LUND UNIVERSITY

The power of slurs in reported speech:

A survey-based study of native speaker intuitions

Jennifer Melckersson

Supervisor: Professor Carita Paradis

Centre for Language and Literature, Lund University
MA in Language and Linguistics, English
SPVR01 Language and Linguistics: Degree Project – Master's (Two Years) Thesis, 30 credits

October 2021

Abstract

When slurs are used in indirect reports (e.g., ‘Lisa said that Tom is a [slur]’), there is a disagreement between theories of language about which speaker is responsible for the derogatory content and offensive potential of the slur—the original speaker or the reporting speaker. This debate is specific to slurs and does not typically extend to non-slur pejoratives. To examine this issue, this thesis compares native English speaker intuitions about the offensiveness of slurs, neutral counterparts, and non-slur pejoratives in isolation, indirect reports, and predicative utterances. These intuitions were elicited in two internet-based surveys. Questions asked by the study concern the offensiveness of slurs in predicative utterances versus indirect reports, the differences and/or similarities of slurs and non-slur pejoratives in these same contexts, and the compatibility of selected slur theories with the results of the surveys.

Results show that while offensive, slurs in indirect reports are less offensive than slurs in predicative utterances. Results also show that non-slur pejoratives are less offensive than slurs in all conditions, and that the difference between reported non-slur pejoratives and predicative non-slur pejoratives is substantially smaller than the difference between reported and predicative slurs.

Therefore, the results contradict theories which attribute offensiveness to the original speaker, and are somewhat compatible with other theories. In the future, research that is better equipped to test for differences in context can further examine the predictive value of such theories. Moreover, the study suggests that most prominent slur theories face issues in explaining the varying degrees of offensiveness found in the judgments of reported slurs. I posit that this issue requires more extensive examination and analysis to fully explore.

Keywords: *slurs, pejoratives, indirect reports, reported speech, offensiveness, native speaker intuition*

Acknowledgements

Firstly, thank you to my supervisor Carita Paradis for her encouragement, advice, and faith in me. I am so grateful for her patience and accommodation every time I fell behind on the intended schedule.

Secondly, I would like to acknowledge my professors and teachers at Lund these past few years, for providing me with an exceptional education and ultimately, the tools to create this thesis. A special thanks goes out to Satu Manninen, for inspiring me and encouraging me many times over since I started taking English at Lund University back in 2016. I also have to mention my peers Michelle and Defne from the Master's program, who have inspired me more than they probably know. Additionally, I would like to thank Matteo Fuoli and Robbie Love for their willingness to share my survey with their students.

Thank you also to my friends and family for helping distribute the surveys in this study, especially Johanna, for her assistance in distributing the pilot survey, and my aunt Sabine, for her invaluable efforts to help distribute one of the surveys to friends and family in the United Kingdom.

Finally, the biggest thank you to my mom Cecilia, for her eternal support.

Table of contents

List of figures

List of tables

Chapter 1 Introduction

Commonly, slurs are referred to as "pejorative terms that target people on the basis of nationality, ethnicity, religion, gender, etc." (Cepollaro, Sulpizio, & Bianchi, 2019, p. 32). For example, such terms include *nigger,* a slur targeting black people on the basis of the color of their skin, and *faggot,* a slur targeting gay people on the basis of their sexual orientation. This can be contrasted with non-slur pejoratives like *idiot, asshole,* or *fucker,* which instead target people based on individual traits or actions. Slurs have been of particular interest to linguists and philosophers of language for the past 20 years, as scholars have attempted to understand how they fit into existing theories of language. Accounts diverge on how the derogatory content of slurs is conveyed, and why they have such—in some ways unique—power to offend. Most would likely agree that slurs, when used directly as in (1), communicate something derogatory and offensive.

(1) John is a faggot.

This derogation is intuitionally evident. However, there are syntactic contexts of slur usage that provoke disagreements among many linguists and philosophers. According to some, the derogatory content of slurs is not neutralized when slurs are placed in, for example, negated, conditional, or reported contexts (Anderson & Lepore, 2013a, 2013b). See (2) and (3), which are examples of such contexts, and consider what these statements can be said to communicate. It is possible to argue that their derogatory content and offensive nature is equal to that of (1).

(2) John is not a faggot
(3) Brian said that John is a faggot

Some other scholars hold that this derogatory content can be neutralized in certain syntactic or contextual circumstances, but do not necessarily agree on which circumstances, or why it works that way (see Hom, 2008; Potts, 2007; Schlenker, 2007).

In this thesis, I home in on the issue of slurs in reported speech, specifically indirect reports such as (3). The use of slurs in reported speech is a particularly interesting phenomenon because expectations pertaining to their derogatory force and offensiveness tend to differ vastly

between the main slur theories. For example, Hom's (2008, 2012) semantic account of slurs, *Combinatorial Externalism* (sometimes referred to as *Literalism*), dictates that the derogatory content of slurs is always attributed to the original slur user, not to the reporter of said slurring utterance. Meanwhile, Anderson and Lepore (2013a, 2013b), progenitors of the *Prohibitionism* approach, dictate that the derogatory content of a slur is always attributed to the reporter, and not necessarily the original slur user. It is easy to see how these opposing theoretical positions could lead to empirical questions. Which of the resulting predictions would be borne out in empirical research? Are these intuitions reflected in the evaluations of regular native speakers of English? I believe that being able to answer these questions, even superficially or preliminarily, would be a significant contribution to our understanding of slurs as a linguistic phenomenon. Therefore, I have chosen to conduct a survey-based study of the offensiveness of slurs in indirect reports compared to that of slurs in non-reported, predicative statements.

At present, there are relatively few studies that examine this issue empirically. Cepollaro et al. (2019), Panzeri and Carrus (2016), and Tenchini and Frigerio (2020) are three such studies that use surveys to compare slur theories with empirical data. In these investigations, the main slur approaches are tested by comparing the offensiveness of slurs with that of non-slur pejoratives and other terms in contrasting contexts, one of which is reported speech compared with non-reported speech. To my knowledge, these three are the only empirical studies which include investigations of slurs in reported speech. Furthermore, all three of these studies were conducted using intuitions of native speakers of Italian as data. This means that no such study has yet been conducted using native speakers of English.

1.1 Aim & research questions

The primary aim of this thesis is to explore the merits of the selected slur theories—as they relate to slurs in reported speech—through an empirical study of offensiveness intuitions in native English speakers. These offensiveness intuitions are elicited through two surveys, whose results are compared to the expectations expressed by the relevant theories. These theories include Hom's (2008, 2012) Combinatorial Externalism, Potts (2007) and Whiting's (2013) *Conventional Implicature theory*, Schlenker (2007) and Cepollaro's (2015) *Presupposition-based theory*, Anderson and Lepore's (2013a, 2013b) Prohibitionism, and Nunberg's (2018) *Affiliationism,* all of which are described in section 2.4. In this endeavor, I am guided by the main research questions of the thesis.

- RQ1: Does offensiveness differ between slurs in non-reported speech and indirect reports? If so, how?
- RQ2: What do the results of the surveys tell us about the differences and/or similarities of slurs and non-slur pejoratives?
- RQ3: Which of the selected slur theories, if any, are compatible with the results of the surveys?

All research questions, but particularly RQ3, address the primary aim of the thesis, which is to examine the selected slur theories. Methodologically speaking, fulfilling this aim is facilitated not only by comparing intuitions about reported slurs with intuitions about non-reported slurs, but also by comparing intuitions about slurs with intuitions about other categories of words (e.g., general taboo words, non-slur pejoratives, neutral descriptors, etc.). As motivated by the sparsity of empirical slur research, this part of the study constitutes its secondary aim—to find out more about slurs by comparing them with other kinds of words in both reported speech, non-reported speech, and in isolation.

1.2 Ethical considerations regarding slurs in academia

In this paper, I discuss various slurs and statements containing slurs, something which in itself runs the risk of offending, promoting harmful attitudes, and exploiting minorities and marginalized people. This is not my intention, but I acknowledge that my best intentions may not be enough to prevent it from happening. With this in mind, it is important to explain how I treat slurs in this paper and why I believe it to be a subject worthy of academic attention.

Firstly, slurs are of particular societal importance. They are associated with numerous hateful attitudes and ideologies, and many heinous acts both historical and contemporary. Their use is sometimes examined in the field of hate speech, and there are researchers who argue that certain slurs, particularly the [n-slur], should be classified as hate speech in a legal manner (Holt, 2018). In a discussion of the American legal system, Holt (2018) says: “As history indicates, hate speech is often the underlying principle that guides hateful conduct. Few words are more likely to incite violence than the word nigger” (p. 419). As such, the [n-slur] is more than an insulting word, partly due to the history of oppression from which it originates. While the [n-slur] may be the most archetypal slur, and the most historically established, this severity can be considered to point to the significance of slurs in general, and thus the need for examining the linguistic mechanisms that underpin slurs as a category.

If it is the case that slurs constitute a legitimate and important subject for research, it calls for sensitive treatment and a premeditated ethical framework. Camp (2018), Anderson and Lepore (2013a), and Cousens (2020), among others, emphasize this need. Anderson and Lepore (2013a), for example, highlight the need to be mindful of the risks of making judgments on slur usage in academic work. They urge scholars to "ask yourself how a targeted member, […], would react to your usage" (p. 31). Additionally, Camp (2018) writes that "even reading [slurs] rightly makes many people deeply uncomfortable" (p. 32). These statements further support my resolve to limit slur occurrences to a certain extent in this thesis. But even if one admits that spelling out a slur is or can be offensive or harmful in academia, there are arguments in favor of doing so. In his paper on disability-related slurs, Cousens (2020) makes two arguments for using the full slurs, albeit sparingly. These are, for one, that "hiding the affective power of slurs behind euphemisms such as 'slur term S' can make it too easy to treat these terms as interesting linguistic puzzles and gloss over their impact", and for another, that "reproducing the use of slurs in canonical examples makes it easier to engage with other accounts directly" (p. 2). Spelling out full slurs can thus be beneficial, despite the ethical risks.

Moreover, there is an idea that contributing something to the field of pejorative language is a sort of compensation for its didactic or academic usage. Camp (2018), for example, who writes out full slurs in her work, says that when tokening slurs in an academic setting, she "incur[s] an obligation to compensate with commensurate insight" (p. 32). I concur, but I also choose not to use the full slurs when it does not directly contribute to context, relevance, or pedagogy. For me, this means that when presenting illustrative examples that require specificity, when quoting from my questionnaire or from its responses, or when quoting relevant literature, I write the full slur. In all other cases, to the best of my ability, I use variations of [slur] to represent members of the category slur. However, what constitutes a slur can be a matter of opinion and theory. Prior to any more advanced discussion, this must be addressed.

Chapter 2 Theoretical background

In terms of a theoretical foundation, this study combines a number of fields, including subfields of traditional linguistics and philosophy of language. This chapter is a description of the current landscape of slur research and literature, which cannot be limited to a single discipline. For example, in his description of existing slur theory, Croom (2011) separates his discussion into two superordinate categories, one of which encompasses work done in a linguistic framework and one work done in a philosophic framework. Furthermore, theories in both disciplines are central to the purposes of this thesis.

In this chapter, relevant definitions of the term slur are presented in section 2.1, followed by descriptions of relevant previous literature in sections 2.2 and 2.3, as well as established slur theories in section 2.4. Finally, the preceding empirical studies on the subject are summarized in section 2.5.

2.1 What is a slur?

The term slur is broadly conceptualized in everyday use, as slur can mean any "shaming or degrading effect" (Merriam-Webster, n.d.). In linguistic or philosophic contexts, slurs are often defined as derogatory expressions that target people or groups on the basis of group membership.

Slurs are also typically classified as a type of pejorative (Hom, 2012). Pejoratives are terms that by virtue of their meaning insult or denigrate. As described by Diaz-Legaspe (2020), many scholars make a distinction between group pejoratives and individual pejoratives and claim that group pejoratives are the ones we commonly define as slurs. Group pejoratives are defined by their targeting of groups based on demographic membership as opposed to behavioral choices. From this perspective, all slurs are group pejoratives (Diaz-Legaspe, 2020).

However, it is possible to make further requirements for inclusion in the slur category. For example, one can make requirements pertaining to the moral value of the target group, or an unbalanced power dynamic between the user(s) of the slur and its target group. Diaz-Legaspe (2020) argues that only group pejoratives for morally neutral groups should count as slurs,

which would exclude pejoratives like *pedo*, which does not refer to a morally neutral target group. She also makes a restriction based on the concept of *dominance*:

> Slurs are not just pejorative terms targeting identifiable neutral groups. They occur in the context of a particular type of social tie between the community of speakers using the slur (call them Ss) and the targeted group (call them Gs). In all cases, the relation between Ss and Gs involves marginalization, subordination, oppression and/or discrimination.
>
> (Diaz-Legaspe, 2020, p. 1413)

This relation between slur users and those targeted is described as a dominance relation (Diaz-Legaspe, 2020). As Diaz-Legaspe (2020) states, this relation is complex and varies depending on many factors, but the term dominance is used to loosely encompass these complexities without going into too much social theory.

This dominance property is also highlighted in sociological material on slurs. For example, Embrick and Henricks (2013) set out from this vantage point in their investigation of the racial inequality of slurs and stereotypes. They write that because of unequal power relations, it is vastly different to use a slur that targets black people by virtue of being black compared to using a slur that targets white people by virtue of being white. More explicitly, Embrick and Henricks (2013) contend that for groups which are based on ethnicity or race, "neutral power relations cannot be presumed and different groups' epithets and stereotypes cannot be judged by the same standard" (p. 198). Moreover, the dominant group, generally white people in this context, are not socially affected by having slurs used against them in the same way that non-white people are (Embrick & Henricks, 2013). This is partially because slurs are "instruments of power" that "legitimize hierarchal arrangements" (Embrick & Henricks, 2013, p. 199). Thus, based on Embrick and Henricks (2013) and their predecessors, slurs are not as effectually wielded by those who have less power against those who have more. With this in mind, we see that although slurs targeting a dominant group can be used in similar pejorative contexts as slurs targeting non-dominant groups, the effects of them are different; Slurs targeting non-dominant groups are generally more effective.

Another potential defining property of slurs is the existence of a corresponding *neutral counterpart* to each slur. A neutral counterpart to a slur contains the same descriptive content

as the slur, but does not include a pejorative, derogatory, or offensive component. An example of this slur-neutral counterpart dichotomy is *kike* versus *Jew*. Some scholars hold that the existence of a neutral counterpart is required in order for a word to be a member of the category slur (e.g., Hom, 2012; Nunberg, 2018; Vallée, 2014; Whiting, 2013). However, it is not universally accepted that all slurs require a neutral counterpart (see Ashwell, 2016). Words such as *slut* or *bitch*, for example, do not necessarily have neutral counterparts, but they can still be considered slurs (Ashwell, 2016). The difference between these gendered pejoratives and, for example, racial pejoratives, can be considered that the former target people based on behavior (such as sexual promiscuity) whereas the latter target demographic groups based on membership in those groups (Ashwell, 2016). This would make the former non-slur pejoratives and the latter slurs. Ashwell's (2016) counterpoint to this is that it is not membership in a group that provokes slur usage but rather a thinking that people are "lesser than others simply because of something that *ought not to be thought of* as marking you out as lesser than others" (p. 236). From this perspective, pejoratives that are based on behavior and do not have commonly understood neutral counterparts would not have to be excluded from the slur category.

Ultimately, these differences of opinion about what constitutes a slur are important to be aware of, but when researching slurs specifically, I believe it has many benefits to focus one's investigation on the expressions that the least amount of people would disagree are slurs. Therefore, the definition of slur in this study is delimited based on a number of central properties, which I have derived from the reviewed literature. These properties are the following:

- An expression that can be categorized as a group pejorative
- An expression that targets morally neutral groups
- An expression that is used against a non-dominant group
- A derogatory expression that has a non-derogatory, neutral counterpart

This is not to say that only expressions that have these traits can be considered slurs. However, for the present study, using slurs that not all relevant theories would agree are slurs would invalidate a comparison of the results with those theories. Using only slurs that have these four central properties ensures that what is being tested is also part of what is being referred to by the literature.

2.2 Indirect reports

The term indirect report refers to the act of reporting what another speaker has said by using some reporting verb or mechanism instead of direct quotation (e.g., 'Jessica said that Martin is good at cooking'). Generally, it is considered that "the aim of reporting is to say what an original speaker said" (Capone, 2012, p. 601).

Theories of indirect reporting are interested in, among other things, the semantic and pragmatic relationship between what is said by the reporter and what was said, or may have been said, by the original speaker. By default, indirect reports are interpreted as conveying the same meaning as the utterance that is being reported (Capone, 2012; Davidson, 1968). Although different words, syntactic structures, or languages may be used, the reporter commits themselves to accurately representing the original speaker's meaning as interpreted (Capone, 2016; Davidson, 1968). However, this default can be contradicted by contextual factors, such as our knowledge of or relationship with the reporter and/or the original speaker, expectations of relevance, circumstances of the speech situation, and more (Capone, 2012, 2016).

2.3 Impoliteness & offensiveness

Another field in which slurs and related terms have been investigated is that of (im)politeness. In the study of (im)politeness, taboo language is of interest because examining it inevitably contributes to our understanding of politeness and impoliteness norms.

A common discussion within the (im)politeness field is whether, or to what degree, impoliteness is conventionalized (Culpeper, 2011). Proponents of the discursive approach to impoliteness emphasize the role of context in determining the impoliteness of an utterance, in contrast with traditional conceptions of politeness and impoliteness as formulaic and "speaker-centric" (Culpeper, 2011, p. 122). However, while acknowledging the importance of contextual variables, Culpeper (2011) cautions against the most absolute versions of the discursive approach. As "people have opinions about how different expressions relate to different degrees of politeness or impoliteness *out of context*", the meaning and function of these expressions have to be "semantically encoded in some way" (Culpeper, 2011, p. 124). Therefore, there are expressions which are conventionalized as impolite to some degree.

The idea of conventionalized impoliteness is easily connected to slurs. It has been claimed that slurs, especially the [n-slur], are some of the most offensive words in the English language (Holt, 2018). Culpeper (2011) hypothesizes that more offensive expressions are both "more context-spanning and more conventionalised" (p. 153). It follows that slurs, assigned on average a high level of offensiveness, are both conventionalized and context-spanning in their impoliteness.

Importantly, how offensive an expression is depends on how offensiveness is defined in the first place, and the concept of offensiveness is intrinsically linked to impoliteness. Impoliteness causes the addressee to take offense, and this phenomenon is vital for determining what impoliteness is and how it functions. In this study, what is meant by *offensive* broadly corresponds with its dictionary definition of "causing displeasure or resentment" (Merriam-Webster, n.d.). However, there are more specific ways of defining and discussing this concept which may be relevant as well. Culpeper (2011) talks about seven different "offence types", which were adapted from several works by Spencer-Oatey (p. 43). Three of the seven stem from the concept of *face*, which was developed as part of Politeness theory by Brown and Levinson (1987), and previously, Goffman (1967, as cited in Culpeper, 2011). Face, in this context, refers to "an image of self delineated in terms of approved social attributes" (Goffman, 1967, p. 5, as cited in Culpeper, 2011, p. 24). Offense can thus be incurred when this image of self is threatened by one's interlocutors. Two of the offense types come from *sociality rights*, which pertain to our expectations of how people around us should act in different situations based on legal or social conventions (Culpeper, 2011). This type of offense deals with instances when "something counters a state of affairs which a participant considers to be considerate and fair" (Culpeper, 2011, p. 39). Culpeper (2011) also adds two of his own categories, in *taboo*, a third subcategory of sociality rights, and *physical self*, which includes speech situations where a participant is made to feel that their physical self is in danger. From Culpeper's (2011) thorough treatment of this concept, it is clear that offensiveness is more complex than its dictionary definition suggests. Offense can stem from various different, yet related, sources, such as derogation, unfairness, the breaking of taboos, and more. In the surveys used in this study, participants were asked to rate words or utterances on a scale from 'not at all offensive' to 'extremely offensive', and no definition of offensive was provided. This left it open for participants to interpret offensiveness in any way they liked. These interpretations may be encompassed by one or several of the offense types laid out by Culpeper (2011).

Another aspect of impoliteness which I believe to be especially relevant to slurs is that of *metadiscourse*. Culpeper (2011) states that "metadiscourse is part of the social dynamic that makes a behaviour count as impolite in the first place" (p. 153). Because the present study is based on surveys that include open-ended questions about slurs and offensiveness, this metadiscourse is part of what it will examine, in the form of individual metalinguistic comments. The field of impoliteness can thus contribute several ideas to a study on slurs such as the present one.

2.4 Slur theories

Often, the linguistic and philosophic literature on slurs provides answers to one or both of the following questions (Tenchini & Frigerio, 2016):

1. How is the derogatory content of slurs conveyed (i.e., what makes them differ in meaning compared to their neutral counterparts)?
2. Why do slurs tend to retain their derogatory content or offensive potential when embedded in for example negated, conditional, and reported contexts?

The following sections outline this literature in the form five slur-based theories. This is done with a focus on their view of slurs in reported contexts specifically. It is important to note that these do not encompass all or even a majority of slur theories, but merely a selection of theories which are deemed prominent due to their explicit stance on reported slurs and/or their recurring relevance in the field.

There are several ways of classifying these theories, but a simple way of doing so is to divide them into three categories (Tenchini & Frigerio, 2016). The first category comprises of semantics based theories like Hom's Combinatorial Externalism, which argues that the derogatory component of slurs is part of their semantic content. The second encompasses more pragmatic-leaning theories, including those based on conventional implicatures, presuppositions, and conversational implicatures, which argue the reverse. The final category is based on social prohibitions against slur usage, and does not take semantic content as relevant to slur theory at all.

2.4.1 Combinatorial Externalism

Hom (2008) analyzes slurs from a semantic point of view, according to which "the derogatory content of an epithet is fundamentally part of its literal meaning" (p. 416). In semantic contexts,

the original description of the literal meaning of a slur such as *faggot* is 'homosexual and despicable because of it'. Hom (2008) argues that this is an insufficient and "naïve formulation", as some slurs are considered particularly derogatory compared to others, and this description does nothing to explain this difference (p. 416). Hom's (2008) approach solves this problem by appealing to semantic externalism. In semantic externalism, "the meanings for words […] are at least in part dependent on the external, social practices of the speaker's linguistic community" (p. 430). Hom (2008) extends this theory to mean that "the derogatory content of an epithet is semantically determined by an external source" (p. 430). One such external source is the "social institutions of racism" that exist in most modern societies (Hom, 2008, p. 430). It is because of the social institution of racism towards (primarily) Chinese people that the slur *chink* has derogatory force, and it is the power of that institution that supports and determines the strength of that derogatory force (Hom, 2008). In the western world, at least, there is virtually no social institution of racism supporting the pejorative content of slurs targeting white people, like *limey* or *honkey*. This, according to Hom's (2008) view, accounts for the difference in derogatory force between *limey* and aforementioned slurs targeting Chinese and black people. By elaborating on this idea, Hom (2008) arrives at the approach he calls Combinatorial Externalism. It is described as follows:

> *Combinatorial externalism* (CE) is the view that racial epithets express complex, socially constructed, negative properties determined in virtue of standing in the appropriate external, causal connection with racist institutions.
>
> (Hom, 2008, p. 431)

In this approach, the derogatory meaning of a slur is also part of its semantic meaning, which separates it entirely from its neutral counterpart. Moreover, this derogatory meaning is dependent on the power of various social and societal forces. This description may intimate a conception of slurs as being derogatory in every utterance and every context, but this is not Hom's (2008) intention, as he later points out that slurs "*express* derogatory semantic content in every context, but they do not actually derogate their targets in every context" (p. 432). Derogation is, according to Hom (2008), "the actual application, or predication, of derogatory content" (p. 432). This view leads to—among other things—a presumed existence of what Hom (2008) calls *non-derogatory, non-appropriated contexts*, or NDNA. It is commonly accepted that slurs can be non-derogatory and non-offensive in appropriated—also called reclaimed—contexts, when used by those that the slurs typically target (Anderson and Lepore, 2013a).

NDNA contexts, however, refer to those uses of slurs that are not reclaimed, but also not derogatory. As such, Hom (2008) posits that although slurs contain derogatory content, it is how they are used that makes them derogatory. Using a slur when reporting or quoting someone else's utterance, then, is not applying derogatory content to someone but merely mentioning a word that happens to have derogatory content.

Later, Hom (2012) clarifies this view on reported slurs further, arguing that slurs are *displaceable*. Displaceability as a term refers to the ability of for example slurs to be embedded within negated, conditional, and reported constructions, as well as other complex constructions. As such, Hom (2012) claims that slurs are displaceable, and therefore their derogatory content is displaced, or neutralized, by attitude reports. In other words, the derogation "successfully remain[s] in the scope of attitude reports" (p. 388). This means that according to CE, the derogatory content of slurs is attributed to the original speaker, not the reporting speaker.

Still, Hom (2012) also acknowledges that even in attitude reports, negated utterances, conditionals, etc., slurs can provoke offense. With this in mind, he argues for a distinction to be made between derogation and offense. For Hom (2012), when slurs are used in reports and negated utterances, they are objectively not derogatory, but they can be subjectively offensive. The offensiveness of a reported slur is explained through conversational implicature as opposed to semantics. However, this argument can be questioned (Jeshion, 2013; Tenchini & Frigerio, 2016). For example, to explain the derogatory and/or offensive component of slurs differently depending on whether the slurs appear in non-reported or reported utterances has been criticized as counter-intuitive, because "they seem equally offensive and for the same reason" (Jeshion, 2013, p. 317). With that in mind, this separation of offense from derogation is arguably not as straightforward as posited by Hom (2012) in his effort to advance his account.

2.4.2 Conventional Implicature

Potts (2007) and Whiting (2013) find the derogatory component of slurs to be conveyed through conventional implicature, which is a notion developed by Grice (1989). From their perspective, slurs and their neutral counterparts contain the same descriptive content and the same truth conditions. For example, *faggot* and *homosexual person* are considered to have the

same extension or referents (Whiting, 2013).[1] The only difference between the slur and its neutral counterpart is that the slur is derogatory whereas the neutral counterpart is not (Whiting, 2013). Since, in this view, the derogatory meaning of a slur is conventionalized but not part of its semantic content (or truth conditions), Potts (2007) and Whiting (2013) argue that its derogatory meaning is a conventional implicature.

The criteria for something to be a conventional implicature are that part of its meaning is detachable, not cancellable, and not something that is calculated on the basis of "conversational exchange and the circumstances of utterance" (Whiting, 2013, p. 365). Whiting (2013) argues that all of these criteria are met by slurs. Their derogatory meaning is "not typically" cancellable, as one cannot state that someone is a [slur] and follow up with 'but I have nothing against [neutral counterpart]' (Whiting, 2013, p. 365). A slur's derogatory meaning is detachable because one can use its neutral counterpart instead of the slur to say the same thing without derogation (Whiting, 2013, p. 365). Additionally, the derogatory meaning is not and cannot be calculated because the derogatory meaning is either something that is already known by the interlocutors or something that is unknown and impossible to infer through the words alone (Whiting, 2013).

Notably, Whiting (2013) speaks less of slurs as conveying derogatory content and more of slurs as expressive of negative *attitudes*. Whiting (2013) and Potts (2007) view epithets, including slurs, as *expressives*. In addition to incorporating slurs into Potts' (2007) view of conventional implicature, this shapes his approach to slurs in reported speech in several different ways. For example, Potts (2007) describes expressives as defined by six properties, all of which are presented in truncated form below:

1. *Independence*: Expressive content contributes a dimension of meaning that is separate from the regular descriptive content.
2. *Nondisplaceability:* Expressives predicate something of the utterance situation.
3. *Perspective dependence:* Expressive content is evaluated from a particular perspective.
4. *Descriptive ineffability:* Speakers are never fully satisfied when they paraphrase expressive content using descriptive, i.e., nonexpressive, terms.

[1] It is my understanding that slurs can have several possible neutral counterparts. The [f-slur], for example, could conceivably take any of the following as a neutral counterpart: *homosexual person, gay person, homosexual male, gay male,* or just *gay.*

5. *Immediacy:* Like performatives, expressives achieve their intended act simply by being uttered; they do not offer content so much as inflict it.
6. *Repeatability:* If a speaker repeatedly uses an expressive item, the effect is generally one of strengthening the emotive content, rather than one of redundancy.

(Potts, 2007, pp. 166-167)

This shows that Potts (2007) believes that slurs are *nondisplaceable*, which contrasts with what Hom (2012) has later argued. The ascription of the nondisplaceability property to slurs means that, in this view, the derogatory content of slurs cannot be embedded in constructions such as reports. Rather, the slur remains derogatory regardless of the syntactic construction. This may lead one to believe that Potts' expects attitudes implicated by slur usage to always be attributed to the reporting speaker. However, the *perspective dependence* property in turn tells us that although this may be the default, it is not always the case. To explain this, Potts (2007) adapts a model from Lasersohn (2005) which is based on the concept of a *contextual judge*. Is an utterance evaluated based on the immediate context of that utterance or some other context (e.g., a context that is being reported)? Potts (2007) says that "[a]s a pragmatic default, the judge is the speaker. But Lasersohn discusses many cases in which predicates of personal taste have another salient entity as their judge" (p. 175). Lasersohn does not extend "predicates of personal taste" to expressives, but Potts (2007) does (p. 175).

In this view, the judge is the agent whose emotions, attitudes, or states of mind are being expressed. In reported utterances, the judge can be either the reporting speaker or the original speaker, depending on whose emotions or attitudes are being expressed. If the original speaker is "salient enough" to be the contextual judge, the attitudes are ascribed to the original speaker (Potts, 2007, p. 175). If not, we default to ascribing those attitudes to the reporting speaker. Therefore, within the Conventional Implicature account, the evaluation made by addressees concerns which of these agents is the most salient. Presumably, this judgment is at least partially subjective, and comes about based on a number of contextual factors.

2.4.3 Presupposition-based theory

Like the Conventional Implicature account, the Presupposition-based account considers slurs to be expressives, and their descriptive content to be identical to that of their neutral counterparts (Cepollaro, 2015; Schlenker, 2007). Where these accounts differ is in their

analysis of slurs' derogatory component. Schlenker (2007) and Cepollaro (2015) advocate for a Presupposition-based theory of slurs. Pragmatic presuppositions are "something taken for granted: […], a requirement on the common ground", and slurs are, according to this view, presupposition triggers (Cepollaro, 2015, p. 39). Just like other presuppositions, "slurs trigger a certain proposition even when they are embedded" (Cepollaro, 2015, p. 39).

Depending on one's view of presuppositions, there are two kinds of presuppositional analyses that can be made (Cepollaro, 2015). Either "the presupposition [triggered by the slur] describes a state of affairs about the target group" (the objective option), or "the presupposition [triggered by the slur] expresses how the speaker feels about the target group" (the subjective option) (Cepollaro, 2015, p. 43). To illustrate using Cepollaro's (2015) examples, the objective option is where (5) is the presupposition of (4), and the subjective option is where (6) is the presupposition of (4) (pp. 42-43).

(4) I don't do business with wops.
(5) Italians are despicable because of being Italian.
(6) The 'wop'-user despises Italians for being Italian.

As Cepollaro (2015) states, these options can be related to the general divide between analyses of presuppositions as either objective, truth relational functions, or subjective, interactional functions. Cepollaro (2015) and Schlenker (2007) both favor the subjective option.

If the presupposition is related to the participants in the speech situation as in (6), reports of slur usage could conceivably be attributed to either the original speaker, the original '[slur]-user', or the reporter, the current '[slur]-user'. Indeed, on the topic of reporting slurs, Schlenker (2007) describes slurs as *shiftable indexicals*. What this means is explained in the following quote.

> Standard indexicals are expressions that must be evaluated with respect to the context of the actual speech act. Shiftable indexicals are more promiscuous, and may be evaluated with respect to any context (e.g. the context of a reported speech act).
>
> (Schlenker, 2007, p. 239)

Essentially, a shiftable indexical can be evaluated based on either context, including the original context of slurring as well as the context of the report. Crucial is that this is an either-or dichotomy depending on the addressee's interpretation (Cepollaro et al., 2019).

2.4.4 Prohibitionism

The Prohibitionism theory argues that slurs are prohibited words. Uttering them in any context breaks a taboo, thus incurring offense. Anderson and Lepore (2013a) state that "slurs are *prohibited* words *not* on account of any content they get across, but rather because of relevant edicts surrounding their prohibition" (p. 26). Prohibitionism also dictates that this is how all profanities and taboo words function—slurs are simply "higher on the prohibition scale" (Anderson & Lepore, 2013b, p. 355).

One of the main benefits of this theory according to Anderson and Lepore (2013b) is that it can explain the difference in derogatory force between two different slurs that have the same target group (i.e., co-extensive slurs such as *nigger* and *coon*, of which the former is considered more offensive). Because the derogatory force of a slur depends on the degree of prohibition associated with that individual word, different degrees of prohibition can accompany co-extensive slurs (Anderson & Lepore, 2013b). For Combinatorial Externalism, in contrast, the derogatory force of a racial slur depends on the strength of the racist institution that supports it. Anderson and Lepore (2013b) argue that as the same institution governs the strength behind all slurs for any particular group, Combinatorial Externalism cannot account for how co-extensive slurs vary in force.

Moreover, Anderson and Lepore (2013b) claim that while other theories—theories based on the content of slurs—can find the behavior of slurs in reported speech to be difficult to explain, the explanation is obvious for Prohibitionism. Generally, it is best practice to re-use exact expressions when reporting someone else's utterance, as doing so affords the reporter the greatest amount of accuracy (Anderson & Lepore, 2013a)[2]. When reporting slur usage, however, this principle has a negative consequence. For example, Anderson and Lepore (2013a) claim that (8) as a report of (7) does not automatically make the original speaker, Eric, the offending party.

[2] However, a certain amount of paraphrasis and summation on the part of the reporter is expected (see Capone, 2012).

(7) A bitch ran for President of the United States in 2008.[3]

(8) Eric said that a bitch ran for President of the United States in 2008.

Rather, "it guarantees an offense by *whoever is reporting him*" (Anderson & Lepore, 2013a, p. 29). This because there is, by default, nothing to say that the reporter did not choose to use the slur independently of what descriptor the original speaker used. In Anderson and Lepore's (2013a) view, it is possible that (8) was uttered as a report of something like (9).

(9) A woman ran for President of the United States in 2008.

Since this possibility always exists in indirect reports, the reporting speaker is always the one to whom the greatest offense will be attributed. This absolute stance is controversial, and many, including Cepollaro (2015), Hom (2012), and Schlenker (2003), disagree with it. Anderson and Lepore's (2013a) defense against their counterarguments is that offense certainly can be attributed to the original speaker over the reporting speaker, but only in cases where certain pragmatic and contextual variables encourage it. In their subsequent work on the subject, Anderson and Lepore (2013b) hold that "should the complement clause of an indirect report contain a slur […], then whatever offensive content that slur carries will be attributable to the *reporter*" (p. 354). The key word here is, I believe, *attributable.* For Prohibitionism, it is always fair to attribute the slur and the offense it incurs to the reporter, regardless of context.

2.4.5 Affiliationism

Nunberg's (2018) theory is that the use of slurs aligns a speaker with a certain belief, ideology, or speech community.[4] Some slurs are so closely associated with certain beliefs that any use of them is tantamount to proclaiming one's support of those beliefs. This phenomenon is explained through a variant on the Gricean notion of conversational implicature, which differs from conventional implicature in that it is highly context-dependent. Nunberg (2018) suggests that slurs and our responses to them can be explained via a type of conversational implicature which he terms "ventriloquistic implicature" (p. 267). A ventriloquistic implicature is when,

[3] Anderson and Lepore (2013a) view *bitch* as a slur.
[4] This theory has been called *Affiliationism* by Jeshion (2016), a practice which I adopt here.

by using marked or nonstandard language, "the speaker evokes or impersonates" members of a different speech community (Nunberg, 2018, p. 267). The phenomenon is explained as such:

> In a particular context, a speaker pointedly disregards the lexical convention of the group whose norms prescribe the default way of referring to A and refers to A instead via the distinct convention of another group that is known to have distinct and heterodox attitudes about A, so as to signal his affiliation with the group and its point of view.
>
> (Nunberg, 2018, p. 268)

This is easily applied to the context of slur usage, because the *group whose norms prescribe the default way of referring to A* is, most of the time, going to be the group that is targeted by the slur, at least according to the logic of this theory. Their norms prescribe the default way of referring to them, i.e., by the neutral counterpart. The other *group that is known to have distinct and heterodox attitudes about A* is, in this context, the slur users, which are often racists or other people with negative attitudes towards certain groups. Essentially, using a slur encourages hearers to infer that the speaker harbors a certain attitude, and it does so by the active selection of the slur over the neutral counterpart. As such, Nunberg's (2018) theory has something in common with Prohibitionism in the sense that the offense lies in flouting norms more so than in the content of the word itself.

Nunberg (2018) has much to say about slurs in indirect reports. For example, ventriloquistic implicatures in general "tend to be speaker-oriented, even in embedded contexts" (p. 268). In other words, they tend to be nondisplaceable (Nunberg, 2018). This implies that slurs are attributed to the reporting speaker as opposed to the original speaker. As with several of the other theories, there are explicit exceptions. One such exception is contexts of reclamation, but Nunberg (2018) also highlights another exception. He describes the neutral counterpart to a slur as an "unmarked term" and the slur as a "marked term", because the former is permissible within the norms of the speech situation and the latter is not (Nunberg, 2018, p. 270). Nunberg (2018) states that "unless the implicated attitude is conversationally relevant", the reporter is not informationally obligated to re-use the marked term, i.e., the slur (p. 270). This means that choosing to use the marked term in a report is offensive or non-offensive depending on the informational relevance of that term in the original utterance. As Nunberg (2018) says, "if there's no reason to suppose that the speaker's choice of words was pertinent to the point at

issue, […] then your pointed decision to use [racial slur] will be taken as an indication of your own racial attitude" (p. 270). Nunberg (2018) does not describe how the addressee of the reporting speaker would gauge the conversational relevance of the slur—presumably such relevance would have to be evident through the reporting utterance and surrounding context alone.

2.5 Previous empirical research

As was described in the introduction chapter, three empirical studies of native Italian speakers have been conducted which include the specific subject of slurs in reported utterances. The first one, chronologically speaking, is by Panzeri and Carrus (2016). They examine the offensiveness of a collection of eight slurs in multiple syntactic contexts, such as negated utterances, conditional utterances, questions, and indirect reports. They also investigate the offensiveness of the same eight slurs in isolation, in order to compare the offensiveness of the slurs themselves to their offensiveness in the relevant syntactic constructions.

In their introduction, Panzeri and Carrus (2016) highlight how different theories account for the offensiveness and derogatory force of slurs in different ways. At the center of the discrepancy between theories is how slurs behave in linguistic operators like embedding, negation, reports, questions, conditionals, etc. (Panzeri & Carrus, 2016). More concretely, Panzeri's and Carrus' (2016) article is based on two different studies which each include a questionnaire. The first of those studies is the one that includes the issue of slurs in reported speech. The relevant questionnaire is divided into two sections. In the first section, participants were faced with 32 words in isolation and asked to rate their offensiveness on a 7-point scale. In the second section, participants were asked to rate, on the same scale, "the offensiveness of a person who utters a sentence that contain [sic] a word […] embedded under the linguistic contexts of negation (NEG), antecedent of conditional (ANT), question (QUE), and indirect report (IND)" (Panzeri & Carrus, 2016, p. 174). The words used in the questionnaire are eight slurs, eight neutral counterparts, eight positive controls, and eight non-slur pejoratives (Panzeri & Carrus, 2016, p. 174). The 132 responses collected show that in isolation, slurs and non-slur pejoratives are both offensive, and the difference between them in offensiveness is "not statistically different" (Panzeri & Carrus, 2016, p. 175). In context, the person who reports an utterance that includes a slur is perceived as offensive. However, they are perceived as overall less offensive than a person who uses a slur in the context of a question or in the antecedent of

a conditional (Panzeri & Carrus, 2016). Ultimately, Panzeri and Carrus' (2016) investigation of slurs in reports specifically is considered, in their words, "not clear" (p. 177).

The second study is by Cepollaro et al. (2019). It focuses entirely on the issue of the offensiveness of slurs in indirect reports compared to non-reported speech. Cepollaro et al. (2019) also motivate their research by outlining the different theoretical approaches to slurs and how several of them make different empirical predictions in regard to the offensiveness of slurs in indirect reports. Cepollaro et al. (2019) make the distinction between Prohibitionism approaches such as Anderson and Lepore (2013a, 2013b) and non-Prohobitionism approaches, such as Schlenker (2007). Their main observation of these approaches is that Prohibitionism predicts that offensiveness is always attributed to the reporting speaker, and the non-Prohibitionism theories predict that "the reporting speaker need not always be taken to be the bearer of the offensive attitude" (Cepollaro et al., 2019, p. 33). This is consistent with my review of the theoretical literature as well. Their study is divided into three surveys. Their pilot survey investigates offensiveness attitudes to words in isolation, including 30 slurs, 22 "non-slurring labels" (i.e., neutral counterparts), 20 non-slur pejoratives, and 19 fillers (Cepollaro et al., 2019, p. 34). All three of Cepollaro et al.'s (2019) surveys use a 7-point scale from 1 (not at all offensive) to 7 (highly offensive). Cepollaro et al. (2019) explain that words from the slur and non-slur pejorative categories were included in order to compare the offensiveness of words that have some descriptive value, like slurs, with words that are "only meant to offend", like most non-slur pejoratives (p. 37).

The second of Cepollaro et al.'s (2019) surveys elicits offensiveness intuitions of the relevant terms in the context of non-reported speech, using stimuli with the formulation "'X is a P', where X is a proper name […] and 'is a P' is a predicate" that can take the form of a slur, non-slurring pejorative, or neutral counterpart (p. 35). The third survey elicits offensiveness intuitions of the relevant words in both predicative formulations and indirect report formulations. That survey uses 104 experimental sentences, half of which are in the non-reported speech form and half in the indirect report form. These are further divided into the three categories previously described, with the addition of an equally numerous category of fillers. They make the following observations:

> "(i) slurs are perceived as more offensive than insults in isolation; (ii) the perceived offensiveness of slurs decreases when they are in atomic predications of the form

> ‘X is a P’ rather than in isolation; (iii) the perceived offensiveness of non-slurring insults increases when they are in atomic predications rather than in isolation; (iv) atomic predications featuring slurs or insults are perceived as less offensive when they are reported, even though the report cannot entirely delete their offensive power.”
>
> (Cepollaro et al., 2019, p. 39)

Cepollaro et al. (2019) suggest that their second and third observations, (ii) and (iii), are due to the fact that slurs have descriptive content and therefore contribute something factual to a statement, whereas non-slur pejoratives only contribute something offensive or insulting. Additionally, reporting slurs is perceived as less offensive than using them directly, but is still offensive. Cepollaro et al. (2019) argue that this result is most compatible with non-Prohibitionism views, especially the presuppositional one, but that some amendments to it may have to be made. In contrast, the result challenges Prohibitionism because that theory cannot explain the lesser degree of offensiveness observed in reported speech (Cepollaro et al., 2019).

The third study is by Tenchini and Frigerio (2020), and is also based on a questionnaire. They too claim that although using a slur is considered offensive and impolite, “it is unclear whether and to which degree reporting a slur can also be impolite” (p. 274). Like the previous studies on this subject, Tenchini and Frigerio (2020) contextualize their study by describing how different slur-related theories make different predictions in terms of which speaker the derogatory component of slurs is typically attributed to in a reporting scenario—the original speaker or the reporting speaker. Additionally, this study serves as an empirical examination of their own slur theory, the Multi-act theory, which is based on the idea that “a speaker performs two different speech acts in uttering a sentence containing a slur: a representative act, […] and an expressive act” (Tenchini & Frigerio, 2020, p. 277).

Tenchini and Frigerio’s (2020) questionnaire focuses on a singular slur, *frocio* (‘faggot’). It also includes the neutral counterpart *omosessuale* (‘homosexual’), the non-slur pejoratives *idiota* (‘idiot’) and *coglione* (‘asshole’), as well as the neutral descriptive term *studente* (‘student’). These words were presented to the participants in several different contexts. For example, they were presented in predicative constructions such as ‘you are *p*’, ‘X is *p*’, and ‘Y: X is *p*’, as well as indirect report constructions like ‘Z: Y has said that X is *p*’. The participants were asked to judge the offensiveness of these sentences on a 5-point scale which contained

five descriptive response options to the question of whether a sentence was offensive. From the responses to this questionnaire, of which there were 121, they find that "the offensiveness of slurs and nonslurring pejoratives embedded in reported speech decreases homogenously" (Tenchini & Frigerio, 2020, p. 287). In other words, although the slur is considered more offensive than the non-slur pejoratives in their study, the effect of reporting on offensiveness is the same. Additionally, they find that their results do not unilaterally support any one slur theory, but they do present problems for the theories that are situated at either ends of the spectrum, that is to say, Combinatorial Externalism and Prohibitionism (Tenchini & Frigerio, 2020).

Chapter 3 Methodology

In order to begin to bridge the gap between the theoretical discussion around slurs and the intuitions of regular native English speakers, I chose to employ a survey-based method. This choice is motivated by the fact that the research questions of this thesis pertain to native speaker intuitions or attitudes, and questionnaires are regarded as a well-suited medium for eliciting attitudes about language (Schleef, 2014). As we have seen, this is also the method used by all three of the previous empirical studies that investigated slurs in reported speech, which further reinforces the suitability of this method.

The study is based on two separate surveys, which I call Survey A and Survey B. Each survey elicited offensiveness intuitions for a number of selected words, which I call *target words*. These words can be sorted into categories that are similar to those used by Panzeri and Carrus (2016), Cepollaro et al. (2019), and Tenchini and Frigerio (2020). For this study, the most important of these categories are *slurs*, *neutral counterparts*, and *non-slur pejoratives*, all of which are used by the aforementioned studies.[5] Including words other than slurs is necessitated by the need for points of comparison and control. For example, it is easy to assume that slurs function differently to their neutral counterparts or to other pejoratives, but in light of the sparse amount of research in this area that would arguably be too strong an assumption to make. Including those categories in the surveys allows for this assumption to be examined. Beyond that, it also has the benefit of adding variety to the surveys, which not only makes it more interesting for participants but also encourages them to carefully read each utterance.

3.1 Participants & distribution

The surveys were created and administered online through a platform called Surveymonkey[6], which allows one to create surveys in exchange for a subscription fee. Links to the surveys were distributed through friends, acquaintances, and various social media platforms such as Twitter, Reddit, and Facebook. All participation was voluntary. The only restriction made on participants was that they be at least 18 years of age and native English speakers. Due to this

[5] While the studies do not use the exact same terminology for these categories, the categories are functionally the same.
[6] https://www.surveymonkey.com/

method of survey distribution, the participants come from a variety of backgrounds and speech communities. More on the participants and their demographics is presented in chapter 4.

3.2 Survey design

Survey A elicited offensiveness intuitions for the target words in isolation. Survey B elicited offensiveness intuitions for the target words in two different conditions—a *predicative* (non-reporting) condition as well as a *reporting* condition. To determine possible issues or weaknesses with Survey B, a pilot version of it was sent out to a small number of people. The main benefit of splitting the study into two separate surveys sent out beforehand was that it allowed for a large and diverse selection of stimuli while keeping the length and time required of the surveys fairly short. It also facilitated examination of a different sample of the population of speakers for their attitudes about the target words in isolation and in context. This created the benefit of providing judgments that could be cross-referenced with each other.

Both surveys asked the participants to rank the offensiveness of each stimulus on a 7-point rating scale, where a score of 1 is equated to 'not at all offensive' and 7 is equated to 'extremely offensive'. In addition, both surveys included a small number of open-ended questions pertaining to the content and/or format of the survey, which are outlined in subsequent subsections. Responses to these questions provided the study with feedback about the surveys as well as a metalinguistic component, and are later analyzed through primarily qualitative content analysis (see section 3.3).

Both surveys also asked for demographic information, which includes native language, country of birth, country of residence, age group, gender, and level of education. Asking for native language ensured that each response counted is from a native English speaker. Asking for country of birth and country of residence allowed the participants to provide an indication of what variety of English they speak. As we have seen in chapter 2, cultural and societal factors are often highlighted as important for a speaker's perception of slurs.[7] This makes it important to collect demographic information like country of birth and country of residence.

[7] The same can be said for all kinds of taboo words (see Jay, 2009).

3.2.1 Pilot survey

When it comes to survey research, it is always advantageous to conduct a pilot study (Schleef, 2014). In this case, it was determined that more benefit would be gained from conducting a pilot of the more complex survey of the two, which is Survey B. Therefore, a pilot of Survey B was constructed. It went out to 14 people, all of which have experience with tertiary education in English. The pilot survey received 10 complete responses. With the exception of one, none of the participants are native English speakers. While not ideal, this is not a significant issue, as the main purpose of the pilot was to test for structural issues, such as the phrasing of stimuli and instructions, length of time required, and other potential large scale problems. Table 1 displays the categories and target words used in the pilot survey.[8]

Table 1. The three categories and 15 target words used in the pilot survey

Slurs	**Neutral counterparts**	**Non-slur pejoratives**
chink	*Chinese*	*asshole*
faggot	*gay*	*bastard*
nigger	*black*	*cunt*
Paki	*Pakistani*	*dickhead*
retard	*disabled*	*motherfucker*

Participants were asked to rate the offensiveness of 35 utterances which include these target words, on a scale from 1 ('not at all offensive') to 7 ('extremely offensive'). See section 3.2.3 for a description of these utterances. The utterances are divided into predicative and reporting utterances. The pilot survey included five slur utterances in the predicative condition and ten in the reporting condition, five neutral counterpart utterances in the predicative condition and five in the reporting condition, as well as five non-slur pejorative utterances in the predicative condition and five in the reporting condition. The mean score on the rating scale, the *average offensiveness value,* was calculated for each category and can be seen in table 2. This is the main metric by which the results are later analyzed.

[8] The selection of target words is discussed in section 3.2.2

Table 2. The average offensiveness values of the different categories and conditions in the pilot survey (where a higher average indicates a greater degree of offensiveness)

Category	**Predicative**	**Reported**
Slur	6	4.03
Neutral Counterpart	2.46	1.38
Non-Slur Pejorative	2.12	1.72

Following the utterances, participants were presented with two open-ended questions pertaining to the subject matter of the survey, after which they were asked for demographic information. To conclude the survey, participants were asked whether they had any thoughts on the content or structure of the survey, as well as whether any parts of the survey were confusing or unclear.

Responses to the pilot survey resulted in a number of small changes to the survey and the stimuli.[9] The target words themselves did not present any issues for the participants, so they were retained. The changes implemented to Survey B as a result of the pilot survey consist of the following:[10]

- A reformulation of some utterances to not include any language that could be interpreted as evaluative.
- The addition of five predicative slur utterances to make the number correspond to the number of reporting slur utterances.
- Slight amendments to the phrasing of the instructions to be as accessible as possible.

3.2.2 Survey A

Survey A investigated attitudes to the target words in isolation. Acquiring these data allows for the offensiveness of these words to be observed on their own without other variables influencing the participants' judgments. In survey A, there are five categories of target words, which can be seen alongside the words themselves in table 3.

[9] The pilot survey results can be found in Appendix A.
[10] Two other changes were made upon internal reflection: The addition of a filler utterance for the sake of variety, and the addition of a concluding open-ended question.

Table 3. The five categories and 29 words used in Survey A

Slurs	**Neutral counterparts**	**Non-slur pejoratives**	**Non-insult taboo words**	**Neutral descriptive terms**
chink	*Chinese*	*asshole*	*bollocks*	*adult*
faggot	*gay*	*bastard*	*fuck*	*brunette*
gook	*Korean*	*cunt*	*Jesus Christ*	*doctor*
kike	*Jew*	*dickhead*	*shag*	*student*
nigger	*black*	*motherfucker*	*shit*	*tenant*
Paki	*Pakistani*			
retard	*disabled*			

These 29 words were selected according to a number of principles. Firstly, the slurs were selected due to their classification as slurs according to the central properties laid out in section 2.1. They were also selected to include both racial slurs and other kinds of slurs. Secondly, commonly accepted neutral counterparts of the slurs were selected. Thirdly, the neutral descriptive terms were copied over from Panzeri and Carrus (2016).[11] Finally, the non-slur pejoratives and non-insult taboo words were chosen due to their appearance in Millwood-Hargrave (2000), who conducted research into "people's attitudes to swearing and offensive language" in the United Kingdom (p. 1). All of the non-slur pejoratives and non-insult taboo words selected for this study were found by Milwood-Hargrave (2000) to be among the 25 most severe swear words in the English language.

The target words were presented to the participants in randomized[12] order, following these instructions:

> Below, you will find a list of words. For each of these words, please indicate how much you think that the word is offensive, using a scale that goes from 1 ("not at all offensive") to 7 ("extremely offensive").

Immediately following this section, participants were asked whether there are any words in the list that they do not understand, and if so, which ones. Preliminary responses to this question were used to inform the selection of target words for Survey B. As can be seen in table 4 in the following section, two of the slurs used in Survey A were removed due to the risk of

[11] Panzeri and Carrus (2016) translate 'castano' to *hair brown*, whereas I have chosen to instead use *brunette*, which is a more natural term in English.
[12] All randomization was performed on https://www.random.org/

comprehension issues. Finally, the survey concluded by asking for certain demographic information, as previously described.

3.2.3 Survey B

In Survey B, as in the pilot, each target word is put into a number of utterances, half of which are in the predicative condition, and half in the reported condition. Each slur appears in four utterances, two predicative and two reported, whereas each of the other types of target words appear in two utterances, one predicative and one reported. As such, including one filler item, there are 41 utterances in total. The target words and the categories they belong to can be seen in table 4.

Table 4. The three categories and 15 target words used in Survey B

Slurs	**Neutral counterparts**	**Non-slur pejoratives**
chink	*Chinese*	*asshole*
faggot	*gay*	*bastard*
nigger	*black*	*cunt*
retard	*disabled*	*dickhead*
Paki	*Pakistani*	*motherfucker*

The utterances containing these target words follow a general structure. The reporting utterances follow the structure of '[A]: [B] [reporting mechanism] [C] [predicate] [target word]', such as (10). The predicative utterances follow the structure of '[A]: [B] [predicate] [target word]', such as (11).

(10) Daisy: This kid called Claire a bastard.
(11) Edward: This guy is clearly an asshole.

Most utterances also contain added elements for the sake of variety, such as (12) and (13).[13]

(12) Stephanie: I heard Cam say that his girlfriend is Chinese.
(13) Jordan: I met a guy the other day, he's a Paki.

[13] See a list of all the utterances in Appendix C.

These utterances were presented to the participants with the following instructions:

> Below, you will find a number of statements. Imagine that you are overhearing these statements as part of a conversation amongst people you don't know. For each of these statements, please indicate how offensive you feel that **the speaker of the statement** is, always on a scale from 1 ("not at all offensive") to 7 ("extremely offensive").

The possibility of having all statements follow the exact same structure was considered, i.e., only changing names and target words, as Cepollaro et al. (2019) did. However, that method risks making participants complacent when filling out the survey, or not interested enough to complete it. I chose to include some variation in order to make the participants read every statement as new and unique, to be judged individually. This also makes it more difficult for participants to see through the survey, so to speak, and realize that the most important factor is the target word itself and the presence or absence of a reporting mechanism.

Steps were also taken to minimize subjectivity in the survey construction when it comes to the selection of the names used in the stimuli and the ordering of the items in the survey. The names were all selected from the 100 most popular baby names for girls and boys in England and Wales in 1996, as reported by the Office for National Statistics (2020).[14] The order of the items in the survey is pseudo-randomized. In this case, this means that the order was run through a random number generator and manually amended only when two utterances containing the same target word happened to be placed one after the other. At that point, the second of those two utterances was randomly placed in a different position.

3.3 Content analysis

This study is both quantitative and qualitative in nature. Both of the surveys included so-called open-ended questions, in response to which the participants wrote freely. These responses are treated partly as a way to evaluate the surveys and thus their results, and partly as metalinguistic comments of the offensiveness intuitions they express. In order to analyze these texts, a simple version of mostly qualitative content analysis is used. Weber (1990) identifies several ends to which one can use content analysis, one of which is the coding of responses to open-ended survey questions. The coding process looks like this:

[14] There is one exception: *Syed*. This name was added due to the lack of ethnic diversity in the most popular baby names. Additionally, some names were shortened to their nickname forms in order to appear less formal.

a) Through reading of the responses, common themes or functions are identified for the responses of each open-ended question.
b) These themes and functions are formalized into a typology of categories (e.g., *context* or *positive feedback*).[15]
c) Each response is sorted into one or two of the categories.

After this process is completed, it is done a second time, which ensures a measure of reliability which is called *stability* (Weber, 1990).

3.4 Ethical considerations

Consent is one of the most important components of ethical research practices (Eckert, 2013). In this study, participants were informed very generally of the contents, structure, and purpose of the survey in the introductory text, where they were also informed that their responses are anonymous. It was then stated that by clicking through to the first page of the survey, they consent to the use of their responses in this study. The anonymity is an aspect of the responses which is ensured by the survey platform—there is a setting one can enable which prevents any IP addresses from being collected by the platform itself. This setting was enabled for all three surveys used in this study.

Another aspect to consider in terms of consent is whether a survey clearly presents what its object of study is (Eckert, 2013). Are the participants made aware of exactly what the researcher is investigating, or is there any deception involved? For this study, there is no deception, but some information is withheld. Neither the focus on slurs nor the focus on reported speech is mentioned in the introduction or in the instructions. However, no secret is made of the interest in attitudes to offensive and taboo language in general. It is possible that some might find slurs to be a particularly upsetting or sensitive subject matter which they are not expecting to be faced with despite this description, and this may be considered an ethical issue. Indeed, some participants of Survey B remark upon the lack of content or sensitivity warning in relation to the slurs. But I did not want to advertise the survey as a survey that is about slurs, because this runs the risk of attracting participants with specific agendas as opposed to participants who are responding in good faith. Therefore, on balance, I believe the lack of content warning to be defensible. Moreover, participants who find the inclusion of slurs to be upsetting are free to exit the survey at any point. This issue is also part of why participants were

[15] A list of the categories identified can be found in Appendix D.

not offered any kind of incentive or reward for completing the survey, even though I may have been able to accrue more responses with the help of such methods. It was important that participants did not feel obligated or coerced to complete the survey, especially if the subject matter was not what they expected it to be or if it made them uncomfortable. As Eckert (2013) says, "[c]onsent should be not only informed, but voluntary" (p. 16).

3.5 Challenges & limitations

A particular limitation for this study is that, for ethical reasons, I was unable to enquire about demographic variables such as the participants' ethnicity, political affiliation, or religious denomination. Doing this would provide additional and important context to judgments about the offensiveness of most slurs in several ways. For one, those who are commonly targeted by particular slurs may find them, and slurs in general, more offensive than those who are not. For another, principles of self-determination dictate that those who are targeted by a slur are the ones with the most agency to judge the offensiveness of that slur (Nunberg, 2018). It follows that knowing this information about the participants would contextualize their responses. The inability to include this dimension in the study is a limitation.

Another issue is the number of variables involved in these offensiveness intuitions. Both the disclosed subject matter of the survey (offensive language), the format of the survey, and the form of the stimuli are variables which may in some way have shaped the participants' judgments. The participants were doubtlessly also affected by what they think is expected of them, both within the survey and as members of a speech community. As established by Labov (1972), the scientific study of language always involves the question of how to minimize the effect of the researcher, the observer, on the object of study, the observed. The fact of the matter is that it is impossible to ethically observe people without making them aware that they are being observed, which ostensibly affects their behavior. What one can do as a researcher is to limit one's effect on the object of study, and to be aware of this effect in one's analysis.

In regard to the stimuli specifically, the disadvantage to encasing the target words in varying utterances as opposed to strictly formulaic ones, is that aspects of the utterances other than the target words and the condition may affect judgments. Steps have been taken to make the relevant factors the most salient differing components of each utterance while retaining variety, such as the use of the common underlying structure. Moreover, comparing the results of Survey

B with those of Survey A should also allow us to see if offensiveness judgments of the target words in isolation are consistent with offensiveness judgments of the words in non-reported statements. If this were to be the case, that would point to a non-significant degree of interference from the varying content of the utterances.

Finally, in a survey that advertises itself as investigating ‘language and offensiveness’, it is perhaps expected that participants will anticipate offensive language and thus be more likely to judge something as offensive.

Chapter 4 Results

In this chapter, the results of the surveys are analyzed and presented in a descriptive manner. This means that they can be understood as an examination of the intuitions of these particular sample groups. Results show that the participants of both surveys find slurs in isolation and utterances involving slurs to be more offensive than any of the other categories of target words. Results also show that reported slur utterances are considered less offensive than non-reported slur utterances. The following sections present these results, and the demographics of the participants, in detail.

4.1 Survey A results

Survey A received 106 complete responses. Of these 106, 3 are from non-native English speakers and have therefore been excluded from the analysis, making the total number of complete responses 103. Of the 103 participants, 58 selected the option ‘female’, 37 ‘male’, and 8 ‘other’. This breaks down to 56% female, 36% male, and 8% other. In terms of level of education, participants were presented with three options: ‘secondary school’, ‘higher education’, or ‘other (please specify)’. All responses fall into the two main options, with 80 of the participants identifying themselves as having obtained some kind of higher education, compared to 23 participants who identified themselves as having taken part in secondary education. With regards to age, the participants are split almost evenly between the youngest age group, which encompasses 18-24, and the other age groups combined. See figure 1 for a breakdown.

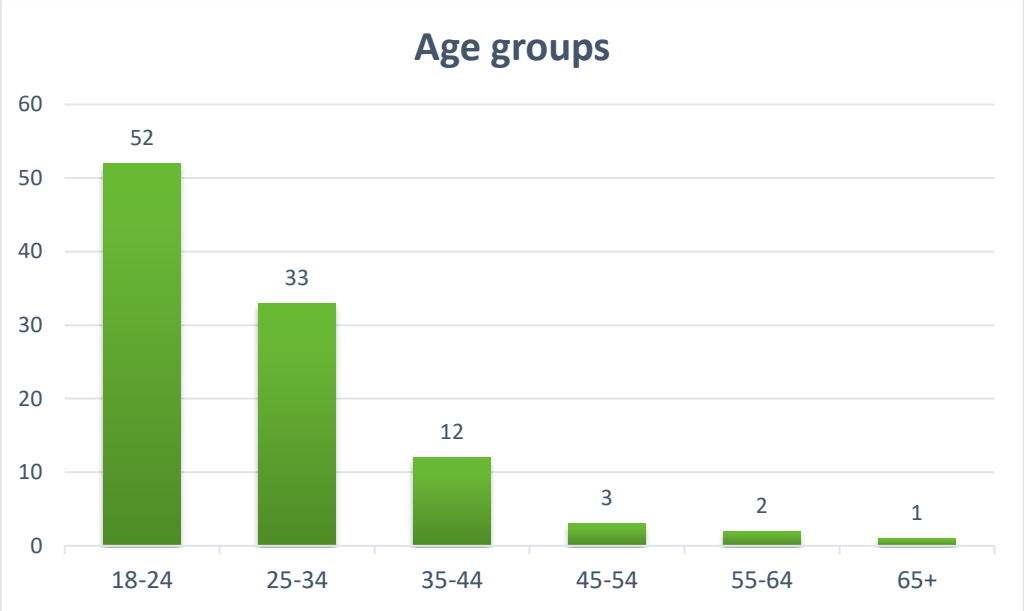

Figure 1. The number of participants of Survey A in the six different age groups

This figure tells us that 83% of the participants in Survey A are somewhere between 18 and 34, and 94% of them are between 18 and 44. Therefore, this survey can tell us far more about the attitudes of the younger age groups than those of the older ones.

Finally, the participants identified their country of birth and country of residence. Country of birth is taken as the best metric for determining the participant's variety of English. The distribution of countries of birth identified in this survey are presented in figure 2.

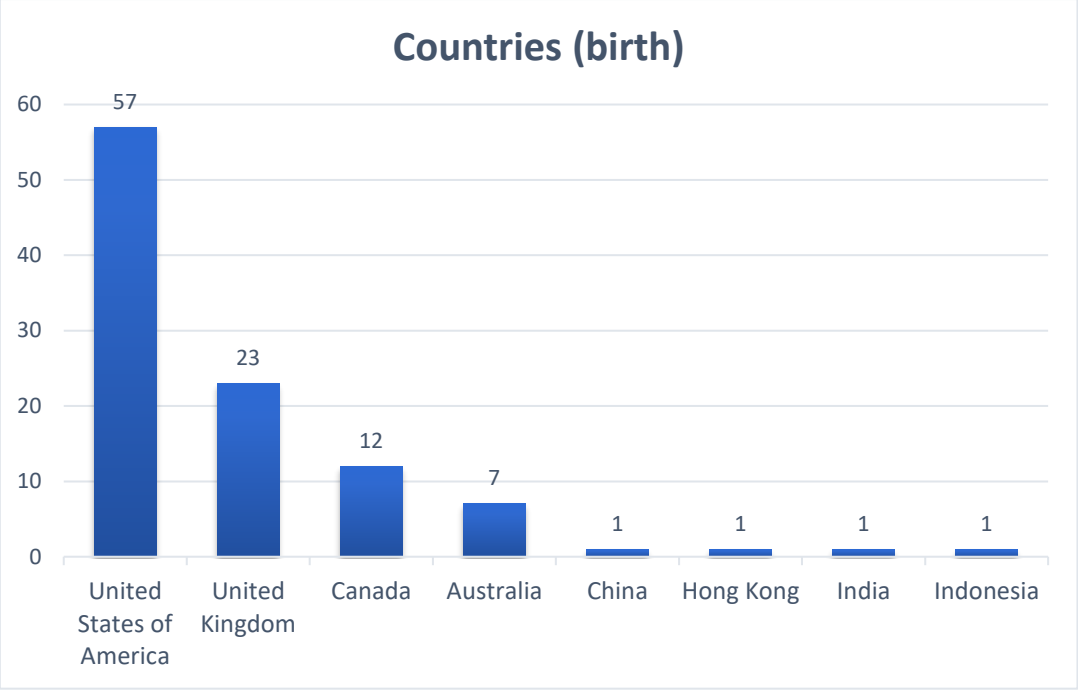

Figure 2. The number of participants of Survey A in the six different age groups

Here, we see that the majority of participants, 55%, are from the United States, but a significant portion are from the United Kingdom as well as other English-speaking countries.

Now that we have an idea of the participants' backgrounds, we turn to their offensiveness intuitions. As has been described, Survey A contains 29 words from five categories: *slurs*, *neutral counterparts*, *neutral descriptors*, *non-slur pejoratives*, and *non-insult taboo words*. Participants were asked to rate these words on a scale from 1 to 7, where 1 is 'not at all offensive' and 7 is 'extremely offensive'. After calculating the mean, the average offensiveness value of each word, we can calculate the average value of each category, which can be seen—rounded to two decimals—in table 5.[16]

[16] The average offensiveness values of the individual 29 words can be found in Appendix B.

Table 5. The average values of the five different categories

Categories	**Average offensiveness values**
Slurs	5.64
Non-slur pejoratives	3.97
Non-insult taboo words	2.57
Neutral counterparts	1.53
Neutral descriptors	1.03

As we can see, slurs are the most offensive and non-slur pejoratives are the second most offensive, followed by the non-insult taboo words. The neutral counterparts are not very offensive at all and the neutral descriptors are almost entirely non-offensive.

Shortly following the list of words, this survey also included a question pertaining to the participants' level of understanding of these words. They were asked the following question:

- Were there any words in this section that you did not understand?

To this question, 53 of the 103 participants chose 'yes'. They were then asked to specify which word(s) they did not understand. Table 6 shows which words were mentioned here and by how many participants they were mentioned.

Table 6. How many participants mentioned specific words as words they did not understand

Target words	**Participants**
Gook	27
Kike	26
Paki	11
Shag	6
Chink	4
Bollocks	1

Incidentally, these responses show some support for my subsequent choice to exclude *gook* and *kike* from Survey B. As we can see in this survey, these two appear to be the most likely to cause comprehension issues. Some participants mentioned in their responses that they looked up the words they did not understand, which means that their judgments of them are not necessarily devalued. However, they may be less valuable—or based on a different kind of understanding—than those of the words that did not cause comprehension issues. Additionally, as Survey B only includes five out of the seven slurs included in Survey A, it may be beneficial

to exclude the two additional ones from our comparison of the two surveys. Therefore, the average offensiveness values of the categories are presented again in table 7, this time excluding the responses for *gook* and *kike* from the slur category value, as well as their neutral counterparts *Korean* and *Jew* from the neutral counterparts category value.

Table 7. The average values of the five different categories, excluding the values from gook and kike and their corresponding neutral counterparts

Categories	**Average offensiveness values**
Slurs	5.73
Non-slur pejoratives	3.97
Non-insult taboo words	2.57
Neutral counterparts	1.47
Neutral descriptors	1.03

4.1.1 Open-ended questions

This survey contained two additional open-ended questions, which were placed on the final page of the survey:

1. Do you have any thoughts about the structure or content of this survey?
2. Was there anything in this survey that you found to be confusing or unclear?

These questions were optional. The first one received 34 responses, and the second one received 32 responses. After going through these responses according to the devised content analysis method, I can summarize them as follows.[17] On the whole, participants remarked on similar issues in responses to both questions. Of the responses to these questions, many are simple *no* responses, or responses with general *positive feedback* or *negative feedback* about the survey. Other responses fall into various categories related to the subject matter of the survey. For example, there are 4 responses that pertain to the issue of *offensiveness as a concept*, such as (14) and (15).

(14) Do I find this offensive... what does that mean? Like for instance, do I find the word fuck offensive? Well, I do not mind using it at all. I do not mind if others use it around me. But I think others find it offensive. So how should I rate the offensiveness of the word?

[17] Content categories are written in italics.

(15) I wasn’t sure how to rate a word like “bastard”. It's offensive to be *called* a bastard, but it's not what I’d think of as a swear word (unlike, e.g., motherfucker).

These responses show differences and uncertainties in the understanding of offensiveness and how to judge it. The other major category of responses are remarks on *context*, either on the *lack of context* provided in the survey, such as (16), or on the *importance of context* in general.

(16) The offensiveness of any word is based on context. Therefore, I included the most common usage for each word in my ranking of overall offense.

These types of answers make up 11 of the responses. Additionally, 3 responses note, in some way, that offensiveness has *cultural or regional differences*.

4.2 Survey B results

Of the 91 participants of Survey B, 49 selected ‘female’, 35 ‘male’ and 7 ‘other’. This translates to 54% female, 38% male, and 8% other, which is very similar to the gender distribution of participants in Survey A. For level of education, 84 identified themselves as having obtained some higher education, 6 selected ‘secondary school’ and 1 described having a practical degree. As for age demographics, there is less of a dominance of the youngest age group in this survey, as can be seen in figure 3. However, the majority of participants can still be placed in the two youngest age groups (63%).

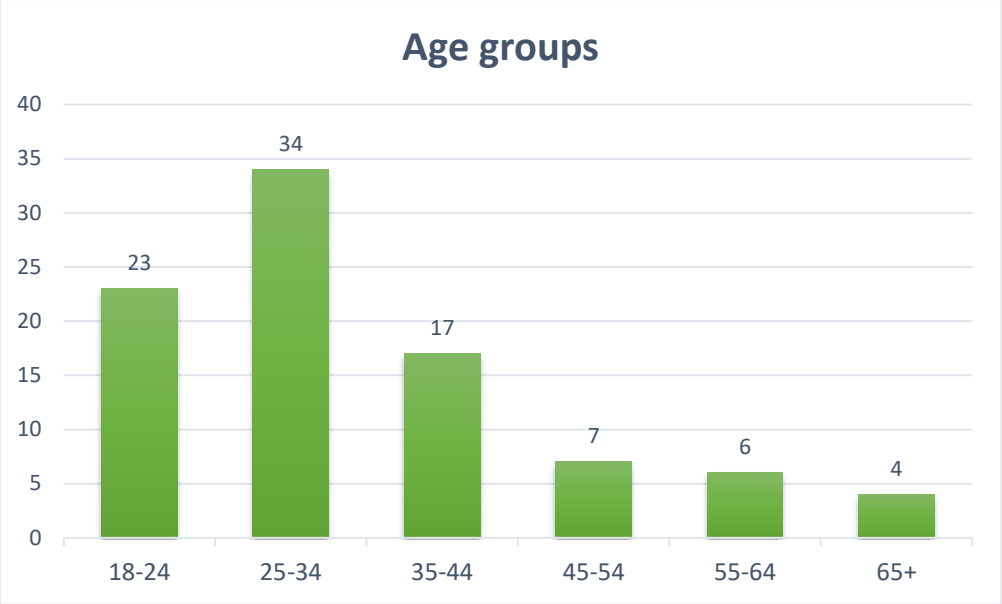

Figure 3. The number of participants of Survey B in the six different age groups

The distribution of countries of birth for the participants of this survey can be seen in figure 4.[18]

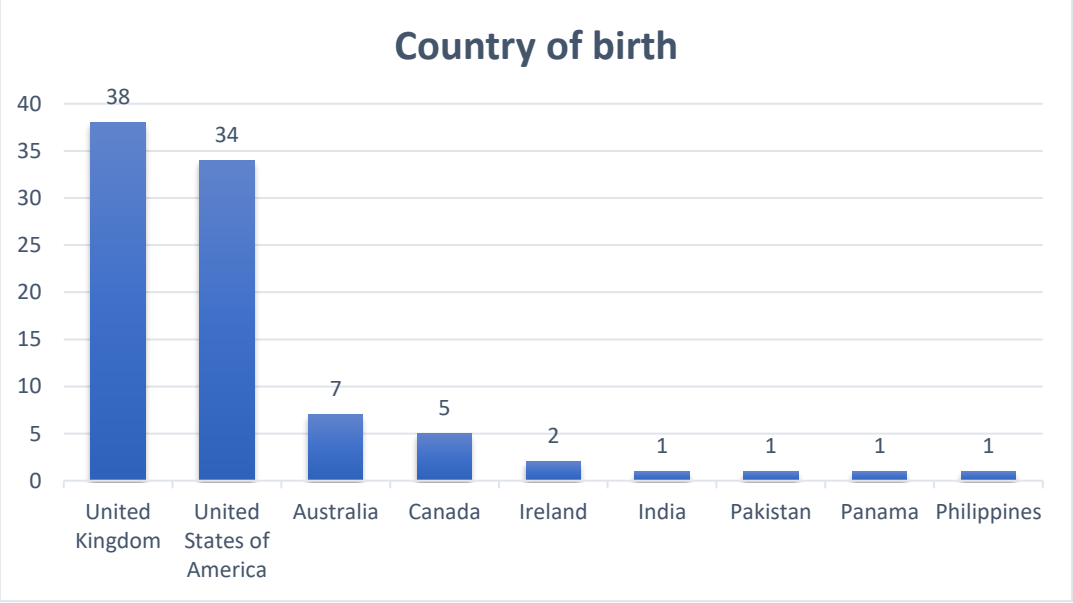

Figure 4. The distribution of countries that the participants of Survey B hail from

[18] One participant did not disclose their country of birth or country of residence, so they are not included in this breakdown. However, their offensiveness judgments are still counted.

This shows that in Survey B, participants from the United Kingdom outnumber participants from the United States, although not by much. These 91 participants provided offensiveness judgments that amount to the following average values, shown in table 8.[19]

Table 8. The average offensiveness values of the different categories and conditions in Survey B

Category	**Averages (predicative)**	**Averages (reported)**
Slurs	5.87	4.5
Non-slur pejoratives	2.85	2.4
Neutral counterparts	1.93	1.47

As table 8 describes, utterances containing slurs are judged to be more offensive than other utterances across the board. That slur utterances are considered the most offensive, followed by non-slur pejoratives, followed by neutral counterparts is aligned with our expectations, and aligned with the judgments of the target words in isolation as well. That the neutral counterparts have scores close to the 'not at all offensive' side of the scale shows that the utterances themselves do not appear to have caused offense when not including offensive terms.

Additionally, the average offensiveness values of the predicative utterances are all slightly higher than those of the reported utterances. To contextualize these averages, we can make use of their *standard deviations*, shown in table 9. The standard deviation value is a measure of how much variation there is in a dataset by how far the values tend to deviate from the average. Essentially, the smaller the standard deviation, the closer the responses are to the average.

Table 9. The standard deviations of the different categories and conditions in Survey B

Category	**Standard deviations (predicative)**	**Standard deviations (reported)**
Non-slur pejoratives	1.88	1.72
Slurs	1.64	2.23
Neutral counterparts	1.51	1.09

Here, we see that the standard deviation for the responses in the reported slur category is the highest, meaning that this category contains the most variation in responses. To make more

[19] In addition, the filler utterance has an average offensiveness value of 1.71.

sense of these numbers, further results will be presented in more detail in subsequent sections and later discussed in chapter 5.

4.2.1 Predicative condition

Utterances in the predicative condition that include slurs and non-slur pejoratives are both expected to be judged as offensive, at least to some degree. This is borne out by the survey responses. In Survey B, all five slurs occur in two utterances per condition. The average values of the predicative utterances containing slurs can be seen in table 10. All predicative slur utterances here have an average offensiveness value that is over 5, which is on the higher end of the scale.

Table 10. The average values of the slur utterances in the predicative condition

Slurs	**Average**
nigger	6.49
faggot	6.13
chink	6.01
Paki	5.39
retard	5.33
All	**5.87**

The judgments of the non-slur pejorative utterances are also interesting, because non-slur pejoratives are expected to behave similarly to slurs in non-reported speech, but not necessarily so in reported speech. To examine this, we must first inspect the average values of the predicative utterances containing non-slur pejoratives. These values are presented in table 11. However, note that each non-slur pejorative is only included in a single utterance within each condition.

Table 11. The average values of the non-slur pejorative utterances in the predicative condition

Non-slur pejoratives	**Average**
cunt	3.79
motherfucker	3.35
bastard	2.45
dickhead	2.35
asshole	2.3
All	**2.85**

Here we see that the predicative non-slur pejorative utterances are considered somewhat offensive, but considerably less offensive than most of the predicative slur utterances. Moreover, the utterances with *cunt* and *motherfucker* are considered the two most offensive ones. This mirrors the judgment of these two words in isolation in Survey A, where they are also the two most offensive words in their category.

4.2.2 Reported condition

Slur utterances in the reported condition are not always expected to be judged as offensive, but most theories include that possibility. In this survey, all reported utterances containing slurs are, on average, judged to be at least somewhat offensive—with the lowest average value being 3.53—but they are, overall, judged to be less offensive than their predicative counterparts. The average values of the reported slur utterances can be seen in table 12. Here, it appears that utterances containing [n-slur] and [f-slur] are less ameliorated by the reporting mechanism than the other utterances in the same condition.

Table 12. The average values of the slur utterances in the reported condition

Slurs	**Average**
faggot	5.42
nigger	5.19
chink	4.32
retard	4.05
Paki	3.53
All	**4.5**

As has been mentioned, non-slur pejoratives are expected, by most, to behave differently to slurs in circumstances like reported speech. Non-slur pejoratives are expected to be displaced by reported speech, unlike slurs. The average offensiveness values pertaining to utterances with reported non-slur pejoratives are shown in table 13.

Table 13. The average values of the non-slur pejorative utterances in the reported condition

Non-slur pejoratives	**Average**
cunt	3.21
motherfucker	2.64
dickhead	2.14
asshole	2.03
bastard	1.96
All	**2.4**

This shows that although there is a difference, the predicative and reported utterances with non-slur pejoratives are judged similarly in terms of offensiveness.

4.2.3 Open-ended questions

Survey B contained four open-ended questions, none of which were required for completing the survey.

1. If you would like to explain or elaborate on any of your responses so far, please do so here:
2. Do you think there are words that are always offensive?
3. Was there anything in this survey that you found to be confusing or unclear?
4. Do you think that it's offensive to use slurs when reporting what someone else has said?

The first two of these questions were placed directly after the last utterance, and the other two were located at the very end of the survey. Question 1 received 55 responses, question 2 received 73 responses, question 3 received 51 responses and question 4 received 87 responses. Here, I present a general summary of their contents according to the present content analysis.

There were 22 responses to question 1 which entailed direct comments on the difference between using insults or slurs in *non-reported versus reported speech*. Some of these responses said that reporting it makes the insult or slur non-offensive, while several others said it

sometimes makes it less offensive but that this depends on context. There were also 4 responses that discussed the issue of *in-group usage*, such as (17).

> (17) I was aaauming [sic] the speakers were white, able-bodied and cis-het. Were they targeted by the words they were using, the offense of them being spoken would be less.

It was also noted by 15 participants that their own or other people's responses might be shaped by *cultural or regional differences,* such as different varieties of English. For example, several participants remarked that *Paki* has a lesser offensive effect outside of the United Kingdom. Others said that *cunt* is less taboo in Australia and/or the United Kingdom compared to the United States. This is a category of responses that pertains to *context* in some respect, and several responses went into other facets of context. For example, (18) includes critique levied at the survey for not providing adequate context to make informed judgments of offensiveness.

> (18) Your survey failed to capture the varying levels of offense depending upon the audience & their relationship to the speaker.

Question 2 is phrased in a binary way and therefore garnered binary responses, although some participants elaborated further. Of the 73 participants to respond to this question, 54 indicated their belief that there are words that are always offensive, while 16 said that this is not the case. The remaining 3 did not take an explicit stance.

Question 3 has to do with the survey itself and the participants' understanding of it. Of its 51 responses, 34 of them are simple negative answers, as in, the participants did not find anything to be confusing or unclear. Of the remaining 17, some participants again commented on the sparse nature of the *context* provided, such as in (19). In (20), one participant remarked on how this lack of context affected their own responses. Additionally, a couple of participants noted that they were not asked for their ethnicity or similar demographics, which was, as has been stated, due to ethical reasons.

> (19) I may have answered to third party statements differently if I knew in what manner and to what purpose they were said.

(20) I was making unintended assumptions about race and gender of the speakers as I considered my response.

Finally, question 4 directly asked participants whether they think it is offensive to use slurs when reporting what someone else has said. Of the 87 participants who responded to this question, 69 believed that this is offensive to some degree, 1 participant said they did not understand the question, and the remaining 17 said they do not find it offensive. Some of the 69 participants wrote *yes* and offered little additional information, but many of them also said that the level of offensiveness depends on *context* and that offensiveness can be stronger or weaker depending on situation, speaker-addressee relationship, culture, and other factors. For some participants, it is clear that the main factor dictating offensiveness is the audience or the addressee. See (21), for example.

(21) Sometimes. It depends whether you think the person you're speaking to might be offended. I think they would be less offended if It's reporting what someone else said, as opposed to the speaker choosing the word themselves. But some people are offended just from hearing a taboo word regardless of context.

This response hints at the idea of a social or cultural conception of offensiveness as different from a personal conception of offensiveness. It also expresses the idea that reporting slurs, while offensive, is *less offensive than using them directly*, which is another content category that 10 of these 69 responses can be sorted into.

Chapter 5 Discussion

5.1 Slurs as a category

Before addressing the research questions, I would like to make a note of the Survey A results. As we have seen, these results show that slurs are the most offensive category of words out of the five categories used in that survey. That slurs would be the most offensive is perhaps to be expected, but it is still interesting that there is such a big difference between the slurs and the non-slur pejoratives (5.73 versus 3.97). This remarkable difference can perhaps be said to empirically reinforce the theoretical perception of slurs as their own category of pejoratives. A caveat to this is of course that results may have been different with other non-slur pejoratives that are considered more or less taboo than the ones used here.

In the following sections, the results of this study, which encompass both offensiveness judgments and metalinguistic comments, are discussed according to the research questions.

5.2 Non-reported & reported slurs

RQ1 asks the following:

- RQ1: Does offensiveness differ between slurs in non-reported speech and indirect reports? If so, how?

Ideally, the results of the surveys would be able to provide us with an answer to this question, at least as it relates to the sampled population. This is possible provided that there are no outliers in the results caused by irrelevant variables. To ensure this, the results of Survey A and the predicative condition of Survey B can be compared. As described in section 4.2, the offensiveness averages of the categories are similar between Survey A and Survey B. This is also, broadly, the case when individual target words are compared, as can be seen in table 14. Arguably, this comparability increases the validity of the judgments provided in Survey B. This is because the predicative slur utterances in Survey B were expected to be judged similarly to the slurs in isolation if no irrelevant variables were able to skew the judgments.

Table 14. The average values of the slurs in isolation and in the predicative condition

Target word	**Average (Survey A)**	**Average (Survey B, predicative)**
nigger	6.68	6.49
chink	6.28	6
faggot	5.88	6.13
retard	5.36	5.33
Paki	4.46	5.4

Although it is not a perfect correspondence between the surveys, the results are arguably similar enough that this could be put down to natural variation between the participants. For example, more of the participants of Survey A were from the United States whereas more of the participants of Survey B were from the United Kingdom. This could have been the cause of *Paki* being judged as more offensive by the participants of Survey B, as this term is more of a slur in British English than American English (Merriam-Webster, n.d.). The fact that there are no other large divergences lends credence to the average judgments provided in Survey B as a whole.

On the other hand, it is also possible that the lack of personal relationship with the speaker, and/or the reported speaker, from the point of view of the participants, made them more likely to judge the utterances in Survey B as offensive. This is an outcome that is implicitly suggested by Capone (2012, 2016), from the reasoning that knowledge of the reporter/reported plays a major part in to whom one attributes the content of the complement clause of a reported slur utterance.

With both of these dimensions established, we turn to the differences between the judgments provided for the predicative (i.e., non-reported) and reported conditions. The average offensiveness values from Survey B indicate that utterances including slurs are considered more offensive than utterances in any other category, but that reported slurs are less offensive than non-reported ones. On its own, this is indicative of two important things, which is that both non-reported and reported slurs are considered offensive, and that there is a difference in degree of offensiveness between them. Additionally, responses to the open-ended questions also show that reported slurs are considered offensive, even though only a small number of participants explicitly mentioned finding reported slurs less offensive than non-reported ones.

Another factor to consider is that the standard deviations for these averages tell us that, of all the categories, the reported slurs category contains the most variance from the mean. This means that some participants may have judged reported slurs to be equally offensive to non-reported ones, some that they are only slightly offensive or somewhere in the middle of the scale as opposed to on the high end, and so on. Additionally, some participants may see all reported slurs as highly offensive, and some may judge certain slurs as more offensive regardless of the condition. Either way, the standard deviations show that the reported slurs category is the most contested.

5.3 The case of non-slur pejoratives

RQ2 asks the following:

- RQ2: What can the results of the questionnaires tell us about the differences and/or similarities of slurs and non-slur pejoratives?

From the results outlined in chapter 4, it appears that non-slur pejoratives are considered to be less offensive than slurs, both in isolated, non-reported, and reported forms. Additionally, the average offensiveness value of non-slur pejoratives is higher in isolation than in predicative utterances (3.97 versus 2.85). Comparatively, the average offensiveness value of slurs is around the same for slurs in isolation and in predicative utterances (5.73 versus 5.87). This result differs from some of the empirical studies previously highlighted, such as Cepollaro et al. (2019), who found that slurs are perceived as less offensive in predicative utterances than in isolation, and that non-slur pejoratives are more offensive in predicative utterances than in isolation. This may be due to a difference in methodology (e.g., their utterances being entirely formulaic, and/or their inferential statistical analysis) or simply a difference in the sampled population. Future research may be able to further elucidate this issue.

It is somewhat surprising that the difference in offensiveness between non-slur pejoratives in the predicative and reported conditions is so small (2.85 versus 2.4). By most accounts, non-slur pejoratives should be less offensive in reports than in predicative statements.[20] That the difference is small may be due to the form of the survey, to variability in the utterances

[20] An exception to this is the Prohibitionism account, at least with regards to non-slur pejoratives that are considered taboo (i.e., vulgar, or profane), such as *asshole*, *fucker*, *dickhead*, etc. Prohibitionism dictates they should be equally offensive in either condition.

themselves, to non-slur pejoratives being more similar to slurs than expected, or to something else, such as an indiscriminate taboo effect. A possible explanation is that because the survey is advertised as pertaining to offensive language, the participants are primed for offensiveness and therefore judge utterances to be offensive to a higher degree than in natural circumstances. Essentially, this could be an effect of the survey that was not mitigated. Alternatively, it is possible that, despite my best efforts, the reported non-slur pejorative utterances were incidentally phrased in such a way that made them more offensive than other utterances.

Another possible explanation is that non-slur pejoratives are more similar to slurs than anticipated. Perhaps their derogatory component also projects out of reports, and thus they too possess some form of the nondisplaceability property. Theorists who see slurs as expressive terms, such as Potts (2007) and Cepollaro (2015), would likely find this easy to accept, as many non-slur pejoratives and other taboo words are also considered to be expressives.

Although there might be several more possible explanations for this result, another that comes to mind is that these terms are simply perceived as taboo and are therefore judged as offensive to some extent regardless of reporting mechanism or syntactic context. Both the individual participants' feelings about taboo words in general, and their understanding of the concept of offensiveness, could have caused them to find these words offensive, regardless of context. Of course, there are sanctions of different degrees against all kinds of taboo words that apply to every use of them, even reporting or metalinguistic uses (Anderson & Lepore 2013b). How offensive one finds taboo words then depends on one's level of dedication to or respect for those sanctions, a sentiment we previously noted in (21). Additionally, a participant responded to the question "[d]o you think there are words that are always offensive?" with (22).

(22) Yes. Swear words and blasphemous statements.

It would be reasonable to surmise that non-slur pejorative utterances are judged as somewhat offensive due to this inherent taboo status. This explanation is one that proponents of Prohibitionism would gravitate towards. Whatever the explanation, the results point to a lesser degree of offensiveness in non-slur pejoratives compared to slurs, and to a very small difference between the conditions.

5.4 Slur theory compatibility

RQ3 asks the following:

- RQ3: Which of the selected slur theories, if any, are compatible with the results of the questionnaires?

To begin answering this question, let us examine the results as they relate to the theories with the clearest predictions—Combinatorial Externalism and Prohibitionism. If the results were in accordance with Combinatorial Externalism, reported slurs would likely be scored either low or very low on the offensiveness scale across the board. This is not the case. On average, participants of Survey B did consider reported slurs to be offensive, and in their open-ended responses, 69 of 87 explicitly stated as much. In other words, our results run counter to Combinatorial Externalism. As has been described, Hom's (2012) counterpoint to the apparent offensiveness of reported slurs is that his theory only dictates the conditions of derogation, and that this is a separate phenomenon from offensiveness. According to Hom (2008), a slur is not derogatory unless it is applied to someone. The question therefore remains: Can offensiveness be separated from derogation? Is this a valid or meaningful differentiation? See (23), written by a participant of Survey B, which directly connects derogation with offensiveness.

(23) There are derogatory words that I find always offensive when I hear them.

As seen from Culpeper (2011), offense can be provoked by many different things, some of which relate to threatening face or going against social conventions. The word derogatory, on the other hand, can be defined as "[h]aving the effect of lowering in honour or estimation" (Oxford English Dictionary). Most of Culpeper's (2011) offense types can arguably be connected to such an effect to some extent. But even if some types of offense are not connected to derogation, this is not necessarily a distinction that is relevant in this context. For example, when asked if it is offensive to use slurs when reporting what someone else has said, one participant wrote (24).

(24) Yes. Well, not exactly. I think it's harmful, and really, my responses were indicating that I found the words to be harmful. Offense is a meaningless concept.

This sentiment of what is and is not harmful recurs in the open-ended responses in conjunction with comments about what is and is not acceptable, appropriate, or offensive. Some participants, like the writer of (24), valued harmfulness over offensiveness, while others did the opposite. However, considering the history of all slurs as derogatory terms, as well as the fact that many participants were clearly aware of the dimensions of derogation, harm, and/or dominance relations inherent to slurs, it does it does not seem likely to me that replacing *offensive* with *derogatory* in the survey instructions would lead to wildly different results. This is effectively what would be required for the results of this study to be in any way compatible with Combinatorial Externalism. Finding evidence to support or falsify this assumption could, however, be a potential worthwhile endeavor for future research.

If the results were in compliance with Prohibitionism, both non-reported and reported slurs would be scored as offensive to a similar degree. This is closer to our actual results, but in terms of average values, reported slurs are considered less offensive, which Prohibitionism cannot easily account for. Recall that Anderson and Lepore (2013a) write that indirectly reporting a slur does not communicate that the original speaker is guilty of using a slur. Instead, it "**guarantees** an offense" by the reporter (p. 29, my emphasis). However, Prohibitionism does dictate that contextual variables have the ability to mitigate the offensiveness of slurs in reported utterances. The present study was not equipped to test for such expanded context. Variation in judgments aside, the fact that both reported and predicative slur statements have an average score of above 4 on a 7-point scale means that the participants mostly agree that reporting a slur does incur offense. In this sense, our results do not contradict Prohibitionism. In another sense, the difference in offensiveness between the two conditions does call into question whether a blanket prohibition against these terms can really be the reason for the offense. Prohibitionism would have to be modified or expanded in some way to be able to explain this differing degree of offensiveness, at least if this finding is something that can be replicated in other studies.

The Conventional Implicature and Presupposition-based theories both dictate that the offensiveness of reported slur utterances can vary between offensive and non-offensive. For Conventional Implicature theory, it is the person whose attitudes are being expressed through slur usage that causes offense. If this is the original or the reporting speaker ostensibly depends on contextual variables and/or subjective interpretation. Observe a statement like (25), which

is part of Survey B, and consider who this statement's 'contextual judge' is—to use Potts' (2007) terminology. Then, take (26), also part of Survey B, and ask the same question.

(25) Sean: But Paul said that Thomas is a nigger
(26) Leah: The presenter called that guy a nigger on national TV

Here, (25) may be expected to cause more offense than (26) if it is more likely for Sean to be interpreted as the contextual judge than Leah. To some extent, this does seem to be the case, as in Survey B, (25) has an average offensiveness value of 5.79 while (26) has a value of 4.58. Both are, however, still significantly offensive. This suggests that if the Conventional Implicature theory is applicable here, then the identity of the contextual judge—the source of offensiveness—is up to interpretation (especially in cases of sparse contextual information).

An argument against the Presuppositional theory of slurs is that presuppositions are expected to be cancellable when embedded, and this does not seem to be the case with embedded slurs (Anderson & Lepore 2013a). Cepollaro (2015) deals with this deficiency by arguing that (a) embedded presuppositions do not have to be cancellable—it is reasonable to say that some presuppositions "can be harder to cancel than others" —and that (b) it is possible to argue that embedded presuppositions triggered by slurs are cancellable (p. 40). Relying on the former counter-argument means that the Presuppositional account of slurs rests on the concession that some embedded presuppositions are not cancellable. Relying on the latter counter-argument means that the Presuppositional account rests on the capacity of slurs' derogatory content to be cancelled. Considering that the results of this study find that slurs are offensive in both reported and non-reported contexts, the latter counter-argument stands on shaky grounds. The former, I leave to those more knowledgeable about the topic of presuppositions than I.

Finally, for Nunberg's (2018) Affiliationism, the offensiveness of reported slurs in Survey B would potentially vary between offensive and non-offensive depending on the conversational relevance of the slur. This idea is reflected in (27), a comment from a participant of Survey B on the offensiveness of reported slurs.

(27) Depends on context and intent – does repeating the exact slur improve communication? If not, then no need to repeat verbatim.

Nunberg's (2018) premise is that usage of a marked term (i.e., a slur), over an unmarked term, has to be justified by informational relevance in order to be non-offensive. Considering the very limited context that is included in the utterances used for Survey B, it is not possible to know whether the slurs in them can be considered justified according to this premise. Likely, most of them would not be able to overcome the default reading. However, some of our utterances may be more likely than others to be interpreted in this way. For example, see (26), (28), (29), and possibly (30).

(28) Ethan: Last time I spoke to Rhiannon, she called Elliott a chink
(29) Kyle: John called Simon a chink the other day
(30) Jodie: Last night, I heard Maria say her boyfriend's a retard

In these cases, it could be argued that the point of the utterance is to convey that the original speaker's use of a slur is notable in itself. One could imagine these utterances being said in a context of outrage, conveying to the addressee that whoever is being reported harbors racist or bigoted attitudes by virtue of their norm breaking choice of words. Replacing the slurs with their neutral counterparts in these statements would fail to convey the original speaker's bigoted attitudes, at least not without additional context, periphrasis, or use of euphemisms (e.g., *n-word, c-word,* etc.). In studies that incorporate more extensive context into the stimuli, this could be investigated further. However, it is important to note that statements which do not have this reading are still expected to be judged as offensive within Affiliationism. Therefore, the results of the study do not contradict Affiliationism.

In summary, Combinatorial Externalism is somewhat incompatible with the results, as the average offensiveness of reported slur utterances is so high. Prohibitionism is not entirely contradicted by the results, but the difference between the offensiveness of predicative and reported slur utterances calls for further development of the theory in order to be compatible with present findings. Conventional Implicature theory, Presupposition theory, and Affiliationism are all neither contradicted nor supported by the results, but some results suggest that further research tailored specifically to the parameters set up by these theories may be fruitful in terms of definitively supporting or falsifying their assertions.

5.4.1 Degrees of offensiveness

One thing that is clear from the results is that there is notable variation in the participants' judgments, particularly as it relates to reported slurs. This begs the question; how are the slur theories positioned in relation to such variation, and could they explain it? As has been discussed, several theories, including Conventional Implicature and Presupposition theory, predict that certain statements involving reported slurs are either offensive or non-offensive, depending on the addressee's or hearer's interpretation of the context or contextual judge (Potts, 2007; Schlenker, 2007). There is little consideration for differences in degrees of offensiveness. This issue has also been pointed out by Cepollaro et al. (2019), whose results also show that reported slurs are less offensive than non-reported slurs.

Differing degrees of offensiveness, at least in this study, likely have something to do with the different degrees of offensiveness of the slurs themselves, as [n-slur] is considered more offensive than [r-slur] or [p-slur], and so on.[21] But responses to the open-ended questions indicate that there are other factors that can influence this judgment as well. For example, see (31) and (32), in which participants are responding to the question about the offensiveness of reported slurs.

> (31) To a greater or lesser degree depending on the slur, the speaker's attitude towards it, and to some extent the awareness of the speaker (for instance, is the speaker a four-year-old or someone with a poor grasp of English?).
>
> (32) Yes, as opposed to a circumlocution or other strategy, but often to a lesser degree, and especially if tone, context or prior knowledge strongly indicate disagreement

It is clear that for these participants—at least based on their metalinguistic comments—these factors do not eliminate the offensiveness of the slur. Rather, they modulate it. A possible interpretation of the variation found in the survey results is thus that we would benefit from changing the way we talk about reported slurs and offensiveness—in reality there will be a large proportion of people who always find them offensive but the degree to which they do so may be dependent on the factors outlined by some of the slur theories. Additionally, we may

[21] This is supported by the offensiveness judgments of these slurs in isolation, which are reflected not entirely but to a large extent by the judgments of these slurs in the reported condition.

want to talk about primary attribution versus secondary. Reported slurs can be expected to be attributed primarily to the original speaker for her act of slurring, and secondarily to the reporting speaker for failing to avoid the slur, or something like that (Capone, 2016). Although, I suspect that this line of reasoning is especially relevant to the behavior of the more central slurs as defined by this study in section 2.1, and perhaps less so to peripheral members of the category, which may be associated with weaker prohibitions and cultural conventions, as well as shorter histories.

Cepollaro et al. (2019) suggest a way of solving this problem for the Presuppositional account. Recall that the Presuppositional account claims that slurs can be "evaluated with respect to any context" (Schlenker, 2007, p. 239). Cepollaro et al. (2019) state that since reported slurs can be varying degrees of offensive, there must be two interpretations of context "available to a speaker at the same time" (p. 41). These two interpretations compete, and when contextual variables do not encourage a particular reading, they lead to perceptions of "intermediate" offensiveness (Cepollaro et al., 2019, p. 41). This addition to the Presupposition theory could explain the current results as well as the similar results found by Cepollaro et al. (2019).

Other theories are not as poised to explain these findings, although there may be ways for them to do so. For one, Combinatorial Externalism does not consider slurs to be derogatory in reported speech, and their potential offensiveness is something that is described as subjective (Hom, 2012). It may be possible for Combinatorial Externalism to explain our results through some interplay between subjective interpretation and the strength of the social institutions that support the derogatory content of slurs. Conventional Implicature theory and Affiliationism, on the other hand, depend on interpretations partially based in context, and do not explicitly discuss how reported slurs may be perceived as lesser degrees of offensive, but rather either offensive or non-offensive.

Thus, that there is so much variation in the degrees of offensiveness of reported slurs, not all of which can be explained by differences in the slurs themselves, is not something that most of the slur theories can readily accommodate for. However, it is not clear how big of an issue this is, considering the sparsity of research on the subject. This is something that future research could delve into further, both to determine the scale of the issue and to explore potential explanations.

Chapter 6 Conclusions

Firstly, I can conclude based on this study that both non-reported and reported slurs are considered offensive, and that there is a difference in degree of offensiveness between them. Reported slurs are, on average, judged to be less offensive. This is not a problematic result for the most pragmatic-oriented theories such as Conventional Implicature theory, Presupposition-based theory, and Affiliationism, but it is also not something that they explicitly expect to see from speakers. It is a problematic result for Combinatorial Externalism and, to a lesser extent, Prohibitionism, as the former expects reported slurs to be non-offensive and the latter expects them to be as offensive as non-reported slurs. These findings therefore have similar implications to those of Cepollaro et al. (2019) and Tenchini and Frigerio (2020).

Secondly, it is clear that the results point to a lesser degree of offensiveness in non-slur pejoratives compared to slurs, and to a comparatively much smaller difference in offensiveness between non-reported and reported non-slur pejoratives. That the difference between non-reported and reported non-slur pejorative utterances is so small is somewhat unexpected, and I have discussed several possible explanations for this result. Ultimately, these findings reinforce the perception of slurs as more offensive than non-slur pejoratives, and pose no particular challenges for the slur theories.

This study has aimed to explore the merits of selected slur theories as they relate to slurs in reported speech. After examining the compatibility of these slur theories with the results of the study, several observations have been made. For example, this study has found results that are explicitly incompatible with Combinatorial Externalism, since reported slurs have been judged as offensive. Hom's (2012) preemptive explanation of this kind of result rests on the separation of derogation and offense, but as I and others have argued (e.g., Jeshion, 2013), this separation is not necessarily effective. As has been discussed, responses to some of the open-ended questions also suggest that this separation is not always found in the minds of speakers.

In contrast, Prohibitionism, Conventional Implicature theory, Presupposition theory, and Affiliationism have been found to be somewhat compatible with the results, to different degrees. Prohibitionism especially requires some modification to be able to fully explain these results. In the future, research that is better equipped to test for differences in context can further

examine the predictive and explanatory value of these theories. Moreover, no theory—except Presupposition theory according to Cepollaro et al. (2019)—is poised to explain the varying degrees of offensiveness found in the offensiveness judgments of reported slurs. This issue arguably requires more extensive examination and analysis to fully explore.

In the meantime, this thesis reinforces the perception that slurs are more offensive than most other types of taboo words and/or pejoratives. It also implies that slurs are likely nondisplaceable to some degree in all reported contexts. This implication, in itself, furthers our understanding of slurs in the English language. It is my hope that this thesis will inspire further empirical research on the topic so that this understanding can continue to grow.

References

Anderson, L., & Lepore, E. (2013a). Slurring words. *Nous, 47*(1), 25–48.

Anderson, L., & Lepore, E. (2013b). What did you call me? Slurs as prohibited words. *Analytic Philosophy, 54(*3), 350-363. DOI: 10.1111/phib.12023

Ashwell, L. (2016). Gendered slurs. *Social Theory and Practice, 42*(2), 228-239. DOI: 10.5840/soctheorpract201642213

Brown, P., & Levinson, S. C. (1987). *Politeness: some universals in language usage*. Cambridge: Cambridge University Press.

Camp, E. (2018). A dual act analysis of slurs. In D. Sosa (Ed.), *Bad Words: Philosophical Perspectives on Slurs*, 29-59. Oxford: Oxford University Press.

Capone, A. (2012). Indirect reports as language games. *Pragmatics & Cognition, 20*(3), 593-613. doi 10.1075/pc.20.3.07cap

Capone, A. (2016). Slurring in indirect-reporting. *International Journal of Language Studies, 10*(4), 19-36. ISSN: 2157-4898

Cepollaro, B. (2015). In defence of a presuppositional account of slurs. *Language Sciences, 52*, 36-45. http://dx.doi.org/10.1016/j.langsci.2014.11.004

Cepollaro, B., Sulpizio, S. & Bianchi, C. (2019). How bad is it to report a slur? An empirical investigation. *Journal of Pragmatics, 146*, 32-42. https://doi.org/10.1016/j.pragma.2019.03.012

Cousens, C. (2020). Are ableist insults secretly slurs? *Language Sciences, 77*, 1-11. https://doi.org/10.1016/j.langsci.2019.101252

Croom, A. (2011). Slurs. *Language Sciences, 33*, 343-358. doi:10.1016/j.langsci.2010.11.005

Culpeper, J. (2011). *Impoliteness: Using language to cause offence*. Cambridge: Cambridge University Press.

Davidson, D. (1968). On saying that. *Synthese, 19*(1), 130-146. https://www.jstor.org/stable/20114635

Diaz-Legaspe, J. (2020). What is a slur? *Philosophical Studies, 177*, 1399-1422. https://doi.org/10.1007/s11098-019-01259-3

Eckert, P. (2013). Ethics in linguistics research. In R.J. Podesva & D. Sharma (eds.), *Research methods in linguistics*, 11-26. Cambridge: Cambridge University Press.

Embrick, D. G. & Henricks, K. (2013). Discursive colorlines at work: How epithets and stereotypes are racially unequal. *Symbolic Interaction, 36*(2), 197-215. DOI: 10.1002/SYMB.51

Grice, P. (1989). *Studies in the way of words*. Cambridge, Massachusetts: Harvard University Press.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania.

Lasersohn, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy 28*(6), 643-686. DOI 10.1007/s10988-005-0596-x

Holt, L. F. (2018). Dropping the 'n-word': Examining how a victim-centered approach could curtail the use of America's most opprobrious term. *Journal of Black Studies, 49*(5), 411-426. https://doi.org/10.1177/0021934718756798

Hom, C. (2008). The semantics of racial epithets. *Journal of Philosophy, 105*(8), 416–440. https://www.jstor.org/stable/20620116

Hom, C. (2012). A puzzle about pejoratives. *Philosophical Studies, 159*, 383–405. https://doi.org/10.1007/s11098-011-9749-7

Jay, T. (2009). The utility and ubiquity of taboo words. *Perspectives on Psychological Science, 4(*2), 153-161.

Jeshion, R. (2013). Slurs and stereotypes. *Analytic Philosophy, 54*(3), 314-329.

Jeshion, R. (2016). Slur creation, bigotry formation: The power of expressivism. *Phenomenology and Mind, 11*, 130-139. DOI: 10.13128/Phe_Mi-20113

Merriam-Webster. (n.d.). Offensive. In *Merriam-Webster.com dictionary*. Retrieved June 30, 2021, from https://www.merriam-webster.com/dictionary/offensive

Merriam-Webster. (n.d.). Paki. In *Merriam-Webster.com dictionary.* Retrieved September 10, 2021, from https://www.merriam-webster.com/dictionary/Paki

Merriam-Webster. (n.d.). Slur. In *Merriam-Webster.com dictionary.* Retrieved April 12, 2021, from https://www.merriam-webster.com/dictionary/slur

Millwood-Hargrave, A. (2000). Delete expletives?: Research Undertaken Jointly by the Advertising Standards Authority, British Broadcasting Corporation, Broadcasting Standards Commission and the Independent Television Commission. http://ligali.org/pdf/ASA_Delete_Expletives_Dec_2000.pdf

Nunberg, G. (2018). The social life of slurs. In D. Fogal, D. Harris, & M. Moss (Eds.), *New work on speech acts* (pp. 237–295). Oxford: Oxford University Press. DOI:10.1093/oso/9780198738831.003.0010

Office for National Statistics. (2020). Baby names for boys in England and Wales. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/babynamesenglandandwalesbabynamesstatisticsboys (2021-04-05)

Office for National Statistics. (2020). Baby names for girls in England and Wales. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/datasets/babynamesenglandandwalesbabynamesstatisticsgirls (2021-04-05)

Oxford English Dictionary. (2021). Derogatory, adj. and n. In *OED Online*. Oxford University Press. Retrieved July 29, 2021, from https://www-oed-com.ludwig.lub.lu.se/view/Entry/50663?redirectedFrom=derogatory

Panzeri, F. & Carrus, S. (2016). Slurs and negation. *Phenomenology and Mind, 11*, 170-180. DOI: 10.13128/Phe_Mi-20117

Potts, C. (2007). The expressive dimension. *Theoretical Linguistics, 33*(2), 165–198. DOI 10.1515/TL.2007.011

Schleef, E. (2014). Written surveys and questionnaires in sociolinguistics. In J. Holmes & K. Hazen (eds.), *Research methods in sociolinguistics: A practical guide*, 42-57. Chichester: Wiley Blackwell

Schlenker, P. (2003). A Plea for Monsters. *Linguistics and Philosophy: An International Journal, 26*(1), 29–120. DOI: 10.1023/A:1022225203544

Schlenker, P. (2007). Expressive presuppositions. *Theoretical Linguistics, 33*(2), 237–245.

Tenchini, M. P. & Frigerio, A. (2016). A multi-act perspective on slurs. In R. Finkbeiner, J. Meibauer, & H. Wiese (Eds.), *Pejoration* (pp. 167-185). Amsterdam: John Benjamins Publishing Company. DOI 10.1075/la.228.o8ten

Tenchini, M. P. & Frigerio, A. (2020). The impoliteness of slurs and other pejoratives in reported speech. *Corpus Pragmatics, 4*, 273-291. https://doi.org/10.1007/s41701-019-00073-w

Vallée, R. (2014). Slurring and common knowledge of ordinary language. *Journal of Pragmatics, 61*, 78-90. http://dx.doi.org/10.1016/j.pragma.2013.11.013

Weber, R. (1990). *Basic content analysis*. (2nd ed.). SAGE Publications, Inc. https://www-doi-org.ludwig.lub.lu.se/10.4135/9781412983488

Whiting, D. (2013). It's not what you said, it's the way you said it: Slurs and conventional implicatures. *Analytic Philosophy, 54*(3), 364-377. https://search-ebscohost-com.ludwig.lub.lu.se/login.aspx?direct=true&db=phl&AN=PHL2214283&site=eds-live&scope=site

Appendix A: Pilot survey results

Table 15. The utterances used in the Pilot Survey and the mean value of their responses

Utterances	**Mean**
Andy: According to Chris, my brother's a faggot.	4.2
Ben: My dad called his friend an asshole the other day.	1.4
Benjamin: Yeah, and Danny's a nigger.	6.9
Bradley: My dad told me that Alex is gay.	1.2
Calum: I mean he's known as a bastard.	1.8
Daisy: This kid called Claire a bastard.	1.5
Edward: This guy was clearly an asshole.	1.5
Erin: That motherfucker Billy doesn't have a clue.	2.4
Ethan: Last time I spoke to Rhiannon, she called Elliott a chink.	3.3
Gemma: Megan said the new guy is Pakistani.	1.4
Heather: He's a Pakistani and he's not looking after his own people.	3.2
Imogen: Yeah, my boyfriend told me this guy's a motherfucker.	1.7
Isobel: I think I might just find out what a cunt Vic is.	3.2
Jodie: Maria said her boyfriend's a retard and that's why he didn't show up.	3.4
Katherine: She's a faggot and so is Aaron.	6.2
Kieran: She said that Anna's black.	1.3
Kyle: John called Simon a chink the other day.	3.4
Laura: My sister told me that Chloe's a retard.	4.2
Leah: The presenter called some guy a nigger on national TV.	4.7
Lewis: They don't like him because he's black.	3.1
Lily: I didn't know he was Chinese.	1.2
Melissa: he's called my mate a Paki in Manchester on a night out.	2.8
Mohammed: Because Joe's disabled, isn't he?	1.9
Natalie: And Mike just called Syed a Paki.	2.9
Nicola: Last week Zoe told me that Robyn's a cunt.	2.4
Oliver: Yeah, Zachary's the one that I kept saying he's gay.	2.9
Reece: According to Lee, our new teacher is a faggot.	5.1
Rory: I'm saying Adam's just a dickhead	1.7
Ryan: Was she a chink, then?	6.4
Scott: Jamie said my cousin was a dickhead	1.6
Sean: But Paul said that Thomas is a nigger.	6.3
Shane: He's a Paki though so that makes sense.	5.7
Stacey: Yesterday Jay said that Owen's disabled.	2
Stephanie: I heard Cam say that his girlfriend is Chinese.	1
Yasmin: Cos like Liam's a retard and he booked it for four PM to pick us up instead of four AM.	4.8

Open-ended questions:

1. If you would like to explain or elaborate on any of your responses so far, please do so here:
2. Do you think there are words that are always offensive?
3. Was there anything in this survey that you found to be confusing or unclear?

Table 16. The categories identified in the 13 responses to the open-ended questions in the Pilot Survey and how many responses fall into each category (primarily or secondarily)

Content categories	**Primary**	**Secondary**
Context	3	2
Yes, there are words that are always offensive	5	
No, there are no words that are always offensive	2	
In-group usage		2
Survey feedback	1	2
Reported slurs are non-offensive	1	

Appendix B: Survey A stimuli and mean values

Table 17. The words used in Survey A and the mean value of their responses

Words	**Mean**
Adult	1
Asshole	3.71
Bastard	3.39
Black	1.47
Bollocks	2.1
Brunette	1.03
Chinese	1.19
Chink	6.28
Cunt	5.07
Dickhead	3.37
Disabled	1.76
Doctor	1.01
Faggot	5.88
Fuck	3.4
Gay	1.78
Gook	5.34
Jesus Christ	1.66
Jew	2.22
Kike	5.48
Korean	1.11
Motherfucker	4.3
Nigger	6.68
Paki	4.46
Pakistani	1.17
Retard	5.36
Shag	2.63
Shit	3.04
Student	1.05
Tenant	1.09

Appendix C: Survey B stimuli and mean values

Table 18. The utterances used in Survey B and the mean value of their responses

Words	**Mean**
Andy: According to my brother, Chris is a faggot.	5.22
Ben: Earlier today, my dad mentioned that his boss is an asshole.	2.03
Benjamin: Yeah, and Danny's a nigger.	6.43
Bradley: My dad told me that Alex is gay.	1.54
Calum: I mean he's known as a bastard.	2.45
Chantelle: That's because Ed's a retard.	5.55
Daisy: This kid called Claire a bastard.	1.96
Edward: This guy is clearly an asshole.	2.3
Erin: That motherfucker Billy doesn't have a clue.	3.35
Ethan: Last time I spoke to Rhiannon, she called Elliott a chink.	4.51
Gemma: Megan said the new guy is Pakistani.	1.37
Heather: That's interesting because he's a Pakistani as far as I know.	1.74
Imogen: Yeah, my boyfriend insisted that this guy's a motherfucker.	2.64
Isobel: I think I might just find out what a cunt Vic is.	3.79
Jess: I'm talking about Luke, who's a nigger.	6.55
Jodie: Last night, I heard Maria say her boyfriend's a retard.	3.63
Jordan: I met a guy the other day, he's a Paki.	5.21
Katherine: She's a faggot and so is Aaron.	6.12
Kieran: She said that Anna's black.	1.41
Kyle: John called Simon a chink the other day.	4.14
Laura: My sister said that Chloe's a retard.	4.47
Leah: The presenter called that guy a nigger on national TV.	4.58
Lewis: Josh is black, which has nothing to do with this.	1.58
Lily: I didn't know he was Chinese.	1.34
Marcus: My friend Naomi is a chink.	5.9
Matt: Jesus Christ you're annoying.	1.71
Melissa: he's called my mate a Paki in Manchester on a night out.	3.54
Mohammed: Because Joe's disabled, isn't he?	1.77
Natalie: And Mike just called Syed a Paki.	3.52
Nicola: Last week Zoe told me that Robyn's a cunt.	3.21
Oliver: Yeah, Zachary's the one that I kept saying he's gay.	3.2
Reece: Lee mentioned that our new teacher's a faggot.	5.63
Rory: I'm saying Adam's just a dickhead	2.35
Ryan: Was she a chink, then?	6.12
Sarah: Did you hear that Harry's a faggot?	6.14
Scott: Jamie said my mate was a dickhead	2.14
Sean: But Paul said that Thomas is a nigger.	5.79
Shane: I didn't know he was a Paki before I met him	5.57
Stacey: Yesterday Jay told me that Owen's disabled.	1.77
Stephanie: I heard Cam say that his girlfriend is Chinese.	1.25

Yasmin: Cos like Liam's a retard and he booked it for four PM to pick us up instead of four AM.	5.12

Appendix D: Content categories

Survey A

Content categories identified in responses to the main open-ended questions in Survey A:

- *No, no thoughts about the survey*
- *No, nothing was unclear*
- *Positive feedback on survey*
- *Negative feedback on survey*
- *Offensiveness as a concept*
- *Lack of context*
- *Importance of context*
- *Cultural or regional differences*

Survey B

Content categories identified in responses to the open-ended questions in Survey B:

- *Non-reported vs. reported speech*
- *In-group usage*
- *Cultural or regional differences*
- *Context*
- *Lack of context*
- *No, nothing was unclear*
- *Yes, there are words that are always offensive*
- *No, there are no words that are always offensive*
- *Yes, it is offensive to use slurs when reporting*
- *No, it is not offensive to use slurs when reporting*
- *Reporting slurs is less offensive than using them*
- *Degrees of offensiveness*