests that in order to adapt, humans have evolved rapid and intuitive mechanisms that do not rely on conscious reflection. Rather, these mechanisms rely solely on conspecifics' attributes (Todorov, Said, Engell, & Oosterhof, 2008). Given that being able to know when to engage or avoid conflict has been imperative to survival, is there any evidence that supports the notion that the cognitive architecture of human beings has evolved to include cognitive mechanisms for assessing physical formidability? Recent research has increasingly focused on representation of formidability to answer such questions.

The anthropologist Daniel M. T. Fessler and his colleagues propose that size and strength constitute a single summary representation of formidability (Fessler et al, 2012). They have for example shown that holding a gun not only makes a person more formidable as a function of the gun but that the gun also makes the person holding it appear larger and stronger (Fessler et al, 2012). In their research on formidability, Fessler and his colleagues have focused on laying bare the range of various factors that influence perception of a foes

relative formidability and in turn how these factors affect the actor's mental representations of the foes physical formidability.

1.1 Mental representation of formidability

Through the course of life, you may at one point or another find yourself in a potentially agonistic conflict with another person. In such a scenario, you will have to assess to what extent this person poses a threat to you. In other words, you must evaluate and decide how formidable this person is. The person you are looking at might appear to be big and strong when he[1] in fact is not. Fessler and his colleagues propose that height, size, and muscularity are not merely physical cues that signal formidability but that they also serve as a summary representation of formidability (Fessler et al., (2012). As to the origin of this summary representation, Fessler and his colleagues argue that because being able to identify a formidable foe could have meant the difference between life and death, human beings developed an internal summary representation of formidability through a process of adaptation and selection.

This system works in such a manner that when it registers cues that signal formidability, the representation of the person displaying these cues makes that person appear taller, bigger, and more muscular than he actually is. However, as mentioned before, the cues that signal formidability are not limited to physical properties of the human body. Factors such as access to weaponry (Fessler et al., 2012) also influence perception of formidability. As such, evaluating formidability entails assessing multiple features extending beyond bodily cues. In turn, "the decision-making involving assessment of multiple features is enhanced by the use of a single summary variable that encapsulates the contributions of these features" (Fessler al., (2012, p. 1).

The two key elements in the conceptual summary representation of formidability outlined by Fessler and colleagues are physical size and strength. As to the origins of this summary representation, the phylogenetic thesis and ontogenetic thesis parallel each other in terms of providing equally plausible explanations (Fessler et al., 2012). As noted by Fessler, they are in fact mutually compatible, as experiences occurring during ontogeny often provide the fundament from which evolved adaptations arise. According to the phylogenetic thesis,

[1] Unfortunately, the research reviewed here has only investigated formidability in men, not women.

the deep antiquity of the contribution of size and strength is an essential factor. It gives rise to the possibility that, as species evolved increasingly complex behavioral repertoires alongside an increase in the variety of factors that influencing relative formidability, size and strength may have come to be employed as the core dimensions of a cognitive representation that summarizes different determinants of relative formidability. This entails that the greater the foe's formidability relative to that of the opponent, the larger and stronger the foe is conceptualized as being, even when the foe's formidability is not derived from actual size or strength, but rather a weapon for example. In this way, size and strength became so important to survival that they eventually came to form the core dimension of a cognitive representation of formidability that summarizes the various determinants (e.g., a weapon) that contribute to formidability (Fessler et al., 2012).

As a result, the greater the foe's formidability relative to oneself, the larger and stronger the foe will be conceptualized as being, even when the foe's formidability does not derive from actual physical size or strength but rather from for example a weapon. This notion can be understood in terms of how a hill slant appear steeper than it actually is if you are carrying a heavy backpack, or how it appears less steep if you are thinking about a close friend (Schnall, Harber, Stefanucci, & Proffitt, 2008). The steepness of the hill does not change. What does change however is our perception and judgment of it and it does so as a function of seemingly unrelated factors. In the same manner, knowing that a person is in possession of a potentially lethal weapon will make the foe more formidable. What does not change though, is the foe's size and strength, even though the foe will appear to be bigger and stronger.

Because the phylogenetic thesis holds that the proposed system that encapsulates representations of formidability is innate, understanding and comprehending the diverse determinants of formidability will in most cases be partially or fully contingent on innate systems that can process input derived from learning (Fessler et al., 2012). On the other hand, the ontogenetic thesis postulates that a wide array of cognitive representations draws on bodily experience and that this often happens without explicit recognition of the relationship between representations and their inherent sources (Barsalou, 2008). By implication, this suggests that representation of relative formidability may be a product of experiences gained from lived events. Children for example, early on learn that bigger often also equals stronger. It is thus likely that through the course of development, size and strength come to play and

essential role in representations of formidability (Fessler et al., 2012). As previously noted, Fessler and colleagues do not believe these two frameworks to be mutually exclusive but rather mutually compatible. The authors themselves do not rule out either option. In fact, they favor a hybrid thesis that largely rests on the premise of an evolved adaptation in which successful function is at least partly dependent on recurrent experiences in the course of development.

1.2 Weapons make the man larger

To test the hypothesis that size and strength constitute a conceptual dimension of a representation used to summarize multiple diverse determinants of a prospective foe's formidability, Fessler and colleagues have conducted several studies. What lies at the heart of all the research Fessler and his colleagues have conducted concerning formidability, is the notion that a range of other factors, independent of size and strength, also influences perceived formidability in terms of size and strength. In the case of their first study, "Weapons Make the Man (Larger): Formidability is Represented as Size and Strength in Humans", Fessler et al. (2012) hypothesized that size and strength constitutes conceptual dimensions of a mental representation used to summarize multiple diverse determinants of a prospective foe's formidability. They argue that, "in humans, weapons are a primary determinant of victory in dyadic violence, and that the modern handgun is prototypic in this regard" (Fessler et al., 2012, p. 2). Based on this premise, they hypothesized that, if representations of a potential foe make use of conceptualized size and strength as a medium for summarizing formidability, enhancing the foe's formidability by "giving" him a weapon should cause the actor's conception of the foe's size and strength to increase.

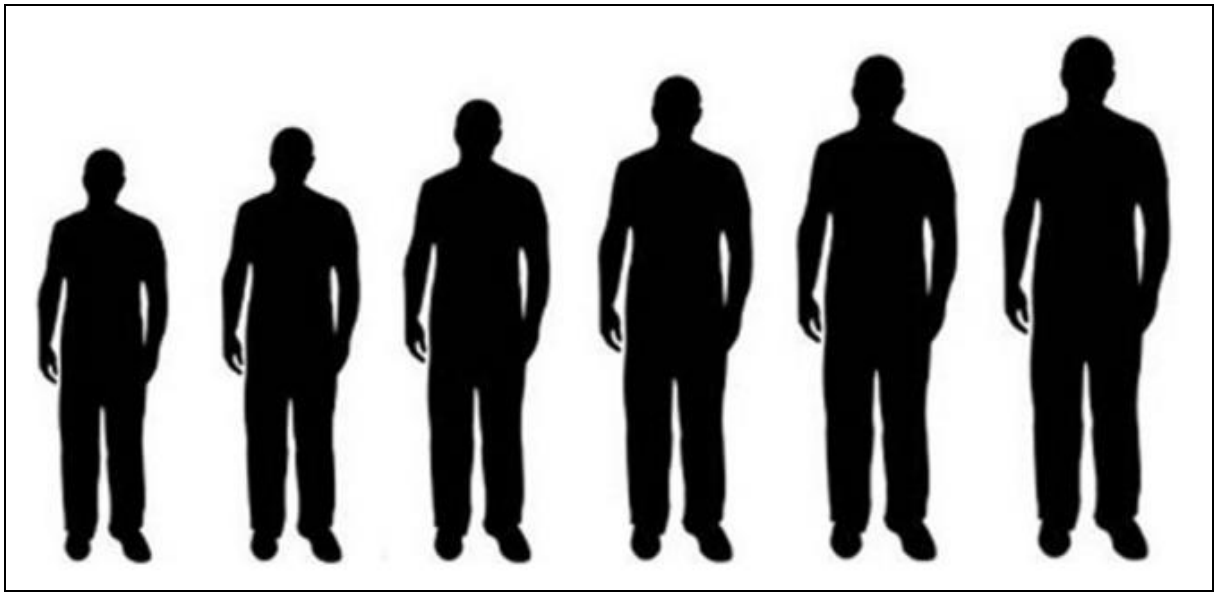


Figure 1. The array Fessler et al (2012, p. 3) gave their participants to provide estimates of size of target individual.

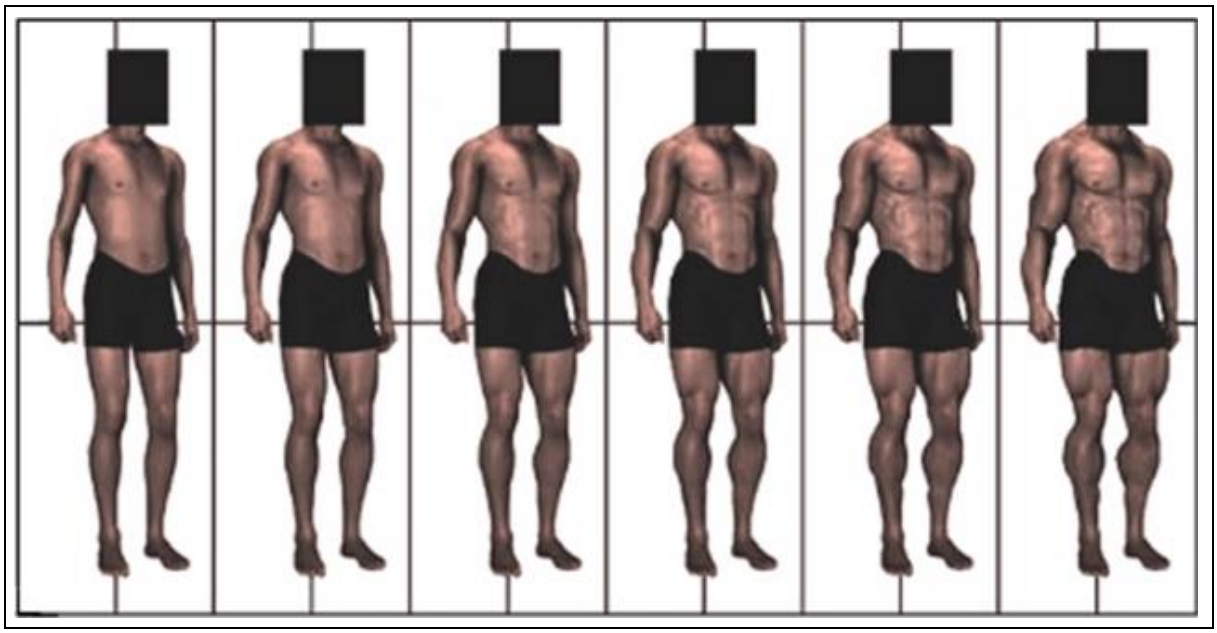


Figure 2. The array Fessler et al. (2012, p. 5) gave their participants to provide estimates of muscularity of target individual.

To test their hypothesis, Fessler et al. (2012) conducted as series of studies where they showed their participants images of hands holding lethal objects (e.g., handgun and a kitchen knife) and non-lethal objects (e.g., a power drill, handsaw, a caulking gun, and a paintbrush). They then had the participants provide estimates of size and muscularity of the person they believed was holding the object by selecting from two sets of arrays. The arrays for size and muscularity can be seen in Figure 1 and Figure 2, respectively. Their results show that the target holding either of the lethal objects was rated as both taller and stronger than the target holding the other non-lethal objects, thus confirming their hypothesis that weapons make the

man larger. In their research, Fessler and colleagues have demonstrated how various other factors influence perception of formidability. They have shown that coalition size reduces perception of formidability by demonstrating that being part of a group (allies) led participants to evaluate the foe as less formidable than when they were alone (Fessler & Holbrook, 2013b). Another study found incapacitation to increase perception of formidability. In this study, they demonstrated that being tied to a chair anchored to the floor caused participants to underestimate their own size and muscularity and conversely, estimate the size and strength of the foe as considerably more formidable than the participants that were not strapped to the chair (Fessler & Holbrook, 2013a). In a series of studies, risk-seeking behavior was found to influence perception of formidability (Fessler, Tiokhin, Holbrook, Gervais, & Snyder, 2014). These studies demonstrated that men who engage in physically risky activities were envisioned to be both taller and stronger and more violent than risk aversive men were. Lastly, Fessler and colleagues has also demonstrated that synchronized walking in synchrony with another person causes people to perceive a purported criminal as less physically formidable than they would if they were walking alone (Fessler & Holbrook, 2014).

The work conducted by Fessler and his colleagues shows that a wide array of features that adds to perceived formidability of a potential foes make him appear bigger and stronger. Features that signal formidability however, are not limited to external factors that add to perceived formidability and physical properties such as size and strength. There is also evidence demonstrating that the human face possess inherent features that signal formidability (Sell et al., 2009; Zilioli et al., 2014). As previously noted, a way to conceptualize formidability is to think about it in terms of fighting ability. Navigating social environments naturally entails scanning the surroundings for potential threats. More often than not, this entails keeping an eye out for people that that might pose a threat to us. Our social nature has made us highly attuned to reading faces. Considering that faces tend to stick out in crowd (Hershler & Hochstein, 2005), they provide a natural source of information. One way to identify whether people pose a threat to us or not entail looking for displays of emotional expressions that signal behavioral intent. In terms of potential threat posed by other humans, we tend to look for displays of aggression and dominance because their corresponding facial features serve as attack expressions (Montepare & Dobish, 2003). As human beings evaluate a prospective foe's formidability as a means of deciding whether to fight or flee, another way of conceptualizing a foe's formidability is in terms of fighting ability. In this regard, displays of aggression and dominance are important cues as they combined serve as attack expressions

(Montepare & Dobish, 2003). Such findings highlight how facial features are highly important because the forecast behavioral intent.

1.3 Facial cues to formidability

Recent evidence suggests that the human neurocognitive architecture contains mechanisms that assess formidability from cues found in human faces (Sell et al., 2009; Zilioli et al., 2014). In their line of research, Aaron Sell and colleagues conceptualize formidability in terms of fighting ability. The researchers argue that, in social species such as humans where aggressive social interactions occur, natural selection typically organizes adaptations designed to enhance the organisms' capacity to inflict damage, and that for ancestral humans the single most important factor driving the differential ability to inflict cost has been upper-body strength. To support their hypothesis, they argue that a number of factors suggest that selection should have tailored strength assessment specializations to use information present in the face alone. They rationalize that, under ancestral conditions clothing, other people, vegetation, and other obstructions would conceal the upper body. If the face also exhibited cues of strength, then this would have provided a separate channel for assessing strength when direct visual assessment of musculature relevant to fighting ability was not possible. Considered in the light of how the brain is known to contain neurocomputational specializations designed to extract dynamic and static social information from faces (e.g., identity, eye direction, emotional state, sex, age), they propose that the ability to detect strength coevolved with the rest of face processing.

The research conducted by Sell and colleagues thus far, provide compelling evidence in support of their hypothesis. In one of their earliest studies (Sell et al., 2009), the researchers demonstrated that people can accurately estimate strength and fighting ability of male targets from photos of their face (and body). A noteworthy finding is that even when only the target face was visible, their subjects were extracting cues of strength that were largely independent of height, weight, and size. Rather, the cues of strength that they used corresponded most strongly to objective measures of upper-body strength (e.g., bench press). Their findings demonstrate that not only do judgments of strength and judgments of fighting ability track each other, but that they track actual upper-body strength. While these findings demonstrate that humans could accurately assess strength and fighting ability from the face, they did not address which facial cues inform these judgments.

In a later study, drawing on research by Windhager, Schaefer, & Fink, (2011) showing how a broad middle face, a widened region between the eyes and a rounded outline (well-curved jaw line and lower forehead) predicted actual strength and perceived masculinity among young men, Zilioli et al., (2014) hypothesized that facial width-to-height ratio (fWHR) is a cue to formidability. In line with their hypothesis, they found that fWHR co-varied with actual fighting ability. Based on the records of Ultimate Fighting Championship fighters and their respective measure of fWHR, they found that fWHR predicted the fighters' number of wins. Their results also revealed that assessments of formidability co-varied with fWHR on both natural faces and computer-generated faces of strong (wide face) and weak faces (narrow face). While their findings demonstrate that fWHR predict fighting ability, it does not explain how or why.

Research showing that fighting ability is associated with physical aggression and success in interpersonal conflict (Sell, Tooby, & Cosmides, 2009; von Rueden, Gurven, & Kaplan, 2008) provide some clues to how fWHR is associated with fighting ability. Sell et al. (2014) reason that if anger functions as a bargaining system in humans, then humans should have evolved to deploy facial morphology in a way that enhances these cues during aggressive bargaining. To test this notion, Sell and colleagues manipulated the seven key facial muscle movements that constitute the anger face so that each target face displayed only one of the key anger movements at a time. In support of their hypothesis, participants rated faces containing any of these key anger movements as physically stronger than the neutral control images. This finding lead the authors to conclude that through a process of selection and adaptation, the constellation of features that comprise the human anger face was selected to enhance cues of physical strength during agonistic bargaining. Their findings therefore demonstrate that the human anger face is not merely a set of arbitrary features evolved to signal aggressive intent but also to enhance cues of strength.

Taken as a whole, the research on facial cues that signal formidability provides a compelling argument supporting the notion that selection pressures has provided humans with a neurocomputational system specialized to detect facial cues that signal formidability. Their research shows that people accurately estimate fighting ability from facial cues and that fWHR is one of the facial features by which they do so. Their research also suggests that facial displays of anger are strategically employed to not only signal intent to cause harm but also enhance cues of strength, signaling ability to cause harm. In an untested hypothesis,

Zilioli et al. (2014) propose that the finding that fighting ability is associated with physical aggression and success in interpersonal conflict suggests that perhaps fighting ability plays a mediating role in the association between fWHR and aggression. If this is the case then, "the relationship between fWHR and aggression is most likely due to the fact that fWHR predicts fighting ability and that those males with greater fighting ability displays more anger and aggression" (Zilioli et al., 2014, p. 7).

1.4 Does formidability have a face and what does it look like?

Thus far, the research by Sell and colleagues have centered on upper body-strength in males, anger, fWHR, and fighting ability. Might these findings also provide some clues to what a formidable face looks like? Sell and colleagues work on fighting ability and its associated features has thus far focused on men. The reason is that the male gender is the sex that shows the strongest evidence for combat design (Sell, Hone, & Pound, 2012). In addition to facial width, male fighters with deep-set eyes had higher combat success (Trebicky, Havlícek, Roberts, Little, & Kleisner, 2013). Recent research by Toscano, Schubert and Sell (2014) found that judgments of dominance was predicted by strength and that brow height, eye and chin length, and width of nose and mouth constitute the common predictors that underlie perceptions of both strength and dominance. Furthermore, other sexually dimorphic features with male bias include prominent cheekbones, wider noses, lower brow ridges, and larger chins (Lefevre et al., 2012; Penton-Voak et al., 2001; Trebicky et al., 2013) . While this research provide some insight as to what a formidable face might look like the only way to ascertain for sure is to investigate within a suitable framework.

1.5 Reverse correlation

Psychological reverse correlation methods comprise a set of data-driven techniques that allow for unconstrained visualization of the information contained in the face that is diagnostic for social perception (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2014). Reverse correlation (RC) techniques represent a rather new approach to face perception that has gained momentum in recent years. One of the strongpoint of this method is that it allows researchers to tap into (probe) and investigate peoples' internal representations of categories without making a priori assumptions about what those internal representations might look like

(Dotsch, Wigboldus, & Knippenberg, 2013).

Another strongpoint is their ability to capture the variance in facial structure that leads to specific social attributions (Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2014). Prior research employing RC techniques have proven strong in assessing cultural differences in the perception of emotional expressions (Jack, Caldara, & Schyns, 2012), expected facial appearance individuals as well as group members (Dotsch & Todorov, 2011; Gosselin & Schyns, 2003), and biases in the representation of social categories (Karremans, Dotsch, & Corneille, 2011), to name a few. As Dotsch and Todorov (2012, p. 569) put it; "The possibilities for laying bare internal representations are endless".

1.6 Two-image forced choice

While there are several different RC techniques, I will outline and describe an image classification task called 'two-images forced choice (2IFC)' variant (used by Dotsch, Wigboldus, Langner, & van Knippenberg, 2008), as this is the method that I will adopt and employ in the research contained within this thesis. In a typical 2IFC task, participants judge noisy images of faces. The noise is created by superimposing random noise on a base face image. The random noise distorts the face at the pixel level, creating variation in the faces, making them appear slightly different.

The process of superimposing random noise patterns also generates negative patterns (the mathematical opposite). This entails that each pixel that is dark in the original noise pattern is consequently bright in the negative noise pattern. In effect, each single trial comprises two version of the base image presented side by side. While one of them is the base image with the original noise, the other one is the base image with the negative noise. The purpose of the 2IFC RC technique is to compute a classification image (CI) that encapsulates a visual representation of peoples' criteria for classifying a face as belonging to a specific category, thereby showing the facial characteristics that drive social judgments.

1.7 Classification images

The 2IFC creates two types of CIs: individual CIs and total average CIs. The individual CIs are unique to each individual and as such, the number of individual CIs depends on the sample size (i.e. one CI is created for each participant). The individual CIs are

generated by collapsing the noise pattern from each image selection trial. The total average CIs vary with the number of selection criteria (i.e. one selection criteria will give one total average CI). The total average CIs are generated by collapsing the noise pattern from each trial across all participants. By convention, the next step is to quantify the typicality of the resulting CIs by presenting the CIs to an independent sample pool for rating. In case of the current research, participants will rate the CIs in terms of perceived formidability.

1.8 Overview of the current research

Considering the theoretical background and evidence reviewed thus far pertaining to mental summary representations of formidability and facial features associated with fighting ability, there is a possibility that there may also be specific facial cues that not only signal formidability but also come together to form a summary representation of facial formidability. In light of these findings, it is hypothesized that activation of this representation causes a person to conceptualize the face of a prospective foe as looking more formidable than the person might actually be. The aim of the current study is to investigate whether factors that adds to relative formidability (e.g., weapons) of a person makes other people conceptualize the face of that person as more formidable.

In their research, Fessler and colleagues have demonstrated the existence of a summary representation of formidability whereby size and strength constitute a summary representation of formidability. The majority of research investigating facial perception in relation to formidability is largely limited to the work conducted by Sell and colleagues. Their research shows that people rather accurately estimate fighting ability (formidability) based on fWHR and that this is most likely a result of people with higher fWHR displaying facial features that also signal aggression. Their research is thus limited to facial cues that signal physical formidability. Unlike Fessler, they have not explored the possibility that the facial features that signal formidability also constitute a constellation of features that form a summary representation of facial formidability.

Among the factors that influence perception of formidability, Fessler et al (2012) has shown that being in possession of a lethal weapon causes people to envision that person as more physically formidable than people in possession of objects of less lethal affordances. Is it plausible that the facial features that signal formidability has come to constitute a facial

summary representation of formidability? Is it possible that having the knowledge that a person is in possession of a lethal weapon not only makes that person appear bigger and stronger (e.g., more formidable) but also that it makes his face appear more formidable? To the best of my knowledge, no research to date has attempted to explore whether objects of lethal affordances enhance facial cues that signal formidability. If this were to be the case then this finding would be a first step towards identifying how people conceptualize facial formidability. The present study will therefore investigate this notion by testing the effects of various objects on facial representation using a reverse correlation technique.

The first port of call is to test the suitability of the two-image forced choice reverse correlation method and discover whether this method is capable of capturing a facial representation of a formidable face. A second port of call is to identify exactly how many trials are required to create a satisfactory classification image. Fessler and colleagues research on formidability defines it in terms of relative resource holding potential, or the ability to inflict cost. Sell's research summarizes this definition as relative fighting ability. The evidence obtained from research by Sell and colleagues suggests that humans possess neurocognitive adaptations that specifically evolved for assessing opponents fighting ability from the face

I will therefore begin by testing the suitability of the reverse correlation method in two separate pre-studies. In the first pre-study, I will present to the participants a classic 2IFC task and ask them to select the facial image they believe most likely to be the winner of a hypothetical fistfight. Conversely, in the second pre-study, I will present participants with the same 2IFC task, only this time I will ask them to select the facial image they believe to the loser of the fistfight. The loser CI from the second pre-study will serve as control, allowing me compare the CI from each experiment to determine if the winner CI indeed appears more formidable than the loser CI does. By framing the instructions in this manner, participants will conceptualize the winner of a fistfight as a fighter, thus allowing me to not only test the reverse correlation method but also Sell's notion that the human faces possess features that evolved to enhance cues of formidability, conceptualized as fighting ability.

In the first main study, I will explore the notion that weapons make the face look more formidable. By showing people images of lethal weapons and have people estimate the height, size, and muscularity of the person people believed to be the owner of the weapon, Fessler and colleagues demonstrated that weapons led people to perceive the person in

possession of a weapon as larger and more muscular than people holding non-lethal objects (Fessler et al., 2012). Research by Sell et al. (2009) has demonstrated that people can accurately assess physical upper body strength from cues in the face. It is therefore possible, given that the facial features that signal formidability also constitute a summary representation of facial formidability, that similarly showing a lethal weapon next to facial images in a 2IFC task, participants will envision the face of the person they believe is holding the weapon as more formidable.

Fessler and colleagues also made some interesting observations concerning the selection of object stimulus that is worth noting (Fessler et al., 2012, Study 4). Deciding which objects to use as stimuli requires some careful consideration. Stereotyping and schema activation may be problematic insofar as participants may infer properties about the individual based on the object stimulus. Handguns are rather uncommon in Norway and therefore more likely to be associated with unwanted stereotypes. Unlike guns, a kitchen knife is a common household object almost equal to a handgun in terms of lethal affordances (Fessler et al., 2012, Study 4). I therefore opted to use a kitchen knife as the lethal object and a pencil as my "neutral" object, as this is an everyday object not typically associated with any type of person in particular.

In the second main study, I will aim to extend upon the findings from the first main study by replicating the experiment using a US sample population with a handgun as the lethal object. The pencil will remain the neutral object. Considering the novelty of this study, it is important to solidify the core findings of the original study. Because I am drawing heavily upon Fessler's research on formidability, in particular the finding that guns make people appear more formidable, it is only natural to attempt to replicate and extend upon Fessler's findings by demonstrating that guns also makes the face appear more formidable. Both Sell and Fessler has produced evidence in support of their findings having cross-cultural validity. Thus, demonstrating that weapons make the face look more formidable using a US population will provide the findings with cross-cultural validity.

Hypothesis 1: People will conceptualize the face of a winner in a fistfight as more formidable than the face of the loser of the fistfight.

Hypothesis 2: Imagining that an individual possesses a knife will lead observers to conceptualize the face of that individual as more formidable than individuals who possess only a pencil.

Hypothesis 3: Imagining that an individual possesses a gun will lead observers to conceptualize the face of that individual as more formidable than individuals who possess only a pencil.

2 Method

2.1 General method

With exception of the two pre-studies, each study comprised two experiments (e.g., 1a & 1b). The first experiment of each study entailed a two image forced choice (2IFC) task. The second experiment in each study entailed having an independent sample of participants rate the CIs produced by the 2IFC task. Whereas study 1 and 2 did include an independent sample of participants rating the CIs from the 2IFC task, the two pre-studies did not.

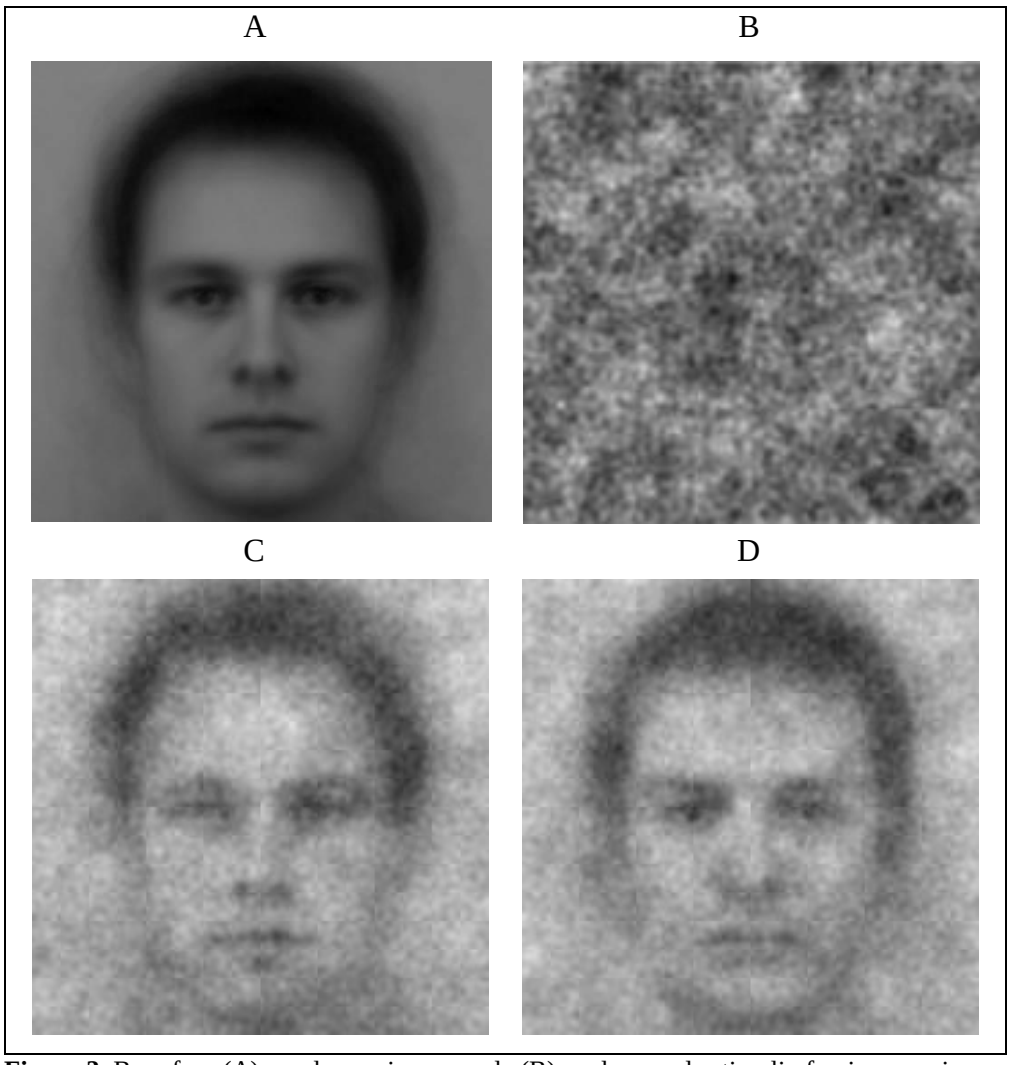


Figure 3. Base face (A), random noise example (B), and example stimuli of noise superimposed on a single base image (C and D). The left stimulus (C) shows the base image with original noise superimposed and the right stimulus (D) shows the base image with the negative noise superimposed. The random noise example (B) is borrowed from Dotsch & Todorov (2012, p. 563).

2.1.1 Stimuli

All the stimuli in the 2IFC task consisted of the same base face with randomly generated noise superimposed on each image in every trial (Dotsch, 2014). The base face used to generate the stimuli images was a grayscale average (morph) generated from more than 2000 male faces from the Karolinska Face Database (Lundqvist, Flykt, & Öhman, 1998; see Figure 3A). The superimposed noise consisted of truncated sinusoid patches of 2 Cycles in 6 Orientations (0˚, 30˚, 60˚, 90˚, 120˚, and 150˚) x 5 Spatial scales (2, 4, 8, 16, and 32 cycles per image) x 2 Phases (0, л/2), with random contrasts (see figure 3B). In sum, the random noise was a function of 4,092 parameters, each defining the contrast value of one truncated sinusoid spanning two cycles. Stimulus size was 512x512 pixels (Adopted from Dotsch & Todorov, 2012, p. 564).

2.1.2 Materials

In all studies, the online surveys used to present the stimuli and collect data was created using Qualtrics Research Suite Software.

2.1.3 Procedure

In all of the experiments, participants' received a web link to the survey. For each trial, presentation of the two facial stimuli images was either side-by-side (pre-studies) or on top of each other (Experiment 1a and 2a). The presented facial stimuli pair was randomly selected from a set of 150 (385 in the pre-studies) pairs of original (positive) and inverse (negative) noise patterns (Figure 3C and 3D, respectively). The on screen placement (negative noise on the left or top vs. negative noise on the right or bottom) of face images with either original or negative noise was also randomized and counterbalanced. Because the random superimposed noise pattern for each image pair is the exact mathematical opposites of each other, one base image had the positive random noise superimposed on it (Figure 3C), while the other base image had the negative random noise superimposed on it (Figure 3D). The procedure followed Dotsch and Todorov (2012) whereby the use of negative noise as opposed to another random noise pattern is done in order to maximize the differences between the two presented images. It also serves to minimize the number of possible stimulus pairs presented, and to simplify data analysis.

2.1.4 Data processing

To compute the classification images (CIs), a script produced by Ron Dotsch was run in R software (Dotsch, 2014). The script calculates the mean of all the noise patterns that the participants selected by averaging the parameters on which those noise patterns are based. As a result, it computes 4,092 mean parameters per participant. It then averages the mean parameters across participants for each cell of the design and generates the classification patterns based on all cell average parameters. Finally, the classification patterns are superimposed on the original base image to generate the CIs.

2.1.5 Data screening and exclusion criteria

For the two main studies, the data were screened for overt suspicion regarding the hypothesis and frivolous responses prior to analysis. In particular, participants who spent less than 10 minutes on the survey were excluded, as more than 10 minutes was deemed necessary in order to provide satisfactory response. Anything less than 10 minutes spent was thus deemed to signal frivolous responses. Comments left by the participants' regarding suspicions concerning the hypothesis or lack of attention (e.g., "boring", "pictures all look the same" etc.) served as means of exclusion. Data exclusion was done before analyzing.

2.2 Pre-studies 1a and 1b

2.2.1 Participants

An initial sample of 16 participants in Experiment 1a were Norwegian and recruited through social media networks. Eight (N = 8) participants did not complete the survey and was therefore excluded from further analyses. The final sample comprised four men and four women (N = 8) between the age of 24 & 65 years (M = 38.25, SD = 16.57).

An initial sample of 18 participants in Experiment 1b was Norwegian and recruited through social media networks. Twelve (N = 12) participants did not complete the survey and was therefore excluded from further analyses. The final sample comprised two men and four women (N = 6) between the ages of 24 & 46 years (M = 32.17, SD = 7.57).

2.2.2 Design

Participants in each pre-study performed a 2IFC task in which they were presented with two facial images and instructed to select the face that they believed would be the winner (Pre-study 1a) or loser (Pre-study 1b) in a hypothetical fistfight. Due to the exploratory and stepwise nature of the investigation, the pre-studies were only a between groups design in the sense that each group of participants was either recruited to partake in separate "Winner" or "Loser" pre-study.

2.2.3 Procedure

All participants in Pre-study 1a and Pre-study 1b received a Facebook message invitation containing a web link and an invitation to partake in the pre-study. Pressing the web link redirected participants to the Qualtrics hosted survey. Participants did not receive any information pertaining to the nature of the study. Prior to the task, the instructions stated that they would always see images of two individuals and that the images would be intentionally noisy. Participants in Pre-study 1a were instructed to always select the face of the person they believed would win against the other in a hypothetical fistfight. Conversely, participants in Pre-study 1b were instructed to always select the face of the person they believed would lose against the other in a hypothetical fistfight. Additionally, instructions urged them to choose intuitively and not spend too much time evaluating their deCIsions and that they could take breaks and come back to the task when desired. There was no screening of participants prior to data processing in either pre-study.

2.2.4 Data processing

Computation of the CIs followed the method outlined in the general method section.

2.2.5 Results

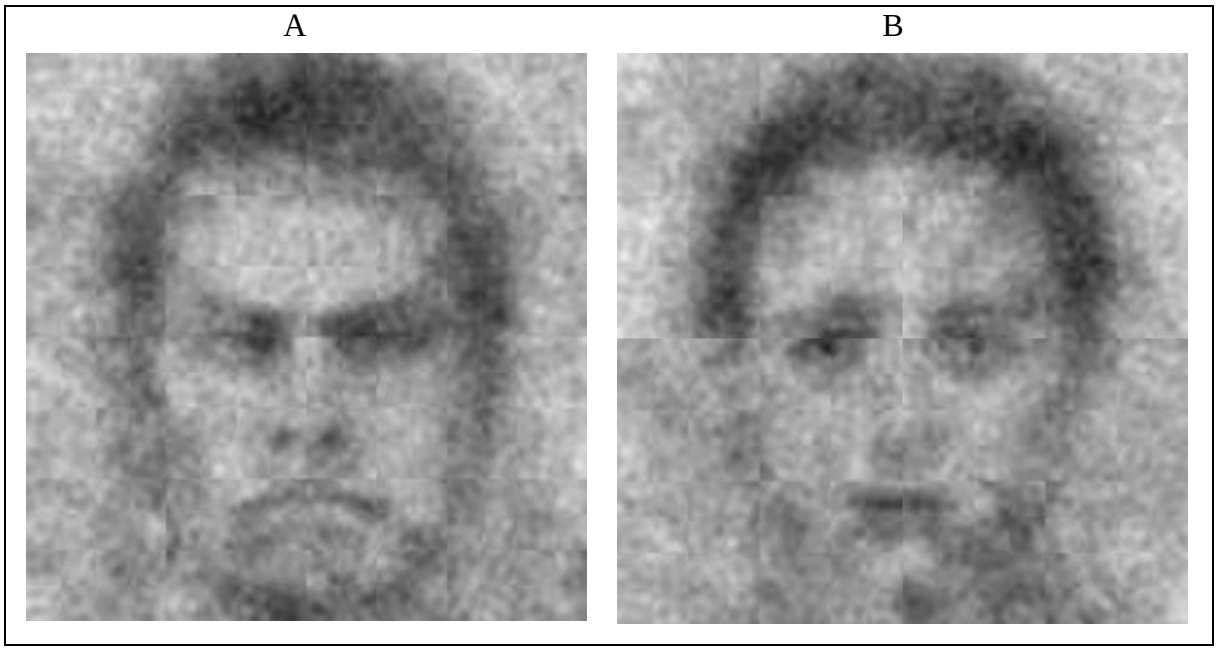


Figure 4. CI from Winner condition (A) and the CI from Loser condition (B) in Pre-study 1a and Pre-study 1b, respectively.

Figure 4 shows the CIs from the Winner (A) and Loser (B) conditions in Pre-study 1a and Pre-study 1b, respectively. Visual inspection of the CIs show that the formidable face from the winner condition display a masculine and square shaped face with a wide nose bone and a square-shaped forehead with a slightly receding hairline. The eyes are deep set and appear cold and angry-looking with upward angled eyebrows. The jawbones appear distinctly muscular, the chin wide and solid, and the mouth downturned with closed thick lips. On the other hand, the weak face from the loser condition displays a childlike and more rounded face with a narrow nose bone and rounded forehead. The eyes are more protruding and appear sad with downward facing eyebrows. The chin is rounded and the mouth is small and slightly open with pouting lips. Because the results were obvious and this was only a pilot test, I did not submit these images to another rating task by an additional sample.

2.2.6 Discussion

Visual inspection of the CIs from the winner and loser conditions clearly show that the winner CI looks considerably more formidable than the loser CI does. These results demonstrate that participants conceptualize a fighter as more formidable than a loser of a fight. More specifically, the results suggests that thinking about someone in terms of fighting

ability activates a facial summary representation of formidability that enhance the facial features in that persons face, making him look more formidable than he really is.

2.2.7 Reducing number of trials

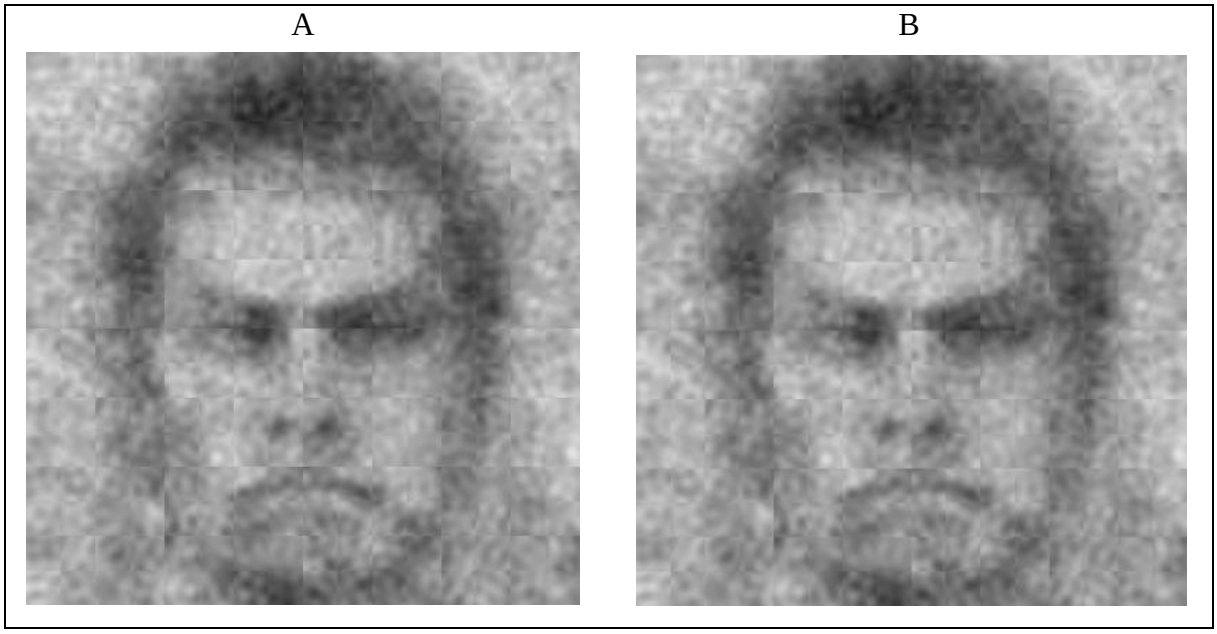


Figure 5. Formidable CI based on 100% of 385 trials (A) and formidable CI based on 40% of 385 trials (B).

The second objective of these two pre-studies was to determine the number of trials needed to produce satisfactory CIs. As expected, feedback from participants confirmed that 385 trials were mentally straining and made the task lengthy and tedious to complete[2]. This was undesirable and of particular concern. Lack of focus may lead participants to pay less attention when making their judgments, which in turn may corrupt the data. To address this concern I randomly selected 80, 60, 40, and 20% of the trials and re-ran the analysis. Visual inspection of the consequent CIs showed that using 40% (154 trials) yielded a satisfactory CI (Figure 5B) that visually did not differ substantially from the original CI (Figure 5A) based on 385 trials. The only difference being that the original CI was slightly sharper.

[2] Which would explain the large number of participants who failed to complete the pre-studies.

2.3 Experiment 1a

2.3.1 Participants

An initial sample of 51 students from the University in Oslo participated in the experiment as part of their course requirement. Based on previously outlined exclusion criteria eight participants were excluded from the dataset. The final sample comprised 43 students including 7 men and 36 women between the ages of 19 and 54 years (M = 25.19, SD = 9.56). The ethnicity was 83.7% White/Caucasian, 9.3% Asian, and 7.0% other or mixed ethnicity. All participants were Norwegian.

2.3.2 Design

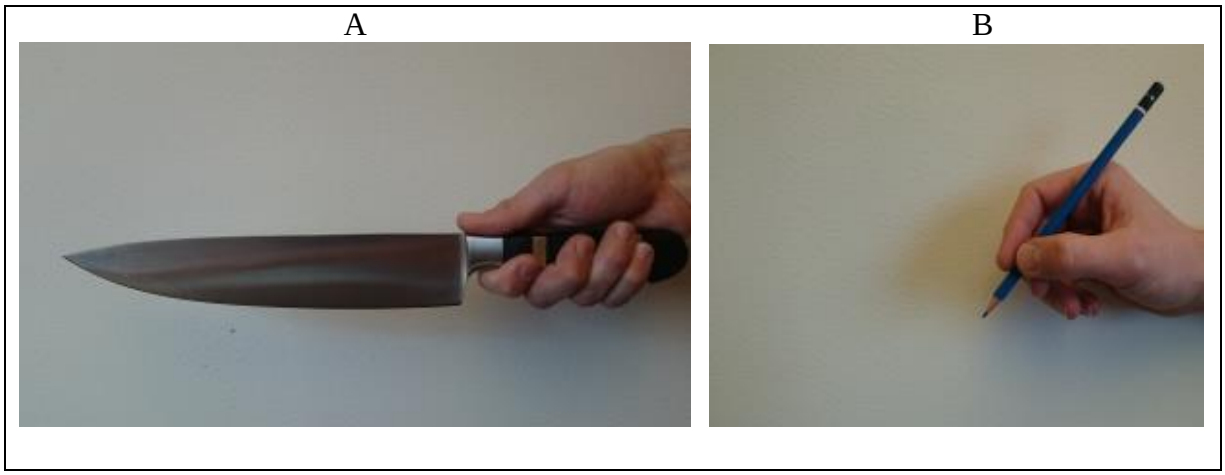


Figure 6. Knife (A) and pencil (B) stimuli.

In a between subject design, participants performed a 2IFC task in which they were presented with two facial images placed on top of each other with a third image depicting a hand holding an object (knife versus pencil) placed to the left of two facial images (Figure 6). Participants were randomly allocated to either the knife or the pencil condition. Instructions informed them that their task was to always select the facial image they believed was holding the object depicted next to the facial images.

2.3.3 Procedure

Prior to beginning the 2IMFC task, participants were informed that in the following task they would always see pictures of two individuals, that the images were intentionally made noisy, and that a third image of an object would be presented next to the individuals.

Instruction then stated that their task was to always select the person they believed was holding the object. Participants made their selection by clicking on the image with their mouse cursor or by pressing the up or down button on their keyboard. Additionally, instructions urged them to choose intuitively and not spend too much time evaluating their deCIsions. In each trial, the two facial stimuli images were presented on top of each other with the object stimuli always presented to the left of them. The object always faced away from the facial stimuli images. All participants completed 150 trials of image pair selection. The presented stimulus pair was randomly selected from the same set used in the pre-studies. Twenty randomly selected trials were presented twice to test reliability of answers (not used in this thesis).

In addition to the 2IFC task, the survey also consisted of demographic questions, including items addressing age, gender, ethnicity, country of residence, and participant height. Additionally, participants were asked to think about the person holding the object and to provide a brief description of what they thought that person was like. They were also given the option to leave comments about the survey, if they had any. Before starting, participants indicated their consent. Following the study, participants were thoroughly debriefed and informed about the hypothesis of the study.

2.3.4 Materials

Some alterations were made to the materials based on findings from the pre-studies. It was discovered that the superimposed noise on the facial images made it difficult to discern the facial features in the images when sitting close to the screen. Given the nature of the task and the importance of being able to distinguish the details in the images, the logical thing to do was to rectify this effect by reducing the image size, thus making the images more distinguishable from each other. The size of the facial stimuli images was therefore reduced from 512x512 pixels to 256x256 pixels.

Secondly, in order to test the hypothesis that weapons make the face look more formidable, I needed to add images of the lethal and non-lethal objects I wanted to use. The object stimuli added to this experiment included an image of a hand holding a kitchen knife as the lethal object as well as an image of a hand holding a pencil as the non-lethal object. The kitchen knife was chosen as the lethal object because guns are less prominent in the Norwegian society as well having less stereotype connotations attached to it. A kitchen knife

was also successfully employed as a lethal object by Fessler et al. (2012) and found to be rated equally lethal to a handgun (Study 2). The pencil was chosen as a non-lethal object because of its objective and neutral properties.

A white Caucasian male was used as a hand model for presentation of both stimuli. The images was set against a neutral white background and the images were cropped by the wrist so not to disclose any details pertaining to the physique (e.g., strength, size) of the model holding the objects. The images were in color and the objects always held by the right hand so to face away from the facial stimuli images.

2.3.5 Data processing

The data processing followed the procedure outlined in the general method section.

2.3.6 Results

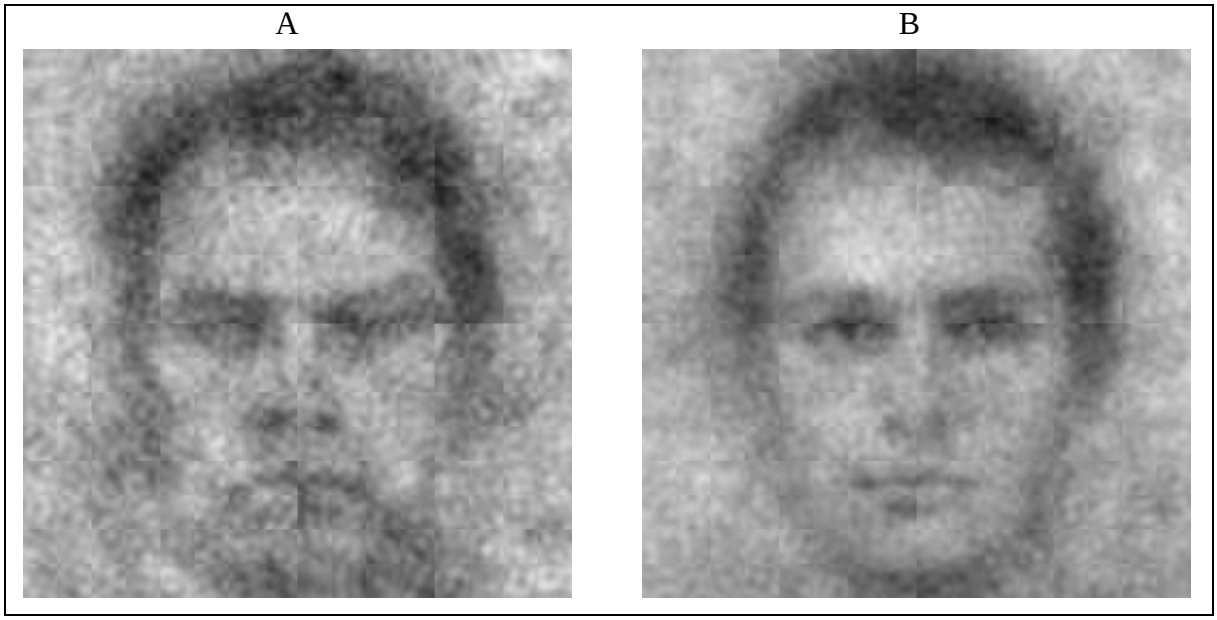


Figure 7. CIs from knife condition (A) and pencil condition (B).

Figure 3 displays the CIs from the "knife" and "pencil" conditions. Visual inspection of the CIs show that the formidable face from the knife condition (Figure 7A) displays a square and masculine facial shape with deep set, cold, cynical, and almost lifeless-looking eyes. His eyebrows appear thick, distinct, and are angled upwards. He has high and distinct cheekbones, sagging cheeks, slightly downturned mouth with thick closed lips, wide and solid chin, wide nose-bridge bone, and a square shaped forehead with a slightly receding hairline.

On the other hand, the CI from the pencil condition (Figure 7B) displays a more rounded face, warm and friendly looking eyes, less distinct and more flat eyebrows, rounded chin, closed mouth with soft and puffy-looking lips, smaller and narrow nose, rounded jaw, and a more rounded forehead with a slightly receding hairline.

2.4 Experiment 1b – Rating of CIs

2.4.1 Participants and design

An initial sample of 54 students from the University in Oslo participated in the experiment as part of their course requirement. Based on exclusion criteria, 10 participants were excluded from the dataset, leaving a final sample of 44 students. The final sample thus comprised 13 men and 31 women between the ages of 19 and 50 years (M = 22.82, SD = 6.34). The ethnicity was 84.1% White/Caucasian, 2.3% Hispanic, 11.4% Asian and 2.3% other or mixed ethnicity. 95.5% was Norwegian, 2.3% was Danish, and 2.3% undisclosed. In a within subject design, all participants performed an image selection task and an image rating task.

2.4.2 Procedure and materials

The stimuli comprised the total average and individual CIs generated by the participants in Experiment 1a. In the first task, participants were randomly presented with either the knife and pencil total average CIs generated by the participants from study 1a. The total average CIs from each condition were presented side by side and participants were instructed to select the person they believed would win against the other in physical fistfight. Placement of the images on left vs. right side was counterbalanced. Prior to the task they were informed that the images would be noisy and that this is intentionally so. In the second task participants were presented with the individual and total average knife and pencil CIs generated by the participants in Experiment 1a. The participants were informed that in the following task they would be presented with intentionally noisy images and that their task was to rate them in terms of perceived formidability on a 5-point Likert scale (1 = not at all formidable, 5 = very formidable). The CI images appeared individually and in random order.

2.4.3 Results

A series of analyses were done to test the hypothesis that being in possession of a knife makes the face of the person holding it look more formidable. Firstly, a binomial test was conducted to test if participants more often selected the average CI from the knife condition as the winner of the hypothetical fistfight compared to the average CI from the pencil condition. Inspection of frequencies show that of the total 44 participants, only 7 chose the average CI of the pencil condition as the more formidable person, while 37 (84.1%) chose the average CI of the knife condition. The binominal test, confirmed that significantly more participants chose the average CI from the knife condition as the winner of the fistfight, Z = 4.37, $p < .001$.

A paired-samples t-test was conducted to compare the ratings of formidability of the average CIs from the knife and pencil conditions. There was a significant difference in the formidability ratings of average CIs from the knife ($M = 3.40$, $SD = .54$) and pencil ($M = 2.90$, $SD = .41$) conditions, $t\ (43) = 6.74$, $p < .001$. A second paired-samples t-test was conducted to compare the averaged ratings of formidability of the individual CIs from the knife and pencil conditions. There was a significant difference in the formidability ratings of individual CIs from the knife ($M = 4.11$, $SD = 1.21$) and pencil ($M = 2.64$, $SD = 1.01$) conditions; $t\ (43) = 5.45$, $p < .001$. These results suggest that a gun makes a person look more formidable. Specifically, our results suggest that knowing that a person is holding a lethal weapon makes his face appear more formidable.

In the previous analysis of the individual classification images, I averaged ratings provided by the participants in the rating studies across the classification images. As a result, the analysis was done with participants of the rating study as the unit of analysis. An alternative approach would be to average ratings across participants of the rating study, and analyze data with the participants of the original study as the unit of analysis. In the literature on reverse correlation in general and the method used here in particular, both approaches can be found, and it is not clear which is preferable, although both should lead to comparable results if the statistical power of both the reverse correlation and the rating study are high.

However, an alternative approach would be to neither aggregate across CIs nor across raters but to use a multilevel approach instead. Using multilevel models, judgments provided in the rating study are the unit of analysis, and shared variance due to same CI and same rater is modeled as random factor. I conducted such an analysis. Using MIXED in SPSS, I set up a

multilevel model of using all 1892 (44*43) judgments from the rating study. Judged formidability was the dependent variable. ID of the CI (e.g., participants in the reverse correlation study) and ID of the rater (e.g., participants of the rating study) were added as crossed random factors, and intercepts were allowed to vary randomly across both. Stimulus in the reverse correlation study (knife vs. pencil) was added as a fixed factor.

The results confirmed the previously reported t-test. CIs from the knife condition were judged on average as more formidable, M=3.40 than CIs from the pencil condition, M=2.86, $F(1, 41) = 41.6$, $p < .001$. The difference between the conditions is thus half a scale point, M = .54, 95% CI [.37, .70]. Inspection of the random factors showed that ratings shared variance due to both rating subject, Wald $Z = 4.17$, $p < .001$, and faceID beyond the condition, Wald Z = 3.5, $p < .001$.

2.4.4 Discussion

The findings suggest that being in possession of a knife makes a person look more formidable. The results from the binominal test showed that significantly more people selected the average CI from the knife condition as the most formidable. Two independent sample t-tests revealed that both the average gun CI and individual gun CIs were rated as significantly more formidable than the average and individual pencil CIs and confirmed the results from the bi-nominal test. Taken together, these results demonstrate that thinking or believing that a person is in possession of a knife makes people conceptualize the face of that person as more formidable than a person holding a pencil. More specifically, the findings suggests that the knife activates a summary representation of facial formidability that in turn enhances the facial features that signal formidability making his face look more formidable than it really is.

2.5 Experiment 2a

2.5.1 Participants

The initial sample comprised 60 participants recruited from Amazon Mechanical Turk. Based on exclusion criteria 14 participants were removed and excluded from further analysis. The final sample of 46 participants comprised 28 men and 18 women between the ages of 22 and 68 years (M = 36.54, SD = 11.87). The ethnicity was 84.8% White/Caucasian, 4.3%

African American, 8.7% Asian, and 2.2% other or mixed ethnicity. All participants were from the United States of America.

2.5.2 Design

Figure 8. Gun (A) and pencil (B) stimuli.

In a between subject design, participants performed a 2IFC task in which they were presented with two facial images placed on top of each other with a third image depicting a hand holding an object (gun versus pencil) placed to the left of two facial images (Figure 8A and 8B, respectively). Participants were randomly allocated to either the gun or the pencil condition. Instructions informed them that their task was to always select the facial image they believed was holding the object depicted next to facial images.

2.5.3 Procedure and materials

With the exception of a gun stimuli image replacing the knife stimuli image used in Experiment 1a and some new exclusion criteria, the materials and procedure mirrored Experiment 1a. A medium sized gun (Colt .45) replaced the knife as the lethal weapon. Because a larger gun has stronger recoil and require more strength to handle, I chose a medium sized gun in order to reduce the likelihood of participants' inferring that because the gun is big, the person holding it must also be big and strong. The same white Caucasian hand model and white backdrop was used to create the gun stimuli image. The pencil stimuli image, previously employed in study 1, remained the non-lethal object. In this experiment both stimuli were presented in black and white as opposed to in color as previously used in

Experiment 1a. Exclusion criteria now also included a question concerning the day of week and questions asking what type of objects they saw (some true and some false) and whether they believed they saw a male or female hand. The rationale behind these new exclusion criteria was to make sure that participants paid attention and did not answer frivolously as well as making sure that they believed that a man was holding the object.

2.5.4 Data processing

The data processing followed the procedure outlined in the general method section.

2.5.5 Results

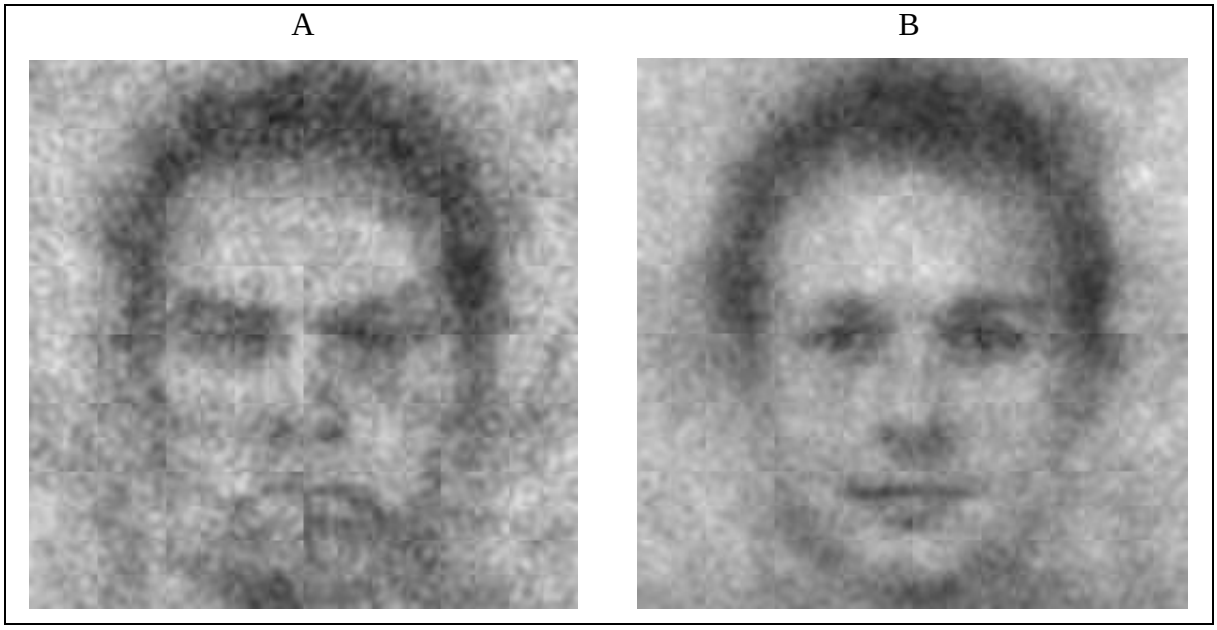


Figure 9. Total average CIs from gun condition (A) and pencil condition (B).

Figure 5 displays the CIs from the gun and pencil conditions. There is hardly any discernable visual difference between the total average CIs obtained in this experiment compared to the total average CIs from Experiment 1a. Visual inspection of the CIs show that the formidable CI from the gun condition (Figure 9A) displays a square and masculine facial shape with deep set, cold, cynical, and almost lifeless-looking eyes. His eyebrows appear thick and solid and are angled upwards. He has high and distinct cheekbones, sagging cheeks, downturned mouth with thick closed lips, wide and solid chin, wide nose-bridge bone, and a square shaped forehead with a slightly receding hairline. On the other hand, the CI from the pencil condition (Figure 9B) displays a more rounded face, warm and friendly looking eyes, less distinct and flatter eyebrows, rounded chin, closed mouth with soft and puffy-

looking lips, smaller and narrow nose, rounded jaw, and a more rounded forehead with a slightly receding hairline.

2.6 Experiment 2b – Rating of CIs

2.6.1 Participants and design

The initial sample comprised 73 participants recruited from Amazon Mechanical Turk. Based on exclusion criteria two participants were excluded from the dataset. The final sample included 71 participants with 37 males and 34 females between the ages of 19 and 67 years (M = 37.63, SD = 11.49). The ethnicity was 81.7% White/Caucasian, 8.5% African American, 1.4% Hispanic, 5.6% Asian, and 2.8% Native American. 98.6% were from the United States of America and 1.4% undisclosed. In a within subject design, all participants performed an image selection task and an image rating task.

2.6.2 Procedure and materials

The stimuli comprised the total average and individual CIs generated by the participants in Experiment 2a. Otherwise, the procedure and materials used in this experiment mirrored Experiment 1b.

2.6.3 Results

Prior to analysis participants' ratings of the average CI of the knife and pencil was averaged across each condition to obtain an average rating of each CI. A binominal test was conducted to test if the average CI from the gun condition was perceived as more formidable than the average CI from pencil condition. Inspection of frequencies show that of the total 71 participants, only 3 chose the average CI of the pencil condition as the more formidable person, while 68 (95.8%) chose the average CI of the gun condition. The binominal test confirmed that significantly more participants chose the average CI from the gun condition as the winner of the fistfight, Z = 7.6, $p < .001$.

A paired-samples t-test was conducted to compare the ratings of formidability of the average CIs from the gun and pencil conditions. There was a significant difference in the formidability ratings of the average CI from gun CI (M = 3.52, SD = .59) and pencil

(*M* = 2.69, *SD* = .64) conditions; *t* (70) = 12.86, *p* < .001. Another paired-samples t-test was conducted to compare the ratings of formidability of the individual CIs from the gun and pencil conditions. There was a significant difference in the formidability ratings of individual CIs from gun (*M* = 4.63, *SD* = .80) and pencil (*M* = 2.04, *SD* = .96) conditions; *t* (69)=15.03, *p* < .00.

I also repeated the multilevel analysis conducted in Study 2. Using MIXED in SPSS, I set up a multilevel model of using all 3266 (46*71) judgments from the rating study. Judged formidability was the dependent variable. ID of the CI (e.g., participants in the reverse correlation study) and ID of the rater (e.g., participants of the rating study) were added as crossed random factors, and intercepts were allowed to vary randomly across both. Stimulus in the reverse correlation study (gun vs. pencil) was added as a fixed factor.

The results again confirmed the previously reported t-test. CIs from the gun condition were judged on average as more formidable, *M* = 3.50 than CIs from the pencil condition, *M* = 2.69, *F*(1, 44) = 32.4, *p* <.001. The difference between the conditions are somewhat larger than in Experiment 1b, *M* = .81 scale points, but the confidence interval largely overlaps with Experiment 1b and includes the Experiment 1b mean value, 95%CI [.52, 1.09]. Inspection of the random factors showed that ratings shared variance due to both rating subject, Wald *Z* = 5.65, *p* <.001, and faceID beyond the condition, Wald *Z* = 4.5, *p* <.001.

2.6.4 Discussion

The findings suggest that being in possession a gun makes a person look more formidable. The results from the binominal test showed that significantly more people selected the average CI from the gun condition as the most formidable. These results were confirmed by two independent sample t-tests revealing that both the average gun CI and individual gun CIs were rated as significantly more formidable than the average and individual pencil CIs. Taken together, these results demonstrate that thinking or believing that a person is in possession of a gun makes people conceptualize the face of that person as more formidable than a person holding a pencil. More specifically, the findings suggests that the gun activates a summary representation of facial formidability that in turn enhances the facial features that signal formidability making his face look more formidable than it really is.

3 General Discussion

Previous research has shown that height, size and strength constitute a summary representation of formidability and that weapons make the person holding it appear bigger and stronger (Fessler et al., 2012). Furthermore, research has shown that people can accurately assess fighting ability (formidability) from cues in the face (Sell, Cosmides, et al., 2009). The present paper aimed to investigate the existence of a summary representation of facial formidability. More specifically, the aim, of the current thesis was to discover if lethal weapons make the face look more formidable.

To put this hypothesis to test, a two images forced choice (2IFC) reverse correlation technique was utilized to demonstrate that conceptualized fighting ability and lethal weapons enhance facial features of formidability. In a series of experiments, it was hypothesized that conceptualizing a person as a fighter will make his face appear more formidable (H1) and believing that a person is holding a knife (H2) or a gun (H3) will make his face appear more formidable. The results lend support for all three hypotheses. In the pre-studies, the results show that conceptualizing someone as a fighter caused people to perceive his face as more formidable than the loser of a fight thereby lending support to the first hypothesis. The results from Study 1 and Study 2 demonstrated that showing people a lethal weapon (gun or knife) also made people perceive the face of the person holding the weapon as more formidable, thus lending support to the second and third hypothesis.

In the following sections, I will first discuss the results in relation to each hypothesis and consider the current findings in relation to fighting ability and the summary representation of formidability proposed by Fessler. Because each hypothesis pertains to their own respective studies, I have chosen to follow the natural timeline of the studies. As such, I will firstly address H1 and the results obtained in the pre-studies with relation to fighting ability. I will then proceed to address H2 and H3 and the results obtained in Study 1 (Experiments 1a and 1b) and Study 2 (Experiments 2a and 2b) with relation to fighting ability and the summary representation of formidability. I will then proceed to outline some implications for future research and society before I lastly address the strength and limitations of the study.

3.1 Facial formidability conceptualized as fighting ability

The main objective in the pre-studies was to explore and discover what the internal representation of facial formidability looks like in an attempt to give formidability a face. One way to think about formidability is in terms of fighting ability (Sell et al., 2012; Sell, Cosmides, et al., 2009). Height, size, and strength are all components that play a crucial role in fighting ability and as such, this definition of formidability encapsulates all three components of Fessler's summary representation of formidability.

Unlike Studies 1 and 2, the Pre-studies did not include independent rating studies testing whether the total average and individual CIs from the winner condition is each study are indeed perceived as more formidable than the CIs from the loser condition. Interpretation of the results is therefore limited to my own visual interpretation of the CIs[3]. Visual inspections of the CIs from the winner and loser conditions clearly show that the winner looks considerably more formidable than the loser of the fistfight does. The observed facial features in the winner CI are consistent with previous findings linking said facial features to fighting ability. The formidable CI from the winner condition has what appears to be a high facial width to height ratio and deep-set eyes, consistent with previous findings linking these features with fighters appearance and higher combat success (Trebicky et al., 2013). Several other facial features apparently displayed in the formidable winner CI is consistent with findings by Toscano, Schubert, and Sell (2014), showing that judgments of dominance was predicted by strength, and that brow height, eye and chin length, and width of nose and mouth constituted the common predictors that underlies perception of strength and dominance. Lastly, the formidable CI from the winner condition appear to display prominent cheekbones, wide nose, low brow ridge and large chins, which corresponds with the sexually dimorphic features with male bias previously identified by research (Lefevre et al., 2012; Penton-Voak et al., 2001; Trebicky et al., 2013). The fact that the winner CI displays facial features previously associated with fighting ability lend support to Sell's findings that the face contains facial cues that signal fighting ability (Sell, Cosmides, et al., 2009). The striking consistency of shared similarities in facial features observed in all the individual CIs suggests that the formidable features contained in the face are universal and appear to form a summary representation of facial formidability.

[3] Unfortunately, the pixel cluster test previously employed by Dotsch & Todorov (2012) is not available for use in the current version of the R software package.

Although it is clear that the winner looks significantly more formidable than the loser of the fistfight, deducing the mechanisms that underlie the evaluations made by the participants remain unclear. Several possibilities present themselves. One possibility is that participants are making inferences based on observed and lived experiences. In this regard, conceptualizing someone as fighter is likely to activate a fighter stereotype schema or mental representation, allowing participants to consult a visual representation of what a fighter's face (and body) might look like. Taking into account, the rising popularity and consequent media coverage of, for example the Ultimate Fighting Championship (UFC), the stereotypical portrayal of male protagonists in the movie industry and so forth, it is highly plausible that the participants have a clear image of what a fighter looks like.

Another possibility is that the participants focused solely on the two images at hand and did confer with any other mental image or stereotype. This would entail that they are basing their evaluations and deCIsions solely on the observed facial features that signal formidability with no reference to any stereotype activation. In turn, this mental representation may or may not include bodily features other than the face. That would entail that they are making inferences based on neurocomputational specializations of the brain designed to detect facial cues that signal formidability. Participants' were instructed to select the winner or loser of a fistfight from two blurry facial images in a 2IFC task. This entails that the participants did not have any means of consulting any direct visual cues pertaining bodily features such as height, size, or strength other than those found in the face. Sells findings have shown that the facial cues associated with fighting ability most likely operate in such a manner that they signal strength and that displays of aggression can be used to signal and enhance cues of strength (Sell et al., 2014).

It is therefore also possible that, while the participants did not actually see any bodily features of the facial image, they may in fact have been making indirect assessments of bodily features such as strength (musculature) and possibly also height, and size. However, the last two components are perhaps less likely given that Sell, Cosmides, et al., (2009) found upper body strength to trump both height and size is evaluations of fighting ability based on facial cues. Therefore, while the participants may not be making inferences about the visual bodily features, they may be making inferences or assessments about strength. The possibility that the participants not only consulted the two facial stimuli for facial cues to formidability but also that they conceptualized what a fighter looks like in terms of physicality cannot be ruled

out. This means they effectively may have had two source of information to draw their inferences from and that our participants consulted not only facial features of formidability but also physical features such as size (height) and strength (muscularity). Either way, there is no way to discern whether the facial summary representation of formidability is separate system or if it is a part of Fessler's summary representation of formidability. Although somewhat exploratory in nature, the findings obtained in these two pre-studies lend some preliminary evidence in support of a facial summary representation of formidability.

As it stands, all of these options are equally likely and plausible and do not cancel each other out. Unfortunately, neither of pre-studies included any measures aimed at teasing apart such questions nor did it ask participants to describe the person they were thinking about when they made their selection. Taking into consideration that the pre-studies were just that, and not full-scale studies, the shortcomings of the studies should not be attributed to weakness in the design but rather the exploratory nature to the approach. It does however highlight several avenues for future research.

3.2 Weapons make the face look more formidable

In Study 1 and Study 2, the aim was to extend upon Fessler's finding that weapons make the man more formidable (Fessler et al., 2012), by showing that weapons also make the face look more formidable. If the facial cues that signal formidability also constitute a summary representation of facial formidability, I would expect that a weapon (knife in Study 1 and gun in Study 2) would also make the face look more formidable. Study 1 and Study 2 successfully demonstrated that believing that someone is in possession of a lethal weapon makes the face of that person look more formidable. In each study, participants completed a 2IFC task where they were shown two intentionally noisy facial images and an object (Knife in Study 1 and a gun in Study 2, or a pencil in both studies). The superimposed noise on the facial images works in such a manner that each of the two images always looks slightly different from each other.

Visual inspection of the CIs from each respective study shows that the CIs from the lethal weapon conditions (knife in Study 1, gun in Study 2) appear considerably more formidable than the CIs from the pencil conditions. The results from the separate rating experiments for each of the two studies confirmed these observations. The results from the

rating experiments show that the independent raters rated the formidable CIs as significantly more formidable than the CIs from the pencil conditions. Although hardly visibly discernible when comparing the CIs from the knife and gun conditions, the results obtained in the rating studies also reveal that the formidable CI from the gun condition was perceived as slightly more formidable than the formidable CI from the knife condition. The reason for this is most likely a result of guns generally considered more lethal than knifes (Fessler et al., 2012, Study 4).

The resemblances between the formidable CIs from Study 1, Study 2 and the pre-study are quite striking. Visual inspection reveal that they all appear to display relatively high fWHR compared to the loser and pencil CIs[4]. This finding is consistent with research showing that emotionally neutral faces became more threatening when fWHR increased (Carré, Morrissey, Mondloch, & McCormick, 2010, Experiment 3). While the base image was neutral, participants in the lethal weapon conditions selected the facial images in the 2IFC task that corresponded with the increased threat posed by being in possession of a lethal weapon.

Visual inspection of the individual formidable CIs reveals that there is slightly less consistency among them. There may be several reasons why we are witnessing this. It may be that the neurocomputational specializations are stronger with some than others are and what we are observing are various degrees of perceived formidability. Another possibility is that some were less vigilant or diligent than others were when making their decisions. Without any specific measure(s) to address this observation, I cannot conclude anything with certainty. This observation does not by any means entail that fWHR is the sole reason why the formidable winner CIs look more formidable than the CIs from the loser and pencil conditions. However, the fact that fWHR is a static facial quality does provide an important clue that also highlights one of the more noteworthy limitations of the studies. There are no means of discerning whether participants based their judgments on static or dynamic facial features when they selected the facial images.

What these results are telling us is that formidability has a face and that weapons make the face look more formidable. There are some fundamental differences in the design between the pre-studies and two consecutive studies. By not pitting the two blurry facial images

[4] fWHR cannot be accurately measured in the CIs obtained in these studies due to the noise.

against each other and framing it as a hypothetical fistfight as in the pre-studies, I was hoping to restrain the participants from consulting any stereotypical portrayal of a fighter in the main studies. Instead of asking participants who is more formidable, the participants in the main studies were required to select the facial image they believed was holding the object (knife/gun vs. pencil). By framing the question in such a manner, I can with greater certainty, assume that the mental representation of formidability captured by the CIs is a true reflection and product of the lethal weapons activating the neurocomputational specialization for detecting formidability in the face. However, there is still a possibility that participants consulted mental images derived from various stereotypes they may associate with guns or knives (e.g., movie stars).

The observed consistency between all the formidable CIs seem to suggest that the neurocomputational specialization for detecting formidability works in such a way that it encapsulates the array of facial features that signal formidability into a single mental representation, a "universal" stereotypical formidable face, or in other words, a summary representation of formidability. This stereotypical formidable face, displaying all the various features that signal formidability, then serves as a point of reference for comparison when assessing relative formidability.

How do the current findings conform to Fessler and Sell's findings? The current findings speak towards the existence of a facial summary representation of formidability much like Fessler's summary representation of physical formidability. Unlike Fessler who has shown that height, size, and strength constitute key components of physical formidability, the current research have yet to discover which specific facial features constitute the key components in facial formidability. Previous research has shown that facial images can be used to obtain physical measure of formidability (Holbrook & Fessler, 2013; Sell, Cosmides, et al., 2009). While Sell et al demonstrated that people can accurately assess upper body strength from facial cues, Fessler and Holbrook successfully employed a facial image depicting a convicted terrorist holding a gun to obtain measures of physical formidability (height, size, and strength).

Considering these findings and taking into account the current findings that weapons make the face look more formidable, it appears that the two representational systems are either part of the same system, or at least closely interconnected. Although I cannot say for certain, it is plausible that, if the facial summary representation of facial formidability is

activated and consequently projected unto a persons' face, it may in turn activate Fessler's summary representation of physical formidability causing the body to appear taller, bigger, and stronger than it really is. Sell and colleagues (Sell et al., 2014) finding that anger and aggression can be used to enhance cues of strength lend further support to this notion.

One way to put this notion to test would be to extend upon the current design employed in the pre-studies by incorporating an independent rating experiment and add some form of measure(s) of height, size, and strength in the 2IFC task. Such a design would increase the validity of the current findings by confirming that the formidable CIs are indeed more formidable that the control CIs and provide a means of assessing whether participants are in fact making bodily inferences. One possibility would be to include a 2IFC variant in which participants are shown silhouette images of bodily features that vary in height, size, and muscularity and have them select the facial image they believe belongs to the body. Such a design would provide a means of discovering whether formidable physical qualities are associated with formidable facial features. If the existence of a facial summary representation of formidability holds true then the CIs from the participants that saw formidable body features should look considerably more formidable than the CIs from the participants who saw less formidable body features. Should the formidable CIs mirror the facial features of those already obtained in the current study then that would lend support to the notion that those features constitute a facial summary representation of formidability.

We can be fairly certain the facial features that signal formidability also form a summary representation of facial formidability. What we cannot tell for certain is how it works. What are the underlying mechanisms that drive evaluations of facial formidability? Does it function in such a manner that it primes participants to scan the facial images for formidable features such as signs of anger, aggression, dominance, fWHR, or does it present itself as a mental image that participants can use to compare for similarities? At this stage of the research, it does not matter. The final product remains the same since the formidable CIs clearly shows what the mental representation of facial formidability looks like, and it appears that the mental image captured in the CIs display many, if not all of the facial features associated with formidability.

The important question is, when we are making evaluations of formidability, are we superimposing a mental image (a summary representation of facial formidability) unto people's faces, like putting on a mask, or are we enhancing facial features already present in

the face? To put it in a different manner, can everybody appear more formidable than they actually are, irrespective of whether they possess the static facial features associated with facial formidability? This last question is particularly interesting because an answer to this question may provide an answer to many still unanswered questions. There are two types of facial features, static and dynamic. Static features are based on facial structure and underlying bone. FWHR is one example of a static facial feature. On the other hand, dynamic features are features based on facial musculature and can thus be used to mimic static facial features. Sell for example has demonstrated that displays of anger and aggression (dynamic features) can be used to enhance cues of strength. This is possible because displays of anger and aggression mimic static features that signal formidability (Sell et al., 2014).

Findings by Todorov et al., (2008) show trait inferences are made along the orthogonal dimensions of dominance/ trustworthiness and power/valence. Their research demonstrates that an evaluation of valence and dominance on emotionally neutral faces is an overgeneralization of adaptive mechanisms for inferring/surmising behavioral intentions and power hierarchies. In emotionally neutral faces, subtle resemblance to facial expressions that signal approach (happiness) or avoidance (anger) functions as the basis for evaluations of valence. Cues that signal physical strength such as masculinity and facial maturity serves as the source of dominance evaluation. In terms of functionality, cues that signal approach or avoidance and strength fuels inferences about intentions to inflict harm and the ability to cause harm (Fiske, Cuddy, & Glick, 2007). In other words, whereas expressions (e.g., anger) signal intentions (e.g., to cause harm), stable/static cues signal ability (e.g., to cause harm).

Several approaches can be taken to explore if people are making evaluations of formidability based on static or dynamic facial features. A logical next step would be to tease apart the relationship between fWHR and formidability. One possible approach could be to take a desired number of computer generated facial images, obtain a measure of fWHR and use a computer software (e.g., Photoshop) to wipe out the internal features of the face (e.g., eyes, eyebrows, nose, and mouth), leaving only the outer ridges, and have participants rate them in terms of perceived formidability. Another possibility could be to use the CIs from the studies in this thesis and smooth out the noise to allow for measurements of fWHR[5] and then wipe out the internal facial features. This method would allow for testing the relationship

[5] The superimposed noise makes it impossible determine the exact location of various facial features needed to obtain a correct and precise measurement of fWHR.

between fWHR and facial formidability – if fWHR, independent of the other facial features contained in the face, is related to formidability, one would expect the facial images with higher fWHR to be rated as more formidable than those with lower fWHR.

Another approach would be to obtain or create an image of person with an average body and neutral face holding a lethal and non-lethal object. The next step would be to crop the head from the body, leaving a headless body image, and use computer software to morph the facial images to vary across the dimensions of valence (anger/happiness). Anger and happiness are expressions that are similar to dominance and trustworthiness, respectively. In addition, the full body image would be manipulated to vary in terms of height, size, and muscularity in order to create arrays for each feature. In a series of trials, participants would see one of the original untampered set of images (lethal object vs. non-lethal object), with the head still attached. The next step would be to show participants an array of the morphed facial images (including the neutral original base face) as well as the arrays of the headless body images and ask them to select the face and bodies they believe were holding the lethal and non-lethal object. If dynamic cues in the face were used to make evaluations of formidability, it would be expected that the facial images that were morphed along the dimension of anger and happiness would be selected as the face belonging to the person holding the lethal object and non-lethal object, respectively. Based on previous findings by Fessler et al. (2012) it would also be expected that the person holding the lethal object to be perceived as taller, bigger and more muscular than the person holding the non-lethal object.

3.3 Strengths and limitations

Most of the strengths and limitations pertain to the design of the studies. The RC technique employed in this thesis proved to be a suitable and strong method. The 2IFC tasks and consequent rating studies provided a suitable platform that allowed me to successfully probe and capture the internal representation of facial formidability, thus giving formidability a face. The RC technique has proven strong in a wide array of previous research and as such it provides solid internal validity to the findings obtained in the studies reported in this thesis. Using different types of populations recruited from two different countries increases cross-cultural validity that in turn increases generalizability of the findings. The sample pool comprised Norwegian university students and Americans recruited from Amazon Mechanical Turk (AMT). Some might question the effort and diligence of participants recruited from

services such as AMT. This concern is unwarranted as research has shown that AMT samples are more diverse than university sample pools and perform just as good, if not better than participants recruited from university sample pools (Buhrmester, Kwang, & Gosling, 2011).

The biggest limitations also concern the design. Although the facial features that cue formidability is there to observe, we do not actually have any means of discerning exactly what they are. Previous RC studies (when relevant) have made use of a pixel and cluster test (Dotsch & Todorov, 2012) to extrapolate the specific regions in the face stimuli images used in judgments. This method however, is not yet readily available and possible to perform in current version of the R software employed in these studies. However, even this method is not without its limitations. Even if we were able to use this method to identify the specific regions in the face that inform judgments of formidability, we would still not be able to tell whether these regions are static or dynamic cues.

This line of research is still in its infancy so naturally there are some limitations. Considering the novelty of this research, it was important that I kept it simple and focused on the key findings of which the research is based. Although it clearly would have been advantageous to identify the specific facial features that informed our participants' judgments, this was not a part of my research question. Moreover, as clearly stated, the visual interpretation of the CIs is subjective and speculative. They are by no means integral to the argumentation in support of our findings. They are only discussed in terms of their possible implication to the findings. As the research further develops, there will be opportunities to address and circumvent the current limitations addressed.

3.4 Implications for society

The current finding that impressions and visual representation of faces vary as a function of perceived formidability may have important implications in the judicial system, in law enforcement, and in healthcare systems, to name a few. In regards to the judicial system and law enforcement, the findings may raise some serious implications for witness testimonies and suspect line-ups. The current findings suggest that if a person for some reason is perceived as formidable this may activate the summary representation of formidable facial features causing a construal of the face whereby formidable facial features is superimposed on the face. If the circumstances surrounding the event a witness is asked to recall entail some

level of perceived formidability in the perpetrator(s), it is fully possible that the face(s) he or she remember is quite different from what it actually is. Consequently, this may lead to faulty prosecution and criminal charges.

The current findings may be used to better understand how perceived formidability may serve to amplify or escalate antagonistic situations. The statement made by Darren Wilson, the police officer that shot and murdered Michael Brown in the Ferguson-shooting read: “Wilson then described Brown becoming enraged, and that Brown looked like a demon” (United States Department of Justice., 2015, p. 14). Had the police officer not perceived M. Brown as a huge demon-like person he might have been less inclined to fire his gun. This scenario underscores the importance of taking into account how our perception of people might change as result of circumstances and training law enforcement officers and people alike to remain calm and clear headed in order to employ strategies aimed at defusing situations before they escalate into potentially threatening situations.

Much in the same manner, the findings may also be relevant for childcare institutions and mental health wards. It may be used to better understand how perceived formidability may serve to amplify or escalate antagonistic situations, or how it may amplify the impressions of psychotic patients with already heightened senses. It contexts like these, body language is an important factor. The current findings suggests that for example body postures that signal displays of dominance or aggression could, in agonistic situations, cause patients to perceive the face as more threatening than it actually is. This in turn may cause them to act out. The above-mentioned implications provide contexts that underscore just how devastating these findings may be in some cases.

3.5 Implications for research

Several implications for future research can be drawn from the findings contained in this thesis. Beyond the avenues for future research previously mentioned when addressing the results, future research might take some other interesting directions. A potential avenue for future research would be to investigate if the current findings can be replicated using a female base face image. Neither Sell nor Fessler have explored if their findings generalize to female targets. Sell reasons that because males are considerably stronger than females and more often deploy physically aggressive strategies, the cognitive specializations for strength assessment

are expected to be better engineered for evaluating males than females (Sell et al., 2012). That is not to say that the cognitive specializations for strength assessment are non-existent in females but that they may be weaker.

Zilioli et al., (2014) findings that fWHR predicts formidability is interesting to consider in this context. Although fWHR has been linked to testosterone, mature features, and dominance – all features commonly associated with the male gender, the finding that fWHR is sexually dimorphic has recently come under scrutiny. A growing body of research suggests that fWHR is not sexually dimorphic (Lefevre et al., 2012). It may be weaker, but not non-existent. Exploring this notion by replicating the current study using a female base face image would therefore be useful to evaluate which facial features in women signal formidability and what the female mental summary representation of facial formidability might look like.

The current findings show that formidable facial features are enhanced or superimposed on the face. Comparing the neutral base image to the formidable CIs demonstrates just how extreme this effect can be. In light of how this effect can construe impressions and memories of faces it would be beneficial for future research to attempt to work backwards in the sense of trying to obtain some sort of measurement to identify to what degree and extent these features are enhanced.

4 Conclusion

The studies contained in this thesis attempted to give formidability a face. This was done by both looking at what kind of face people assume for a winner of physical fight and people who has gun. The results found that they look close to identical. In order to do so, two pre-studies and two main studies employed a two-image forced choice task to assess internal representations of facial formidability. In both main studies, participants completed a two-image force choice task in which they saw an image of a hand holding a lethal weapon or a non-lethal object and two noisy facial images as part of reverse correlation design. Their task was always to select the facial image they believed to be the person holding the object. The lethal object in the first study was a kitchen knife while the lethal object in the second study was medium sized gun. The non-lethal object in both studies was pencil. For each main study, a pool of independent raters confirmed the results by rating the lethal weapon classification images resulting from the reverse correlation paradigm as significantly more formidable than the non-lethal classification images. Results from the pre-studies revealed that conceptualizing someone as a fighter makes him look more formidable while results from the two main studies showed that lethal weapons make the face of the person holding it look more formidable. These findings illustrate that conceptualizing someone as a fighter and thinking that a person is in possession of a lethal weapon activates a mental summary representation of facial formidability which in turn is superimposed on the neutral facial image, making that person look more formidable than he actually is. The current findings lend support to the notion that human beings possess neurocomputational specializations for identifying facial features that signal formidability. The finding that perception of faces change as function of perceived level of formidability may have important implications for how people remember and recall faces and how those people are ultimately perceived and treated. Extending on the current research by identifying whether the underlying mechanisms behind facial evaluation of formidability rely on static or dynamic facial cues is necessary in order to determine who is subject to the overgeneralization effect of facial formidability.

References

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–45. http://doi.org/10.1146/annurev.psych.59.103006.093639

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). How to make a risk seem riskier : The ratio bias versus construal level theory. *Perspectives on Psychological Science*, *6*(1), 3–5. http://doi.org/10.1177/1745691610393980

Carré, J. M., Morrissey, M. D., Mondloch, C. J., & McCormick, C. M. (2010). Estimating aggression from emotionally neutral faces: Which facial cues are diagnostic? *Perception*, *39*, 356–377. http://doi.org/10.1068/p6543

Dotsch, R. O. N., Wigboldus, D. H. J., & Knippenberg, A. D. V. A. N. (2013). Behavioral information biases the expected facial appearance of members of novel groups. *European Journal of Social Psychology*, *125*(December 2012), 116–125.

Dotsch, R., & Todorov, a. (2011). Reverse Correlating Social Face Perception. *Social Psychological and Personality Science*, *3*(5), 562–571. http://doi.org/10.1177/1948550611430272

Dotsch, R., & Todorov, A. (2012, December 12). Reverse Correlating Social Face Perception. *Social Psychological and Personality Science*, *3*(5), 562–571. http://doi.org/10.1177/1948550611430272

Dotsch, R., Wigboldus, D. H. J., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science*, *19*(10), 978–80. http://doi.org/10.1111/j.1467-9280.2008.02186.x

Fessler, D. M. T., & Holbrook, C. (2013a). Bound to lose: physical incapacitation increases the conceptualized size of an antagonist in men. *PloS One*, *8*(8), e71306. http://doi.org/10.1371/journal.pone.0071306

Fessler, D. M. T., & Holbrook, C. (2013b). Friends shrink foes: the presence of comrades decreases the envisioned physical formidability of an opponent. *Psychological Science*, *24*(5), 797–802. http://doi.org/10.1177/0956797612461508

Fessler, D. M. T., & Holbrook, C. (2014). Marching into battle : synchronized walking diminishes the conceptualized formidability of an antagonist in men. Biol. Lett. 10: 20140592. http://dx.doi.org/10.1098/rsbl.2014.0592

Fessler, D. M. T., Holbrook, C., & Snyder, J. K. (2012). Weapons make the man (larger): formidability is represented as size and strength in humans. *PloS One*, *7*(4), e32751. http://doi.org/10.1371/journal.pone.0032751

Fessler, D. M. T., Tiokhin, L. B., Holbrook, C., Gervais, M. M., & Snyder, J. K. (2014). Foundations of the Crazy Bastard Hypothesis: Nonviolent physical risk-taking enhances

conceptualized formidability. *Evolution and Human Behavior*, *35*(1), 26–33. http://doi.org/10.1016/j.evolhumbehav.2013.09.003

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. http://doi.org/10.1016/j.tics.2006.11.005

Gosselin, F. G., & Schyns, P. G. (2003). Superstitious perceptions reveal representations of spatial propositions, *14*(5), 505–509. Retrieved from http://eprints.gla.ac.uk/30033/

Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision Research*, *45*, 1707–1724. http://doi.org/10.1016/j.visres.2004.12.021

Holbrook, C., & Fessler, D. M. T. (2013). Sizing up the threat: the envisioned physical formidability of terrorists tracks their leaders' failures and successes. *Cognition*, *127*(1), 46–56. http://doi.org/10.1016/j.cognition.2012.12.002

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology*, *141*(1), 19–25. http://doi.org/10.1037/a0023463

Karremans, J. C., Dotsch, R., & Corneille, O. (2011). Romantic relationship status biases memory of faces of attractive opposite-sex others: Evidence from a reverse-correlation paradigm. *Cognition*, *121*(3), 422–426. http://doi.org/10.1016/j.cognition.2011.07.008

Lefevre, C. E., Lewis, G. J., Bates, T. C., Dzhelyova, M., Coetzee, V., Deary, I. J., & Perrett, D. I. (2012). No evidence for sexual dimorphism of facial width-to-height ratio in four large adult samples. *Evolution and Human Behavior*, *33*(6), 623–627. http://doi.org/10.1016/j.evolhumbehav.2012.03.002

Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, *27*(4), 237–254. http://doi.org/10.1023/A:1027332800296

Penton-Voak, I. S., Jones, B. C., Little, a C., Baker, S., Tiddeman, B., Burt, D. M., & Perrett, D. I. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proceedings. Biological Sciences / The Royal Society*, *268*(1476), 1617–1623. http://doi.org/10.1098/rspb.2001.1703

Pinker, S. (2011). *The Better Angels of our Nature: The Decline of Violence in History and its Causes*. London: Penguin Books Ltd.

Schnall, S., Harber, K. D., Stefanucci, J. K., & Proffitt, D. R. (2008). Social support and the perception of geographical slant. *Journal of Experimental Social Psychology*, *44*(5), 1246–1255. http://doi.org/10.1016/j.jesp.2008.04.011

Sell, A., Bryant, G. a, Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., … Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice.

Proceedings. Biological Sciences / The Royal Society, *277*(June), 3509–3518. http://doi.org/10.1098/rspb.2010.0769

Sell, A., Cosmides, L., & Tooby, J. (2014). The human anger face evolved to enhance cues of strength. *Evolution and Human Behavior*, *35*(5), 425–429. http://doi.org/10.1016/j.evolhumbehav.2014.05.008

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings. Biological Sciences / The Royal Society*, *276*(October 2008), 575–584. http://doi.org/10.1098/rspb.2008.1177

Sell, A., Hone, L. S. E., & Pound, N. (2012). The Importance of Physical Strength to Human Males. *Human Nature*, *23*(1), 30–44. http://doi.org/10.1007/s12110-012-9131-2

Sell, A., Tooby, J., & Cosmides, L. (2009). Formidability and the logic of human anger. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(35), 15073–8. http://doi.org/10.1073/pnas.0904312106

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2014). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology*, (August 2014), 1–27. http://doi.org/10.1146/annurev-psych-113011-143831

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–60. http://doi.org/10.1016/j.tics.2008.10.001

Toscano, H. Schubert, S. Sell, A. (2014). Judgments of Dominance from the Face Track Physical Strength. *Evolutionary Psychology*, *12*(1), 1–18.

Trebicky, V., Havlícek, J., Roberts, S. C., Little, A. C., & Kleisner, K. (2013). Perceived aggressiveness predicts fighting performance in mixed-martial-arts fighters. *Psychological Science*, *24*, 1664–72. http://doi.org/10.1177/0956797613477117

United States Department of Justice. (2015). Department of Justice Report Regarding the Criminal Investigation Into the Shooting Death of Michael Brown by Ferguson, Missouri Police Officer Darren Wilson, 86. Retrieved from http://www.justice.gov/sites/default/files/opa/press-releases/attachments/2015/03/04/doj_report_on_shooting_of_michael_brown_1.pdf

Von Rueden, C., Gurven, M., & Kaplan, H. (2008). The multiple dimensions of male social status in an Amazonian society. *Evolution and Human Behavior*, *29*(6), 402–415. http://doi.org/10.1016/j.evolhumbehav.2008.05.001

Windhager, S., Schaefer, K., & Fink, B. (2011). Geometric morphometrics of male facial shape in relation to physical strength and perceived attractiveness, dominance, and masculinity. *American Journal of Human Biology*, *23*(6), 805–814. http://doi.org/10.1002/ajhb.21219

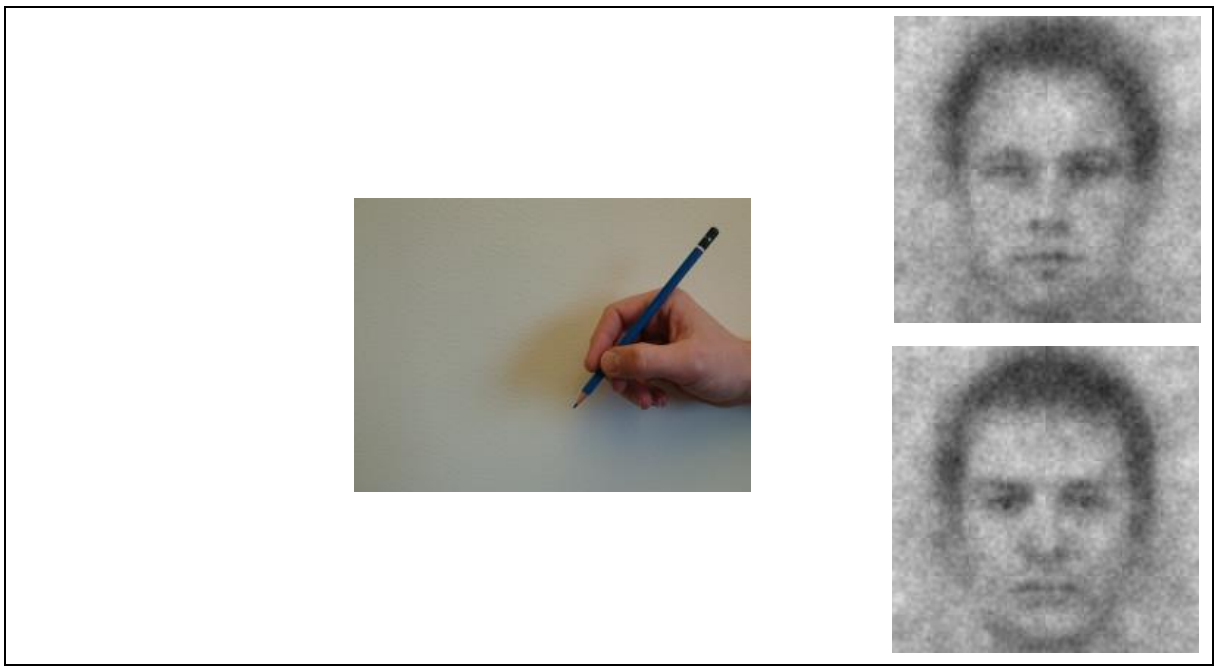
Zilioli, S., Sell, A. N., Stirrat, M., Jagore, J., Vickerman, W., & Watson, N. V. (2014). Face of a fighter: Bizygomatic width as a cue of formidability. *Aggressive Behavior*, *9999*(April), 1–9. http://doi.org/10.1002/ab.21544

Attachments

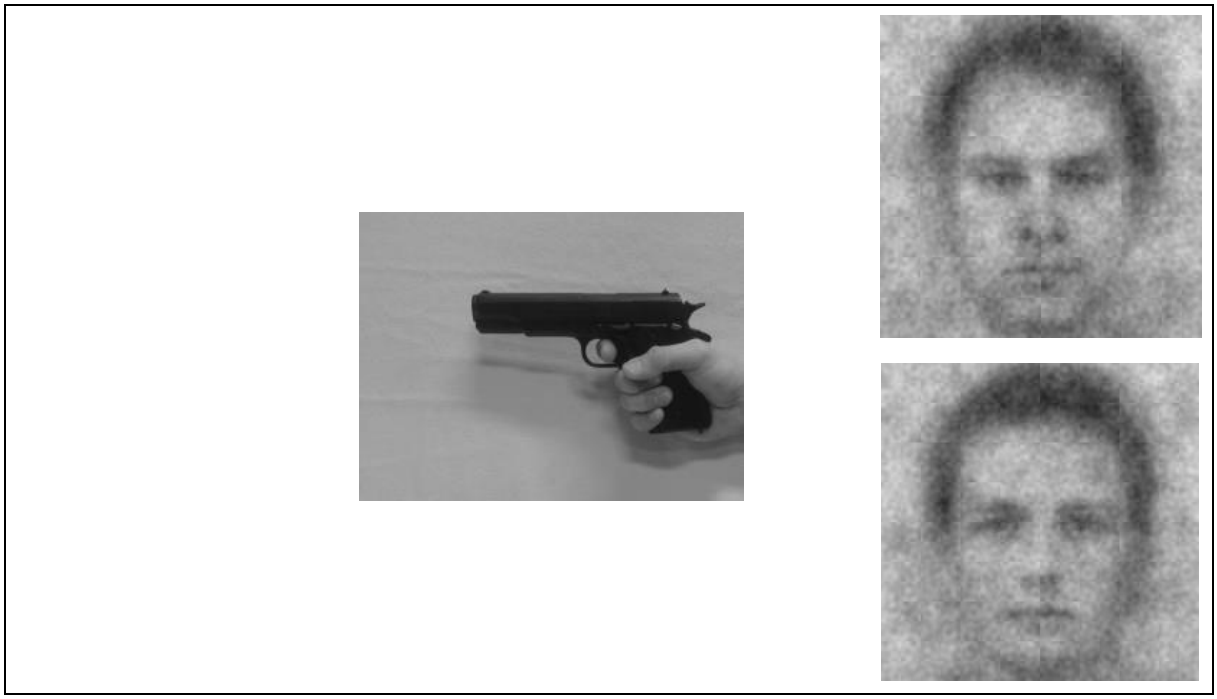
Example two-image forced choice trial taken from the survey used in the pre-studies.

Example two-image forced choice trial taken from the lethal weapon condition in Experiment 1a.

Example two-image forced choice trial taken from the non-lethal weapon condition in Experiment 1a.



Example two-image forced choice trial taken from the lethal weapon condition in Experiment 2a.



Example two-image forced choice trial taken from the non-lethal weapon condition in Experiments 1a and 2a.