Social Identity Enactment Through Linguistic Style: Using Naturally Occurring Online Data to Study Behavioural Prototypicality

Submitted by Alicia Cork, to the University of Exeter

as a thesis for the degree of

Doctor of Philosophy in Psychology

in March 2021

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signature:…………………………………………………………………

Acknowledgements

First and foremost, I would like to thank my three supervisors, Dr Miriam Koschate-Reis, Professor Richard Everson and Professor Mark Levine, without whom this research would not have been possible. I wish to express my sincerest gratitude to Miriam, for her unwavering support, immense knowledge and exacting attention to detail every step of the way. Without her, I can guarantee that I would still be stuck in some rogue research rabbit hole; thank you for always trying to keep me on the right track. Secondly, a huge thank you to Richard for patiently guiding me through the more computational aspects of this research project and igniting my interest in data science and artificial intelligence. And finally, a sincere thanks to Mark, for his constructive suggestions and contagiously calm approach to research.

I would like to also thank the Alan Turing Institute for the opportunity to partake in numerous multi-disciplinary projects which have truly enriched my research capabilities. Further, a great deal of thanks goes to the PsyID lab group in Exeter for their continued support and constructive suggestions throughout the last three and a half years. I would also like to give thanks to the Engineering and Physical Sciences Research Council for generously funding this research.

Finally, on a personal note, I would like to give a huge thank you to all my friends and family who have supported me throughout this PhD journey. Specifically, I would like to thank my best friend and partner, Alex Brown, for always cheering me up when things got tough and for equipping me with endless cups of tea and coffee. Writing a thesis during a national lockdown has been exceptionally arduous at times, and I am not sure that I would have been able to complete this thesis without Alex's unyielding support. Thank you.

Abstract

Social identity prototypes refer to the quintessential representation of a particular social identity; prototypes define and prescribe the characteristics, behaviours and attitudes of a particular group, as distinguished from other groups (Hogg, 2001). For the most part, identity prototypicality is studied using self-reported methods used to assess perceptions of the prototypicality of self and others. However, in this thesis we provide behavioural evidence to demonstrate how linguistic style data can be used to measure identity-prototypical behaviour in real world contexts. Combining naturally-occurring online data with experimental data, the first chapter demonstrates that individuals behave in an identity-prototypical way regardless of the context in which they are communicating. Further, we show that this identity-prototypical style of communication is robust to topic, demographics, personality and platform, and moreover that the same identity-prototypical communication style can be detected in experimentally controlled conditions. In the second chapter, we demonstrate the small but statistically significant link between identity-prototypical communication and influence in real-world forum data. This finding provides insight into how group members respond to other ingroup members based on their prototypical communication style in real-world situations. Finally, in the third chapter, we use the group prototypical behaviour observed in naturally occurring online forum data to construct a typology of social identities, demonstrating the existence of five different types of social identity in line with the research of Deaux et al. (1995). We also demonstrate that it is possible to use this measurement of behavioural prototypicality to observe identity change over time. Using eight years' worth of forum data, we illustrate the slow movement of the transgender identity from being a stigmatised identity in 2012, to shifting towards a collective action identity in 2019. In sum, the

findings outlined in this thesis provide evidence to support the idea that it is possible to use machine learning algorithms and naturally occurring online data to study behavioural prototypicality in real world environments. Moreover, this methodology enables us to study identities 'in the wild' thus transcending the limitations associated with using self-reported methodologies or experimental approaches to study how individuals express and enact their group memberships. Further, we also demonstrate the value in using naturally-occurring online behavioural data to test and extend the key components of social identity theory.

Table of Contents

List of Tables

List of Figures

Author's Declaration

I declare that parts of the Literature Review and Studies 1a, 1b and 2a of this thesis have been published in the Group Processes and Intergroup Relations Journal and have been reproduced here with the permission of the copyright holder. My contribution to this multi-authored paper included developing the main idea, undertaking data collection, completing the data analysis, and writing the paper. My co-authors contributed to the development of the main idea and also provided comments on drafts of the paper.

Reference

Cork, A., Everson, R., Levine, M., & Koschate, M. (2020). Using computational techniques to study social influence online. *Group Processes & Intergroup Relations, 23*(6), 808–826. https://doi.org/10.1177/1368430220937354

1   Literature Review

Self-categorisation theory is a prominent social psychological theory which seeks to understand how behaviours and cognitions are impacted by psychological group membership. Specifically, self-categorisation theory suggests that individuals who share a common social identity develop shared sets of behavioural and attitudinal norms that guide their perceptions and behaviour (Turner et al., 1987; Turner, 1991). Whilst the focus of self-categorisation research has revolved around understanding how the salience of a shared social identity impacts cognition and perceptions, in this thesis we explore directly how shared social identities impact behaviour in the real world. Moreover, we also reverse this relationship by exploring what we can learn about social identities through observing and studying the uniformities in group behaviour.

1.1   Overview of Self-Categorisation Theory

Social identities are the aspects of self that are rooted in psychological group memberships and thus form the social-cognitive basis of group behaviour (Turner et al., 1987). The fundamental aim of self-categorisation theory is to explain how a collection of individuals comes to perceive and define themselves as a united group and to explore the impact of salient psychological group memberships on behaviour and perception (Turner et al., 1987). A psychological group, also known as a social identity, is a group that is psychologically important for the members and constitutes more than just membership in a collective. Thus, whilst an individual may be part of many 'groups', it is the psychological connection and recognition of oneself as a group member that influences cognition and behaviour. In this way, individuals relate themselves to the group through processes of social comparison and in turn derive

their values, norms and attitudes from their psychological group membership. This social comparison may be both intergroup, where individuals compare and differentiate their ingroup with reference to other outgroups, as well as intragroup, where individuals compare themselves to other ingroup members (Turner et al., 1987).

Individuals often have many different social identities, such as political or vocational identities, however, these identities only exert influence over the individual's behaviour and cognitions when the identity is rendered salient (Haslam, 2004; Oakes, 1987; Turner & Haslam, 1991). The salience of a particular social identity is said to depend on the cognitive accessibility of it (Gurin & Markus, 1988) and the fit of the identity in the given context (Oakes, 1987; Oakes & Turner, 1990; Turner, 1985, 1991). In relation to accessibility, specific categorisations may be more cognitively accessible than others because they are frequently used (chronic accessibility), or alternatively they may be particularly relevant in a given context (situational accessibility). With regards to the fit of a categorisation, categorisations may 'fit' a situation to the degree that the categorisation is able to account for within-group similarities and between-group differences (comparative fit), and whether these differences can be used to make sense of behaviour (normative fit) (Oakes, 1987; Turner & Oakes, 1990). For example, an English-Scottish categorisation would fit a situation whereby the English and Scottish individuals behaved more like their own ingroup members than the outgroup members, and this behaviour was in line with their group stereotypic expectations (Oakes, 1987). When a specific social identity becomes salient then, the individual is theorised to behave, think and perceive others in ways that are normative to that salient social identity (Hogg, 2007; Turner et al., 1987).

More specifically, when a social identity becomes salient, self-perception and behaviour become ingroup normative as individuals adopt the norms, behaviours, and beliefs of the group (Turner et al., 1987). In addition to this, the perception of outgroup members becomes outgroup stereotypical such that outgroup members are evaluated and judged in line with outgroup stereotypes, as opposed to as unique individuals (Abrams et al., 1990; Haslam & Turner, 1992). Concordantly, when a social identity becomes salient, individuals are hypothesized to behave in ways that are congruent with their group prototype – the group position which maximises the ratio of intragroup similarity and outgroup difference (Turner, 1991; Turner et al., 1987). In this way then, the psychological group membership exists within the individual and constitutes more than just the individual operating within an external collective (Tajfel, 1981; Turner, 1981; Turner et al., 1987). Additionally, prototypes are said to be context-dependent rather than fixed – they vary as a function of the comparative context (Doosje et al., 1998; Haslam et al., 1992, 1999; Hogg & Reid, 2006; Turner et al., 1987). As a result of this variability, the most prototypical position of a group will change dependent on the salience of a particular outgroup (Haslam et al., 1995); prototypes thus operate by maximising between-group differences whilst simultaneously minimising within-category differences (Abrams et al., 1990; Oakes, 1987; Turner et al., 1987; Turner, 1991). For example, the position of the most prototypical vegetarian would be different dependent on whether the outgroup of comparison was meat-eaters or vegans.

In the following section of the literature review, I will explore the existing literature on identity prototypes (Hogg, 2001) and identity prototypicality (Turner et al., 1987). I will point to the subtle disparities between the hypothesized role of the identity prototype as posited by Hogg (2000; Hogg & Reid, 2006; Hogg & Rinella,

2018), and the notion of identity prototypicality as originally suggested by Turner (Turner, 1991; Turner et al., 1987). I will make the argument that where Turner (1991) suggests that a group's prototype is synonymous with the most prototypical position and is thus entirely locally derived, Hogg instead argues that prototypes are the cognitive representation of group norms, where the group norms captured by the prototype represent meaningful context-dependent similarities within and differences between groups (Hogg & Reid, 2006). This subtle difference is important for how we conceptualise and understand uniformities in social group behaviour.

Further, I will then go on to explore how prototypes and prototypicality have been operationalised within the current literature, making the argument that whilst a plethora of studies have investigated perceptions of prototypicality, few studies have directly observed how salient social identities affect prototypical behaviour in real world contexts. Given the importance of behaviour in self-categorisation hypotheses (Turner, 1991), the need to directly test this aspect of the theory appears long overdue.

1.2 Identity Prototypes

In the original conceptualisation of self-categorisation theory (Turner et al., 1987), Turner states that 'the prototypicality of any ingroup member is the degree to which he or she exemplifies or is representative of some stereotypical attribute of the group as a whole – also defined and operationalised by means of the meta-contrast principle' (p. 79). More specifically, the prototypical position is determined through dividing the mean differences between an ingroup member and outgroup members, by the mean differences between the ingroup member and other ingroup members on the relevant stereotypical dimensions of comparison. Thus, the most prototypical position of the group depends on the salient outgroup with which one compares

one's ingroup, and moreover, the prototypical position emerges directly from the relationships between the groups. Turner states that the relative prototypicality of an individual and the stereotypical dimensions of comparison are 'relative and situation-specific, not absolute, static or constant' (p.80). In his later writing, Turner goes on to equate 'the prototype' with the most prototypical position, stating that 'the prototype is *the position* that best defines what the group has in common in contrast to other groups' (Turner, 1991, p. 77). Therefore, the prototype or prototypical position emerges directly from the local context; it mathematically encapsulates the position which maximises differences between groups and similarities within the group. Prototypes do not exist outside of an explicit local context and are thus entirely context-dependent.

In contrast to prototypes being seen as the specific position which maximises differences between groups and similarities within the group, Turner suggests that (self-)stereotypes comprise the attributes and characteristics which make up the cognitive representation of a group (Turner et al., 1987). Therefore, when a social identity is rendered salient and individuals become depersonalised, this leads to "'self-stereotyping' whereby people come to perceive themselves more as the interchangeable exemplars of a social category" (Turner et al., 1987, p. 50). From Turner's perspective then, the self-stereotype is the name given to the cognitive representation of the group which comprises group normative attributes and features.

Interestingly however, this notion of self-stereotypes as comprising the normative attributes of a group appears to show a great deal of overlap with Hogg's definition of a prototype. According to Hogg and Rinella (2018), 'people cognitively represent groups as prototypes — fuzzy sets of attributes (e.g., thoughts, feelings,

behaviours) that simultaneously embody intragroup similarities and intergroup differences, thus capturing the essence of a group as a distinct and internally coherent entity. Group prototypes define a group's identity and, in the case of ingroup prototypes, prescribe how members should behave.' (p.7). Comparatively, this definition of a prototype is less context-dependent than Turner's original definition (1991). However, both Turner (1991) and Hogg and Rinella (2018) suggest that a cognitive representation impacts group perception and behaviour through depersonalisation and self-stereotyping, although where Hogg refers to this as the prototype (Hogg, 2001; Hogg & Reid, 2006; Hogg & Rinella, 2018), Turner refers to this representation as the self-stereotype (Turner et al., 1987; Turner, 1991).

As expected, the differentiation between prototypes as locally derived (Turner, 1991) and prototypes as pre-existing cognitive representations (Hogg, 2001) has implications for how group prototypes and prototypicality is studied. Specifically, in the next section, we explore how prototypes are conceptualised and measured both at the local level in line with the definition provided by Turner (1991), as well as at a more cognitive, global level as suggested by Hogg (2001). By teasing apart this distinction between the two approaches, we can see how self-reported trait studies tend to focus on a more global and context-independent idea of the prototype (e.g., Hogg & Hains, 1996), whereas behavioural studies such as those within the polarisation literature tend to focus on a more localised idea of the prototype (e.g., Mackie, 1986). In turn, there appears to be a gap in the literature for studying more 'global' prototypes using behavioural measures. Before we address the empirical measures of prototypes and prototypicality however, we will briefly explore why prototypes and judgments of prototypicality are considered important within social identity research.

1.2.1   Prototypicality Gradient

As outlined above, some group members may be perceived as more prototypical than others to the extent that they embody the stereotypical attributes which maximise intraclass similarities and intergroup differences (Turner et al., 1987). This is referred to as the 'prototypicality gradient' (Hogg, 2001; Hogg & Terry, 2000; Rosch, 1978; Turner, 1985). Importantly, it has been found that group members are highly attentive to the group prototype and are thus strongly motivated to learn which behaviours, attitudes and cognitions are group normative (Hogg, 2005; Hogg & Abrams, 1993; Hogg & Terry, 2000). Further, being perceived by other group members as prototypical is of great value to individuals (Gómez et al., 2014). Research into members perceived as non-prototypical (also known as marginal members) has shown that due to the enhanced motivation to be seen as prototypical, non-prototypical members will go to great lengths to prove their loyalty to the group often through extreme measures (Goldman & Hogg, 2016; Jetten et al., 2003; Jetten, Branscombe, et al., 2002; Noel et al., 1995; Simon & Stürmer, 2003). Members are more likely to go to extremes for their group when they identify highly with the group or when the group is an important or central aspect of their sense of self (Jetten et al., 2003; Noel et al., 1995). When the identity is less important to the individual, and the identity boundaries are more permeable, marginal members may instead disengage (Ellemers & van Laar, 2010).

Being perceived as a prototypical group member is advantageous due to the power and influence that is conferred upon prototypical members by other group members. For example, prototypical members are more able to influence attitude change than peripheral members (Kameda et al., 1997; Knippenberg & Wilke, 1992; Reicher & Hopkins, 1996; Turner, 1991; Wood, 2000) and have greater sway over

defining the group's norms (van Knippenberg, 2011). As a result, group members who are perceived as prototypical are more likely to emerge as leaders (Fielding & Hogg, 1997). Further, as leaders, those who are perceived as more prototypical are given greater leeway to make mistakes than others (Giessner et al., 2009), due to the fact that greater trust is placed in their abilities and motivations (Giessner & van Knippenberg, 2008). They are also perceived as fairer (De Cremer et al., 2010) and more charismatic (Platow et al., 2006; Steffens et al., 2014) and thus receive greater endorsement from others (Ullrich et al., 2009). This is because group prototypical leaders derive their power from the implicit perception that they represent what is group normative (Hogg & Reid, 2006). Consequently, followers are more open to being influenced by group prototypical leaders as they implicitly believe that prototypical leaders are motivated to pursue the group's best interests in a way that non-prototypical leaders are not (Giessner & van Knippenberg, 2008; van Knippenberg & van Knippenberg, 2005). In this way, prototypical leaders are believed to be 'one of us' and their actions are seen as 'doing it for us' (Steffens et al., 2014). Thus, prototypicality and judgments of prototypicality play a key role in understanding various social influence processes central to self-categorisation theory. In the next section, we will therefore explore how prototypes, self-stereotypes and prototypicality have been empirically measured within the literature.

1.3 Studying Group Prototypes, Self-stereotypes and Prototypicality

In this section, I will compare the approaches taken to study localised prototypicality (in line with Turner et al., 1987 and Turner, 1991) with the approaches taken to study globalised prototypicality (in line with Hogg, 2001). Notably, the distinction between the two relies on the dependence of context; where Turner

suggests that prototypes arise directly within the context and are thus inherently dynamic (Turner, 1991), Hogg (2001) suggests that prototypes are instead a cognitive representation which are therefore relatively independent of context (I say relatively, as there is evidence that the cognitive representation of the group does change when a particular outgroup is made salient, Haslam et al., 1995). Therefore, in this section I will first focus on the literature which studies a more globalised cognitive representation of the group. As noted previously, the context-independent ingroup representation may also be referred to as the self-stereotype (Turner, 1991).

Consequently, this section will first look at the self-stereotyping literature which explores group members behavioural or attitudinal assimilation to the stereotype of the group. It will then go on to explore other approaches to the study of the ingroup representation which tend to rely on participants describing the traits of their identities. Next, I will outline the more localised approaches, exploring the polarisation literature and the formation of locally prototypical positions as developed through interaction. This literature is often behavioural in nature and uses experimental data. In turn, I will argue that whilst local prototypes tend to be studied in a behavioural manner as they arise out of local interactions, there are fewer studies which consider the influence of the ingroup representation from a behavioural perspective. Whilst some self-stereotyping research does focus on behavioural assimilation in experimental conditions, we note the absence of research using naturally occurring behavioural data to understand how group identities impact behaviour outside of laboratory conditions. This thesis aims to address this gap in the literature.

1.3.1 Global Prototypicality

1.3.1.1 Self-stereotyping

Self-stereotyping seeks to explain behavioural or attitudinal assimilation within groups (Oakes et al., 1991; Pehrson & Reicher, 2012; Spears et al., 1997; van Veelen et al., 2016). Self-stereotyping is said to occur 'on all and any dimensions which are believed to be correlated with a categorisation' (Hogg & Abrams, 1988, p.74), and thus may include both positive and negative attributes. For example, research into gender stereotypes has demonstrated that individuals attribute both positive and negative stereotypes to themselves (Latrofa et al., 2010). Conversely, Biernat et al. (1996) noted that fraternity and sorority members selectively self-stereotype with regards to the positive stereotypes associated with their group, but not the negative stereotypes. Here then, we see that self-stereotyping in terms of the traits one reports as being stereotypical of one's group may indeed be impacted by the valence of the attributes. Selective self-stereotyping may be explained by a desire to keep a positive image of one's group (Rubin & Hewstone, 1998). In turn then, self-stereotyping refers to the process of attributing stereotypical group attributes to oneself, although the valence of those attributes may impact the attribution process.

The process of self-stereotyping is often studied through deriving a list of stereotypical attributes about an identity from laypeople, and then observing how participants behave in line with these socially stereotypical attributes (e.g., Haslam et al., 1992; Latrofa et al., 2010; van Veelen et al., 2016). In this way, self-stereotypes comprise both ingroup and possible outgroup representations of the group. For example, in their pilot study, Latrofa et al. (2010) developed lists of attributes deemed stereotypically female (e.g., affectionate) or non-stereotypically female (e.g.,

powerful) through asking laypeople what they think society thinks of women in general. In their main study, they then assessed how female participants rated these attributes for women generally, and then for themselves. This notion of using a pilot study to determine societally stereotypical attributes of an identity is common throughout the literature (e.g., Haslam et al., 1992; van Veelen et al., 2016). Consequently, these studies can help us to understand how individuals may assimilate to a general or societally held stereotype of their own group (i.e., a stereotype held by laypeople).

Notably, the idea that individuals behaviourally or attitudinally assimilate to a general stereotype of the group is distinct from research which focuses directly on how an ingroup may respond or behave based on what they believe a specified outgroup thinks of their ingroup. This difference lies in whether the ingroup representation comes from a specific outgroup, or whether it is a more generalised layperson perspective of a group. Whilst the latter (a generalised layperson perspective) is often the focus of self-stereotyping research, the former (a specific outgroup perspective) is the focus of meta-stereotyping research. Thus, meta-stereotyping research focuses on how group A perceives group B perceives group A. For example, research from Peters et al. (2019) has noted how older workers perceptions on how they are viewed by younger workers impacts their perceived employability. Specifically, older workers who thought that the younger workers held negative stereotypes about them reported feeling less employable than those who did not believe that younger people held negative stereotypes. Here we can see how an ingroup member's belief about an outgroup members' stereotype of their group comes to impact the ingroup members' behaviours and cognitions. This demonstrates the importance of understanding from whose perspective a stereotype

is from. Whilst self-stereotypes are said to comprise 'any dimensions which are believed to be correlated with a categorisation' (Hogg & Abrams, 1988, p.74), prototypes may be seen to comprise only the ingroup's representation of their own group (Hogg & Reid, 2006). Of course, ingroup members perspectives of their own identity may be influenced by possible outgroup perspectives, but the definition of a prototype as per Hogg (2000, 2001) is more explicit in being defined from the perspective of ingroup members. For this reason, one way to assess the content of the group prototype is to ask group members what traits or attributes they think are characteristic of their group. We will explore this literature next.

1.3.1.2 Self-Report Trait Studies

Prototypes and self-stereotypes are also studied using basic self-report methodologies wherein participants are asked to describe their identity. For example, Haslam et al. (1995) demonstrated that when Australians describe their own identity in the absence of a comparative context, they use traits such as happy-go-lucky, sportsmanlike and straightforward to describe their identity. However, when asked to rate themselves on traits in relation to an American outgroup, the Australian participants now report being more sportsmanlike, but less happy-go-lucky and now also pleasure-loving. Consequently, Haslam et al. (1995) suggested that the comparative context (in this case, the introduction of the American outgroup) changed the defining attributes of the group self-stereotype. The participants thus responded in line with the predictions of the metacontrast principle; they highlighted the features of their identity that maximally distinguished them from the outgroup, whilst simultaneously maximising within-category similarities. This comparative context led to a new attribute – pleasure-loving – becoming important to the Australian self-stereotype, as it allowed them to distinguish their Australian identity

from the American identity. These results thus support the idea that not only do groups have an idea of their global ingroup representation outside of a comparative context, but that certain aspects of the ingroup representation are emphasised when a particular outgroup is made salient. In this way then, the ingroup representation exists both independently of context, but may at the same time be impacted by changes in the local context.

In other self-reported trait research, participants have been asked to report the behaviours and traits of their identity which differentiated them from other groups. Hogg and Hains (1996) gave participants a definition of what they meant by a team's 'identity' before asking participants to describe their own team's identity. Hogg and Hains's research focused on the identity of netball teams and thus their definition of a netball team's identity was as follows: 'Your team's "identity" is a cluster of things about your team which describes how your team operates and how your team is different from other teams. Think about your team in terms of how you play netball, how your team behaves on and off the court, and particular things which are unique to your team and which serve to differentiate your team from other teams' (p. 298). Using this methodology, Hogg and Hains (1996) found that each of their 11 netball teams reported similar prototypes, albeit with differing levels of emphasis on the social and task functions of the team. Moreover, they found that in nine out of the 11 teams, individuals showed strong agreement on what their team's prototype was. This finding of consensus demonstrates the existence of a shared cognitive representation of the group which exists as a result of the shared social reality of the ingroup members. Interestingly however, despite the clear focus in Hogg and Hains's (1996) study on perceived behavioural differences between an ingroup and possible outgroups, it is notable that no research has directly assessed the behaviour of

groups in order to understand the differences between different groups' global prototypes.

Conversely, research into more localised group prototypes has often used behavioural methods to understand how prototypes form as a result of direct interactions with other group members. In this research, prototypes are considered as the most prototypical position in line with Turner's original conceptualisation (1991). In the following section, I will explore this more localised understanding of the group prototype through an overview of the polarisation literature and group norms research.

1.3.2 Localised Prototypes and Prototypicality

1.3.2.1   Polarisation Studies

Polarisation refers to the well-established finding that following group discussion, individuals often report a more extreme opinion than the average of the discussants' opinions prior to interaction (Mackie & Cooper, 1984; Turner et al., 1989). However, polarisation only occurs when individuals share a salient social identity; when a shared social identity is salient, ingroup members post-discussion opinions polarise in the direction of the ingroup members pre-discussion positions (Abrams et al., 1990; Hogg et al., 1990; Mackie & Cooper, 1984; Mackie, 1986; McGarty et al., 1992; Moscovici & Zavalloni, 1969; Myers & Lamm, 1976; Turner et al., 1989). Polarisation, therefore, is a convergence upon the perceived in-group prototypical position as opposed to the actual normative position of the individuals in the discussion. Ingroup members converge on this extreme position so as to maximally differentiate themselves from alternative and opposing viewpoints (McGarty et al., 1992; Turner et al., 1989). When individuals perceive other

discussants to be outgroup members however, they fail to see others' perspectives as subjectively valid and do not update their own opinions (Turner, 1991).

The process of polarisation as a result of perceived in- and outgroup prototypical positions provides compelling evidence demonstrating the impact of the comparative context on the perceived prototypical position. Further, Hogg et al. (1990) state that polarisation is a 'conformity phenomenon in which individuals who identify with a group conform, through the process of self-categorisation, to the local norm which best represents the group.' (p. 81). Of note here, is the term 'local norm'; the locality of this prototypically normative position is distinct from a more globally prototypical position. However, polarisation studies may also allow us to understand how global prototypes and local prototypes exist together. For example, where ingroup members consensually agree on the direction of their pro/anti stance prior to communication (often determined in a pilot study; e.g., Postmes et al., 2005), this can be said to represent the identity's generalised globally prototypical position. This pro or anti stance on an issue may be deduced from an understanding about 'who we are' and 'what we think'. In this way, this starting position exists in the absence of an explicit local context. By comparison, the locally determined prototypical position is context-dependent; it depends on the explicit position of other group members within discussion. Here then, we can identify the role of both the more context-independent ingroup representation, as compared to a context-dependent locally derived prototypical position.

Outside of experimental studies of polarisation, some online research has focused on understanding how prototypes and prototypical positions arise out of naturally occurring direct interactions between group members. Below, we explore

the different ways that researchers have utilised naturally occurring behaviour to study and conceptualise prototypicality within specified contexts.

1.3.2.2 Naturally Occurring Behavioural Prototypicality

In their research paper, Postmes et al. (2000) were interested in understanding the natural formation of group identities and group prototypes which arise through repeated online interactions. They noted that in student groups discussing course materials, group-specific communication norms naturally emerged over time. Using human coders, they noted that each group had a distinctive communication style, such that some groups were found to be more humorous or rude than others. Postmes and colleagues suggested that this was evidence of group prototypical styles forming through repeated interactions which served to distinguish each group from other potential outgroups (Brewer, 1991). Additionally, they found that over time students in each group conformed more to these norms, but only when engaging in intragroup communication. They thus concluded that interactive groups develop their own communication norms through repeated interactions and that over time individuals become more prototypical through using and adhering to the group communication norms. In this way, individuals express their ingroup identity to others through using the linguistic strategies that are indicative of group identity.

Postmes et al.'s (2000) study indicates how communication norms can form over time and how these may be indicative of a sense of shared identity. This idea has been complemented by research using social media data to understand the development of shared meaningful collective action identities. Smith et al. (2015) used linguistic analysis to study markers of the shared social identities that formed during the Occupy Wall Street movement. Using both the content of social media

posts in combination with the style of the posts (i.e. use of pronouns, negations and assent words), they noted that in a Facebook group dedicated to discussing the Occupy Wall Street movement, discussants formed a shared identity when they were able to agree on both the injunctive norm of their identity ('revoking corporate personhood') as well as the necessary action required to achieve this aim ('occupy Wall Street'). Specifically, they noted that ingroups and outgroups formed around different injunctive norms and desired action paths. Smith and colleagues specified that this formation of shared social identity occurred via processes laid out in the identity-norm nexus formation model (INN-formation model; Smith, Thomas, et al., 2015). The INN-formation model posits that new movements such as the Occupy movement are not defined by pre-existing ideologies and categories but are instead negotiated through interpersonal communication. Thus, individuals share a common motivation for social change and collectively decide upon the best way to enact that change through repeated interactions. In this way, the shared identity is created when individuals behaviourally and attitudinally assimilate on what they believe to be the right course of action (injunctive norm).

In both of the aforementioned studies, the authors have used longitudinal analyses to observe the process of behavioural assimilation which occurs as a result of repeated interactions. In these studies, shared social identities and identity prototypical positions emerge from the local group norms (Postmes et al., 2000) or through the conscious negotiation of the content of an identity (Smith et al., 2015a). Similar research from computer science and communication domains has observed how linguistic norms form over time and are used by individuals to signal their ingroup identity (Bagozzi & Dholakia, 2002; Danescu-Niculescu-Mizil et al., 2013; Johnson et al., 2015). For example, Danescu-Niculescu-Mizil et al. (2013) explored

how the adoption of these local group norms can be used to predict the length of a user's stay within a particular online community. Further, Johnson et al. (2015) indicated that those who more frequently employ the communication norms of the group are likely to have a greater influential ability over the virtual community than those whose behaviour is less aligned with the norms.

Here then we can see how direct interactions between group members in real world environments leads to the formation of group prototypical positions. Further, we can see that adherence to these prototypical group norms is related to the amount of influence that group members have (Johnson et al., 2015), in line with the premises of Turner's self-categorisation theory (Turner, 1991). However, when assessing the relationship between a more globalised notion of prototypicality and influence, this research has focused almost entirely on using self-reported judgments of prototypicality as opposed to behavioural prototypicality. This is because global prototypicality is not assessed or measured behaviourally in the same way that local prototypicality is. Below, we will critically evaluate some of the ways that global prototypicality and influence have been measured, pointing to the difficulties in defining and measuring what is meant by the most prototypical person when only self-report measures are used.

1.3.3   Social Identity Theory of Leadership

Within the body of literature often referred to as the 'Social Identity Theory of Leadership' (Hogg, 2001; Hogg et al., 2012), researchers aim to assess the relationship between the prototypicality of an individual, and the influence that they have over other group members. In contrast to Turner's mathematical formulation of the most prototypical position (Turner et al., 1987), this research tends to focus on self-reported judgments of another's prototypicality by asking other group members

to rate their leader's exemplariness as a group member or similarity to other group members (Steffens et al., 2021). In this way, this literature aims to understand how a leader (or other group member) embodies the norms and attributes that are perceived to define the group identity (Hogg et al., 2012).

Notably, much of this prototypicality literature uses measures taken from van Knippenberg and van Knippenberg (2005) or Platow and van Knippenberg (2001), whereby participants are asked to rate how similar a member or leader is to the rest of the group, whether they are a 'good example' of the identity, and whether they represent the characteristics of the group. Conversely, others have argued that similarity alone is not a theoretically sound measure of prototypicality and instead propose that the most prototypical individual must be an 'exemplary' or 'ideal' group member, as opposed to just a 'good example' (Bartel & Wiesenfeld, 2013; Halevy et al., 2011; Hogg et al., 2012; Steffens et al., 2021; Steffens et al., 2014). This notion of similarity not being a sufficient enough measure of prototypicality fits with the idea that prototypicality constitutes more than just within-group likeness (Turner et al., 1987).

In a recent meta-analysis, Steffens and colleagues (2021) sought to examine the effects of both ideal prototypicality and average prototypicality on leadership evaluations. They found that leader prototypicality was more strongly correlated with leader evaluation and behavioural outcomes in studies where prototypicality was conceptualised as the most ideal-type as opposed to the most average group member. Table 1.1 below outlines how they distinguished between studies that categorised prototypicality as exemplariness, and those that defined prototypicality with regards to similarity. It must be noted that Steffens and colleagues observed

robust effects for the impact of leadership prototypicality on leader evaluation and behavioural outcomes, irrespective of the measures used.

Table 1.1

Table taken from Steffens et al., (2021) Distinguishing Between Ideal-type and Average-type Measures of Prototypicality

Scale name (Reference)	*k*	Scale Items
Leader Relative In-group Prototypicality (Platow & van Knippenberg, 2001)	*49*	1. [This leader] represents what is characteristic about [in-group]. 2. [This leader] is representative of [in-group] members. 3. [This leader] is a good example of the kind of people who are members of [in-group]. 4. [This leader] stands for what members of [in-group] have in common. 5. [This leader] is not representative of the kind of people who are members of [in-group]. (reversed) 6. [This leader] is very similar to most people in [in-group].
Leader Group Prototypicality Scale (van Knippenberg & van Knippenberg, 2005)	*36*	1. [This leader] is a good example of the kind of people that are members of [in-group]. 2. [This leader] represents what is characteristic about [in-group].

		3. [This leader] has a lot in common with members of [in-group].
		4. [This leader] is very similar to the members of [in-group].
		5. [This leader] represents what [in-group] stands for.
		6. [This leader] is an embodiment of our [in-group] group norms.
Identity Leadership Inventory–Leader Identity Prototypicality Subscale (Steffens et al., 2014)	*33*	1. [This leader] embodies what it means to be a member of[in-group].
		2. [This leader] is representative of members of [in-group].
		3. [This leader] is model member of [in-group].[a]
		4. [This leader] exemplifies what it means to be a member of[in-group].[a]
Supervisor Organizational Embodiment Scale (Eisenberger et al., 2010)	*25*	1. When [this leader] encourages me, I believe that [in-group] is encouraging me.
		2. When [this leader] is pleased with my work, I feel that[in-group] is pleased.
		3. When [this leader] compliments me, it is the same as[in-group] complimenting me.
		4. When [this leader] pays attention to my efforts, I believe that[in-group] is paying attention to my efforts.

		5. [This leader] is characteristic of [in-group].
		6. [This leader] and [in-group] have a lot in common.
		7. When I am evaluated by [this leader], it is the same as being evaluated by [in-group].
		8. [This leader] is representative of [in-group].
		9. [This leader] is typical of [in-group].
Leader Group Prototypicality Scale created for purpose of study (Janson et al., 2008): Study 2)	*1*	1. In terms of interests and values, [this leader] does not have a lot in common with the team. (reversed)
		2. [This leader] is a great fit for the team.
		3. [This leader] exemplifies what is best about the team.[a]
Leader Group Prototypicality Scale created for purpose of study (León et al., 2009)	*1*	1. [This leader] represents the most characteristic traits of the group.
		2. [This leader] would be elected by the group members to represent the group before others.
		3. [This leader] is well-liked by the group members.
		4. [This leader] personifies the most positive values of the group.[a]
Leader Group Prototypicality Scale created for purpose of study (Sluss, 2006)	*1*	1. [This leader] is an exemplary member of the organization.[a]

	2. [This leader] acts according to the organizational goals and values.[a] 3. [This leader] represents what the organization stands for.

Note. *K* = number of independent samples that used a given scale.

[a] Indicates items assessing ideal-type notion of leader group prototypicality as coded in Steffen et al's (2021) analysis

Despite this robust effect, here we question whether the self-report questions outlined in Table 1.1 are in fact actually capturing different conceptualisations of context-independent (global) prototypicality. For example, Steffens et al. (2021) suggest that when participants are asked to rate whether a leader is a 'good example of the kind of people who are members of the in-group' (Platow & van Knippenberg, 2001; van Knippenberg & van Knippenberg, 2005), they are capturing how similar to the group a leader is, however when asked whether a leader is a 'model member of the ingroup' (Steffens et al., 2014) they are instead capturing the exemplariness of the leader. However, whether this semantic difference is acknowledged by participants remains unstudied; would participants rate the same person differently dependent on whether the question was regarding their position as a 'model member' or as a 'good example'? Herein lies the difficulty in using only self-report scales to measure perceived prototypicality. Whether participants respond differently to questions of similarity and exemplariness is currently unclear.

In addition, research which has sought to compare the perceived prototypicality of exemplary and average group members has been confounded by the choice of

prototypicality questions employed by the researchers. For example, Steffens and colleagues (2013) explored whether group members with exemplary performance were perceived as more prototypical than those with average performance. In their study, they used vignettes which described either an exemplary academic or an average academic. Their exemplary academic was said to have published nine articles as opposed to the two published by the average academic, taught three modules as opposed to only two, supervised two PhD students as opposed to only one, had a recent PhD student pass their viva as opposed to fail it, and won a grant worth £300,000 as opposed to just £10,000. The academics were then 'reviewed' by an independent reviewer who rated both academics performance on a scale from 1 to 7. The independent reviewer made clear that the midpoint of 4 was the average across all academics and consequently gave the average academic scores of 3s and 4s, whereas the exemplary academic achieved scores of 6s and 7s. Participants were then asked to judge how prototypical these academics were using the following items: i) This academic embodies what the group stands for, ii) This academic exemplifies what it means to be a member of the group, iii) This academic epitomises what it means to be a member of the group and iv) This academic is representative of members of the group. Steffens and colleagues (2013) noted that the academic with exemplary performance was perceived to be more prototypical than the average academic.

However, one has to question the degree to which this finding was based on identity prototypicality specifically. Firstly, participants were given four scores from an 'independent reviewer' as to the academics' performance all of which were proposed to define an academic's role (i.e., publishing papers, supervising students etc). However, without the inclusion of other metrics that are equally desirable albeit not

prototypical of the academic identity (for example, having a large LinkedIn network), it is not possible to tell whether participants thought the exemplary academic genuinely 'epitomised what it means to be a member of the group', or whether they simply saw that an individual scoring 7/7 on four performance metrics was more exemplary than an individual scoring 4/7.

In addition to this, Steffens et al. (2013) suggest that their research provides support for the idea that leaders are exemplary group members and not just average members. However, the prototypicality questions used in this research are not directed at measuring similarity to other group members. Despite the fact that the majority of prototypicality research uses measures such as 'This group member is very similar to most people in the group' (Platow & van Knippenberg, 2001), Steffens and colleagues omitted measures of similarity and instead focused on measuring prototypicality using terms such as 'epitomising', 'exemplifying' and 'embodying'. Whilst their fourth question enquired about how 'representative' the individual was of the identity, this question still does not capture similarity explicitly. It therefore appears that they have presented the vignette of an exemplary academic, confirmed that the academic was exemplary using an independent reviewer and then asked participants to judge their exemplariness. Without performance metrics that do not pertain directly to the academic identity, and without the inclusion of measures of perceived similarity within (or possibly even distinct from) the prototypicality measure, this study does not coherently illustrate whether prototypicality is in fact defined by exemplariness or similarity to others.

Based on the above literature then, I suggest that self-report measures may not be completely coherent in measuring prototypicality due to subjectivity of both participants and researchers. In Steffens et al. (2013) research, we observed how

researchers can use different questions to operationalise prototypicality; in this way, the definition of prototypicality is left to the researcher and their choice of questions. Thus, not only is the measurement of prototypicality confounded by participant interpretations of a 'good example' compared to a 'model member', but we can also see how decisions made by researchers regarding what questions to ask can impact the validity of the conclusions drawn. This therefore points to the need for more objective and observable measures of prototypicality.

1.3.3.1 Visualisations of Prototypicality

An alternative approach used to manipulate perceptions of prototypicality is through the use of visualisations displayed to artificial, experimental groups in the laboratory. These studies appear to capture the idea of intergroup distinctiveness and metacontrast theory more coherently (Turner et al., 1987). In these studies, participants are shown a bogus distribution curve of ingroup members' responses to a particular task whereby the tail of the curve overlaps with a distribution curve of outgroup responses to the same task (Platow & van Knippenberg, 2005). Participants are then shown an arrow pointing to either the peak of the curve, a position closer to the outgroup, or a more extreme position away from the outgroup. In some studies, participants are then explicitly told how similar the target is to other ingroup members, and how different they are from outgroup members (Platow & van Knippenberg, 2005, p. 1511). In this way, researchers define the prototypical position at the peak of the curve. Research using this methodology has found that leaders at the mode of the distribution are perceived as fairer and more likely to be endorsed than leaders at either end of the distribution curve (Platow & van Knippenberg, 2005).

Whilst this methodology coherently indicates how the perception of similarity to one's ingroup affects ingroup members' evaluations of leaders, there are two key limitations of this methodology worth mentioning. Firstly, in these studies it is rarely reported whether the distribution curve is skewed and thus it is difficult to assess whether targets are simply occupying the mean position of the group. It is therefore not clear whether it is within-group similarity or between-group differentiation that is truly driving these results. In fact, other research using visual cues to manipulate perceptions of prototypicality places individual's bogus group scores on a scale from 0-90, wherein individuals at both the extremes of the scale are told that they are more typical of their group (Noel et al., 1995; van Kleef et al., 2007). In this way, prototypicality is conceptualised as the 'extreme' end of the scale whereby the individual is furthest from the outgroup and is thus not the same as the mean position. This second methodology appears to be more in line with the principle of meta-contrasts put forward by Turner et al. (1987) as it captures both ingroup similarity as well as differentiation from an outgroup. However, by comparing these two approaches we can see how the visual communication of prototypicality may impact who participants perceive as being prototypical of their identity.

In addition to this, a key limitation of both of these methodologies concerns the ecological validity of giving participants 'objective knowledge' regarding how similar a person is to the rest of the ingroup, as well as a hypothetical outgroup. In real world settings, perfectly presented distribution curves or numerical scales of the one attribute of interest compared neatly with a defined outgroup are often not accessible[1]. Similarly, some studies opt not to use visual aids in their prototypicality

---

[1] An exception to this may be in sports or organisational settings where participants are provided with key information as to their position relative to their teammates/colleagues (e.g., in performance review meetings).

manipulation, but instead provide a written description of an individual. Often, this description is prefaced with 'As a typical [group member]' or mentions that the target is 'very representative of the kinds of persons in the team' as opposed to 'an "outsider" [..], someone who was very different from the other team members and who had a different background, different interests and different attitudes' (Giessner & van Knippenberg, 2008; Pierro et al., 2009). By directly communicating to the participants which individual is representative and which one is an outsider, it is not clear how individuals make prototypicality judgments themselves. In fact, it is possible that simply being told that an individual is representative has a significant impact on the participants' resultant answers, regardless of whether the participant would have actually judged the target to be representative themselves. One's subjective perceptions of one's own (and others') prototypicality plays an important role in influencing behaviours and attitudes (Gomez et al., 2014), and thus the level of certainty regarding a group member's position within the group leads us to question the validity of using objective appraisals of the individual to study prototypicality.

Nonetheless, there are some studies which have provided a description of an individual without explicitly stating the intended level of prototypicality of the individual. For example, Giessner et al. (2009, Study 1) created fictional vignettes of Green party candidates that had been developed in consultation with Green party supporters. After presenting participants with the description of either a prototypical or non-prototypical Green party candidate, participants were asked to rate whether the candidate 'stands for what members of the Green Party have in common' and whether they 'represent what is characteristic about the Green Party' (adapted from van Knippenberg & van Knippenberg, 2001). Giessner and colleagues'

prototypicality manipulation worked, in that the prototypical candidate was rated as being more representative of the Green party than the other candidate. Here, we can see that through developing an understanding of the global group prototype (through consultations with other ingroup members), it was possible to anticipate how other group members would perceive the level of prototypicality of the candidates. In this way, participants did not need to be told that a candidate was similar to other group members as is frequently the case in prototypicality research, instead there was evidence of a shared social reality that influenced their own perceptions of each person's prototypicality. This study provides strong evidence to suggest that a shared understanding of the global group prototype exists, and moreover we can use knowledge of this group prototype to predict other's perceptions of prototypicality.

In sum, the results from Giessner et al. (2009) have coherently demonstrated the link between context-independent prototypes and perceptions of influence. To extend this research however, it would be interesting to assess whether judgments of prototypicality are based on the way in which individuals behave. That is, are individuals who behave more prototypically more influential than those who do not? This is the focus of chapter 3 of this thesis, but before we can address this question it first becomes necessary to develop a validated measure of behavioural content-independent prototypicality. In the next section, we explore why online behavioural data may be well suited to achieving this aim.

1.4 Social Identity Model of Deindividuation Effects

The SIDE model (Reicher et al., 1995) suggests that individuals communicating in online groups may have a pre-existing social identity salient due to the anonymity afforded by computer mediated communication. Here we explore the literature examining how the anonymity afforded by online interactions impacts social

identity salience – both with regards to pre-existing social identities, as well as those which are formed within communication. We thus argue that in online communities with defined pre-existing identities, it is reasonable to assume that individuals communicating in these forums will have a pre-existing identity salient. In turn, this makes online data well suited for pre-existing social identities in a naturally-occurring context.

The SIDE model incorporates two distinct aspects of anonymity that lead to the increased salience of social identities; namely, the anonymity of others to the individual and the anonymity of the individual to others (Reicher et al., 1995). These two pathways can be thought of in terms of i) the salience of a social identity caused by the depersonalisation of others (cognitive component), and ii) the performance and construction of the individual's identity resulting from their own anonymity to others (strategic component) (Sassenberg & Postmes, 2002). Often in research, these two aspects are not disentangled; there are relatively few studies that have explicitly focused on asynchronous anonymity such as situations in which an individual is identifiable whilst their group is anonymous (or vice versa). Nonetheless, I will make explicit those which have teased apart the role of identifiability of self and identifiability of others within this review.

The SIDE model has explored the role of anonymity both in computer mediated communication (CMC) (Lea et al., 2001; Spears & Lea, 1994) as well as in face to face communication (Reicher, 1984). It contends that when a social identity is salient, the anonymity of others leads to their greater depersonalisation and perceived similarity resulting in the enhancement of group related outcomes. In this way, depersonalisation makes people seem less like unique individuals and more like 'interchangeable exemplars of a social category' (Turner et al., 1987, p.52). At

the same time, anonymity of self to others leads individuals to behave more in line with their group norms.

The first evidence to support the impact of anonymity on depersonalisation was found offline using a standard deindividuation procedure. Reicher (1984) found that individuals communicating in face to face groups were more likely to adhere to group normative behaviour when made visually anonymous through the use of masks and overalls. Since then, similar findings have been replicated online; for example, the combination of social identity salience and anonymity has led to heightened group polarisation in CMCs (Lea & Spears, 1991; Lee, 2006, 2007), stronger intergroup differentiation (Postmes et al., 2002), and stronger conformity to group norms (Postmes et al., 2000; Postmes et al., 2001). Similarly, group-based attraction (Lea et al., 2001), cohesion (Lea et al., 2007), trust, (Tanis & Postmes, 2005) and social support (Spears et al., 2002) have all been found to increase under conditions of social identity salience and anonymity in computer mediated interactions. Whilst this body of work demonstrates that social identity salience and anonymity can lead to a large range of group related outcomes, it must be noted that the majority of this work has taken place in laboratory conditions where both anonymity and social identity salience were carefully controlled. This is likely due to the difficulty in measuring group processes (i.e. salience and prototypicality) using naturally occurring data.

Nonetheless, we can use this research to understand how individuals behave when anonymity and identity salience are experimentally manipulated. For example, Lea and Spears (1991) experimentally manipulated the salience of either a personal or social identity by informing individuals that the study was interested in either group communication styles or in individual communication styles. They then went on to

refer to the participants by either their group number or their participant number. Additionally, anonymity was manipulated through using participants sat in the same room or in different rooms where they were visually anonymous. In their research, Lea and Spears (1991) found that local group polarisation was strongest when participants were interacting anonymously and had a social identity salient. When a social identity was salient and members were identifiable, polarisation still occurred in the group normative direction although to a lesser extent than in the anonymous condition. Conversely, when participants were anonymous but had a personal identity salient, Lea and Spears observed depolarisation whereby individuals sought to define their individuality in contrast to the group norm. They therefore concluded that when a social identity is salient, increased anonymity can enhance group-based outcomes such as local group polarisation. The crux of the SIDE model however rests on the salience of the social identity. In this study, the social identity referred to a local group identity inferred from communication with anonymous others.

Further research has explored how pre-existing social identities, such as university department, gender and nationality, impact group processes such as polarisation and group cohesion in computer mediated communications (Lea et al., 2007; Postmes et al., 2002). For example, Lea et al. (2007) used discussion groups of four participants (two British and two Dutch, with a male and female of each nationality) to test hypotheses regarding the psychological impact of visibility on the salience of gender identities. They found that when individuals were visually anonymous, it was possible to manipulate the salience of national identities using discussion topics that were relevant to nationality. Thus, in the discussion about nationality, they found that group cohesiveness was enhanced by anonymity and was moderated by prototypicality and self-categorisation. Notably however, they

found that being visible (as opposed to anonymous) led to greater self-categorisation in the gender identity manipulation. This was argued to be because gender is a visually cued category and therefore depersonalisation is believed to occur through visual fit. Conversely, as European nationality is a non-visually cued category, depersonalisation was enhanced through anonymity. In this way then, we can see how anonymity can lead to greater group cohesion when pre-existing social identities are made salient. This suggests that individuals may be impacted by both local group identities that are created within the interaction (Lea & Spears, 1991) as well as pre-existing categorical level identities (Lea et al., 2007).

Interestingly however, it must be noted that in the majority studies that have attempted to study the hypotheses of the SIDE model in real-world contexts, self-categorisation and social identification have been measured with regards to the immediate local group. Individuals are asked to report on their identification or self-categorisation with a specific forum or Facebook group as opposed to with the wider superordinate categorical identity. For example, studies have shown that the more individuals communicate with others in an online community, the more strongly identified they feel with the local group (Jans et al., 2015). Additionally, Le Henaff et al. (2015) observe how local team-based identification in social gaming impacts group performance. Conversely, few studies have explicitly measured how anonymity and the salience of a pre-existing categorical social identity impact real-world behaviour (cf. Chen & Wu, 2015). Here, we find synthesis with the behavioural research discussed in the prior section regarding the distinction between social identity induction and social identity deduction. Once again, we note the gap in the literature regarding how pre-existing salient social identities may impact behaviour in the real world.

One of the few studies which has focused on a global, pre-existing categorical identity in real-world contexts is Chen and Wu (2015). Chen and Wu studied the effect of anonymity on cheating behaviour in online games and found that cheating behaviour was moderated by identification with a generalised gaming identity. In this study, participants were asked to report how often they played anonymously with groups of gamers, how identified they felt with the gaming community and also asked to report on their own cheating behaviour. Chen and Wu noted that individuals who reported playing anonymously more often, were also more likely to report greater cheating behaviour. However, the relationship between anonymity and cheating was moderated by identification with other gamers, such that those who identified with the gaming identity were less likely to engage in cheating behaviour regardless of their level of anonymity. Here, we can see how anonymous individuals with greater identification to a global group identity engage in more group-serving behaviour regardless of their identifiability. On the other hand, anonymity and low identification results in self-serving behaviour, in line with the hypotheses of the SIDE model. Whilst this study does not directly measure the salience of a social or personal identity during gaming itself, it is assumed that individuals with greater identification to the gaming identity are more likely to have a gaming social identity salient during playing, than those who report low levels of identification with the gaming identity.

This study helps us to understand how pre-existing social categories impact group related behaviour in online situations. Thus, taken together with the findings of Lea et al. (2007) who manipulated identity salience using discussion topics, we can assume that in real-world, topical, online discussions in which a pre-existing shared social identity exists, individuals will self-categorise with this pre-existing social

identity when communicating with others. We therefore contend that it will be possible to observe group prototypical behaviour in anonymous online groups with pre-existing shared social categories. Moreover, we suggest that this prototypical behaviour will transcend local group norms; the behaviour of individuals communicating with a specific social identity salient will be comparable across different forums. In this way, identity-prototypical behaviour will be deduced from the global identity prototype as opposed to solely induced from local group norms. That is not to say that local group norms will not have some effect on behaviour induction (e.g., Jans et al., 2015; Le Henaff et al., 2015), just that we will also be able to find evidence that individual's behaviour is impacted by a global identity prototype.

In addition to the cognitive effects of depersonalisation discussed above, researchers have suggested that it is in fact greater anonymity of self to others that enhances group normative behaviour (Klein et al., 2007). This is because anonymity of the individual to others gives the individual the freedom and power to enact their identity how they see fit. In this way, the strategic component of the SIDE model has been extended to incorporate the 'purposeful expression (or suppression) of behaviours relevant to those norms conventionally associated with a salient social identity' (Klein et al., 2007, p.30). This component of the SIDE model is associated with social identity consolidation, namely the process of securing one's place within a group (Klein et al., 2007).

For the most part, the strategic aspect of the SIDE model has been studied with regards to how visibility to one's in- or outgroup impacts upon one's propensity to adhere to group norms. For example, in Reicher et al. (1998, Study 3) it was observed that when low status ingroup members (students) were accountable only to high status outgroup members (staff), they were less likely to endorse ingroup

behaviours that were deemed unacceptable to the outgroup, such as lying after missing class. However, when the students were visible to other ingroup members in addition to being visible to staff, participants increased their endorsement of ingroup activities that were deemed unacceptable by the staff outgroup. In this way, it was argued that visibility to ingroup members leads to the desire to behave group normatively and demonstrate oneself to be group prototypical regardless of the apparent scrutiny of high-status outgroup members. However, when outgroup members are the only ones to which a participant is held accountable, the need to express ingroup normativity decreases – at least within lower status groups.

In sum, research exploring the impact of anonymity and deindividuation suggests that when unidentifiable individuals communicate with anonymous ingroup members, the individual is more likely to perceive the others as similar to self and is also more likely to behave group normatively. Importantly, this depersonalisation and adherence to group norms only occurs when a shared social identity is explicitly salient. Whilst the research into this phenomenon in real-world environments has focused on local group identity salience (e.g., Le Henaff et al., 2015), we contest that research from the gaming self-report literature (Chen & Wu, 2015) indicates that wider social category prototypes also impact the behaviour of ingroup members. Specifically, we propose that in online groups whereby a pre-existing social identity is made explicit – i.e., groups centred around a well-defined social identity – individuals will become depersonalised in line with the global group prototype and thus behave accordingly.

However, in order to test this hypothesis, it is necessary to use a behavioural indicator that is relatively robust to the influence of local group norms. Previous research has indicated that individuals' opinions are highly influenced by direct

discussion with other ingroup members. For example, Koudenburg et al. (2019) directly observe how local groups who share a category level superordinate identity may polarise in different directions based on the perception of the local group norm. Similarly, we have observed from the research of Smith et al. (2015a) that the content of an identity in terms of beliefs as to what the identity stands for are flexible to change through local group interaction. In this way, local groups appear to have too large an impact on the content of attitudes and opinions for us to be able to detect a superordinate category level identity across different local groups. Instead, a behavioural indicator that is focused on the form of conversation as opposed to the content of conversation may be a more robust yet subtle indicator to how individuals perceive and share their social psychological reality (Koudenburg et al., 2017).

1.5   Psychology and Linguistic Style Literature

Whilst the majority of linguistic research within psychology has focused on the semantics of communication, a more recent trend has shifted towards the analysis of linguistic style due to its ubiquity in communication (Boyd & Pennebaker, 2017; Boyd & Schwartz, 2021; Pennebaker, 2011; Pennebaker et al., 2003; Tausczik & Pennebaker, 2010). In lay terms, style refers to 'how' a message is communicated, as opposed to 'what' is being said. A message can be articulated in many different ways whilst still retaining its meaning, and thus the stylistic (non-semantic) words used to convey a particular message are thought to be integral to understanding how individuals construct their own realities (Chung & Pennebaker, 2007; Pennebaker, 2011). Based on this assertion, a plethora of research has sought to identify the link between the way a person communicates and their individual personality (Boyd & Pennebaker, 2017; Mairesse et al., 2007; Tong et al., 2020), values (Boyd et al., 2015) and psychopathologies (Junghaenel et al., 2008) as well as demographic

factors such as gender (Newman et al., 2008) and age (Löckenhoff et al., 2008). Moreover, others have taken this approach further through looking at intra-individual changes in communication style. Communication Accommodation Theory (Gallois et al., 2016; Giles et al., 2010) and Linguistic Style Matching (Gonzales et al., 2010; Welbers & de Nooy, 2014) research suggest that social identities impact linguistic style in dyadic communication, whereas Koschate et al. (2021) demonstrate that social identity salience impacts communication style regardless of the direct presence of interlocutors. This section will provide an overview of this literature and make the case for using linguistic style to understand and empirically measure how individuals construct their social realities through shared cognitive prototypes.

A large body of research has suggested that personal identity traits such as personality can be predicted using linguistic style measures. In their seminal paper, Pennebaker & King (1999) noted correlations between Big 5 personality measures (openness to experience, conscientiousness, neuroticism, extroversion and agreeableness) and linguistic style measures. They showed that in essays written by students, aspects of individuals' language such as positive emotion words and social words corresponded to self-reported traits such as extraversion and neuroticism. Following this research, a large number of researchers have developed this initial finding using real world blog posts (Yarkoni, 2010), in-class writing (Komisin & Curry, 2012), day to day verbal conversations (Mehl et al., 2006), writings about recent personal loss (Baddeley & Singer, 2008) or other specific topics (Fast & Funder, 2008). Based on this literature, it has been assumed that individuals write in linguistic patterns that are relatively stable and unchanging across time and contexts (Boyd, 2017), and thus we can glean information about stable aspects of self, such as personality. However, for the most part, research into the impact of individual traits

on linguistic style has not factored in the role played by demographics and other social factors on linguistic style. By looking only at individual level comparisons, personality research misses the fact that whilst linguistic style may provide insight into the way individuals construct their personal realities, groups of individuals have shared social realities that are detectable in language (Koschate et al., 2021; Newman et al., 2008). By changing the lens of analysis from the individual level to the group level, we are thus able to observe how different types of groups construct their realities through analysis of their linguistic style.

Exploring how we can use linguistic style to understand differences between demographic groups has been on the research agenda for decades (e.g., Nguyen et al., 2016). However, findings are predictably inconsistent across studies. For example, when studying differences in linguistic style between males and females, Mulac et al. (1990) find that young boys are more likely to offer direct opinions whereas young girls are more likely to ask questions. However, this finding is reversed in men and women when looking at the ways in which professional criticism is given by managers with Mulac, et al. (2000) finding that women were more likely to give directives and men were more likely to ask questions. Similar contradictions have been noted when seeking to understand which gender uses more first person singular pronouns such as 'I'; Mehl and Pennebaker (2003) noted that women tend to use more first person singular pronouns such as I, whereas others have noted that men tended to have a higher usage of singular personal pronouns (Mulac et al., 2006). Whilst most of these studies used relatively small sample sizes and only one data source (i.e., descriptive essays, Mulac et al., 2006), Newman et al. (2008) attempted to address these inconsistencies by analysing 14,324 texts written in a variety of contexts including fiction stories, personal recounts on emotional events,

journal entries, exams and conversations. They found that across these different contexts, women and men's linguistic style substantially differed. For example, on aggregate women were more likely than men to use pronouns, negations and use more words per sentence. Meanwhile, men were more likely than women to use long words, articles, prepositions and numbers. Newman et al. (2008) thus concluded that they had found definitive evidence for differences in communication style across genders.

However, Newman et al. (2008) also reported a significant main effect for context; in the 37 different linguistic categories where a significant interaction between gender and context was noted (p <.001), 18 categories were found to be significantly different in the reverse direction. In other words, whilst on aggregate women's and men's linguistic style appears to differ, these differences are moderated by the context of communication. For example, whereas on aggregate women are noted as using more words per sentence than men, in fiction and journal entries they tend to use fewer words per sentence than men. Similarly, whereas on aggregate men are noted as using fewer negations and more numbers than women, in conversational contexts this finding reverses. This interaction therefore questions the extent to which stable communication style differences across genders can really be said to exist. In their introduction, Newman et al. (2008) note that by studying linguistic gender differences, we can gain an 'insight into how men and women approach their social worlds' (p. 212). However, the significant impact of context on their results coupled with the inconsistent literature outlined above, suggests that men and women may not *always* 'approach their social worlds' in different ways. Instead, and in line with self-categorization theory, it could be argued that men and women only communicate in gender prototypical ways when their gender identity is

salient. At other times, for example when giving professional criticism (Mulac et al., 2000), it is possible that gender identities are not salient and instead it is the salience of a work identity that influences communication style. This would explain the inconsistent findings of Mulac et al. (2000) when studying gender differences in work settings. This intra-individual variation mediated by social identity salience is an idea that has been further explored by Koschate et al. (2021).

Koschate et al. (2021) provide initial evidence to support the claim that individuals with a salient social identity communicate in predictable ways, regardless of the specific domain of communication. Using online forums as a proxy for identity salience, they demonstrated that feminists communicate using a distinctive linguistic style when compared to parents. Moreover, they also showed evidence of intra-individualistic style shifts wherein the same individual commenting in a parent forum had a different linguistic style fingerprint than when they commented in a parenting forum. Further, Koschate et al. (2021) provided direct evidence to indicate that this result could be explained by the salience of a social identity by directly manipulating social identity salience. Through activating the salience of a parent or feminist social identity and asking participants to write about a pre-specified issue, Koschate et al. (2021) noted that their classifier (which had been trained on online forum posts) was still able to detect whether a parent or feminist identity had been made salient within the participant using only linguistic style features. In this way, they demonstrated that even when individuals are not communicating directly with other ingroup members or members of the community that the model was trained on, the salience of a social identity impacts their style of communication. This thesis therefore seeks to extend this initial finding to explore whether the linguistic style fingerprints of different social identities can be used to understand group prototypes.

Previous research from Communication Accommodation Theory (CAT) has also pointed to the importance of social identity in communicative strategies (Giles et al., 2010). CAT is both an interpersonal and intergroup theory of communication which focuses on how and why individuals change their communication style during social encounters with others (Gallois et al., 2016; Giles et al., 2010; Giles & Ogay, 2007). In its initial formulation, CAT aimed to 'explore the sociopsychological parameters underlying the moves speakers make in their speech behaviors' (Gallois et al., 2005, p.7). Central to CAT is the idea that individuals adapt their linguistic style strategically to create and maintain personal or social identities in interpersonal communication. Thus, individuals 'converge' towards their interlocutor's communication in order to reduce the social differences between them and establish rapport. Convergence refers to the process of adapting one's language patterns to more closely resemble the language patterns of another individual or group whom one wishes to be associated with or whose approval one is attempting to gain. Conversely, individuals 'diverge' against an interlocutor's language patterns when attempting to accentuate differences in social identity. This may be observed when the individual dislikes the interlocutor (Doise et al., 1976), however may also be observed in asymmetric relationships, i.e. doctor-patient relations. In these scenarios, speech patterns between dyads diverge in order to maintain a professional level of social distance between the two individuals (Ahmed & Bates, 2016).

Interestingly however, Thakerar et al. (1982) noted that individuals' self-reporting of their accommodative strategies was not always in line with their linguistic behaviour. For example, in dyads characterized by status inequality, low status participants reported accommodating to high status individuals' communicative

styles, and yet empirical analysis of their linguistic behaviour indicated the opposite.

Whilst high-status participants slowed their speech and made their accents less

standardized, low status participants increased their rate of speech and used more

standardized accents. In this way, the low status individuals had accommodated to

their stereotypes of high-status speakers, but not to their actual speech patterns

(Thakerar et al., 1982). Similar findings have been demonstrated across groups such

as gender-based identities (Mulac et al., 1983), age groups (Caporael et al., 1983;

Ryan et al., 1986) and disability groups (Klemz, 1977; Ryan et al., 1986). This

relatively robust finding (Gallois et al., 2016) provides crucial evidence to support the

idea that linguistic style may be mediated by the way that individuals cognitively

represent groups; in this case, linguistic style behaviour is indicative of the

stereotypical representation of high-status speakers. Thus, to extend this finding, we

suggest that not only do individuals change their style based on stereotypical

attributes of outgroups in interpersonal communication (Thakerar et al., 1982), but

we suggest that style may also be used by participants to enact their own ingroup

prototypes. Thus, we suggest that the cognitive representation of the ingroup (i.e. the

prototype) can be studied through an analysis of an individual's linguistic style.

    The difference in approach between communication accommodation theorists

and the approach outlined by Koschate et al. (2021) lies in the domain of

communication. CAT focuses on how individuals enact their identities in response to

and in reactance from other individuals who are immediately present. In this way,

CAT focuses on how the communication and social identity of others may impact

upon one's own communicative strategies in interpersonal interactions; it is therefore

a theory of movement towards or away from others. In contrast to this however,

Koschate et al. (2021) suggest that individuals' linguistic style does not only vary in

relation to immediate others, but is mediated by an internal, cognitive understanding of who the individual is at a given point in time. Thus, in the same way that self-categorisation research has indicated that the salience of an identity impacts how one perceives and experiences the world (Turner et al., 1987; Abrams et al., 1990), Koschate et al. demonstrate that this salience of an identity is detectable in subtle linguistic cues. Of course, it is entirely possible that the presence of an interlocutor leads to the salience of a particular identity (Turner et al., 1987), but the interlocutor is not a prerequisite for observing identity congruent behaviour – identity salience is.

In sum, based on the notion that identity salience impacts linguistic style (Koschate et al., 2021), and that individuals behave in identity prototypical ways when an identity is salient (Turner et al., 1987), this thesis looks at measuring and modelling prototypical linguistic style behaviour. In this thesis therefore, the research questions of interest are:

1. Can linguistic style indicators be used to observe identity-prototypical behaviour that transcends local group norms? (Chapter 2)

2. Is there a relationship between identity prototypical behaviour and influence in real-world environments? (Chapter 3)

3. How can we use this measure of behavioural prototypicality to understand groups in real-world contexts? (Chapter 4)

2   Validating a behavioural measure of group-level prototypes

In this chapter, we aim to create a behavioural measure of context-independent prototypicality using naturally-occurring online linguistic data. As outlined in the literature review, self-categorisation theory proposes that a salient group membership leads to depersonalisation and hence prototypical behaviour. Whilst the majority of studies have focused on depersonalisation and prototypical behaviour at the local group level (e.g., Le Henaff et al., 2015; Postmes et al., 2000; Smith et al., 2015a), we suggest that in online communities where a shared pre-existing identity already exists, individuals will be influenced by a more global, context-independent group prototype.

This research extends the methodology of Koschate et al. (2021) who found that individuals who share a salient social identity communicate in group prototypical ways. They noted that forum names provided an effective way of predicting which pre-existing social identity would be salient in a given context. Thus, in their research, they established that individuals communicated in a group prototypical way when a particular identity was made salient. In this research, we extend this by i) using a measure of group-level prototypical behaviour to understand individual-level prototypicality, ii) using more advanced machine learning classifiers to capture the diversity of styles that may be prototypical of a single salient social identity, iii) training this classifier on identities not used in the original research, and iv) controlling for forum-specific topics in naturally-occurring data in order to rule out topic specificity as a moderating cause of prototypical behaviour.

2.1 Identities of Interest

The present research was conducted in liaison with the National Crime Agency (NCA). The aim of this industry relationship was to help the NCA better

understand the social identities of individuals operating in darknet cryptomarkets. This research therefore focuses on the social identities believed to be present on these darknet cryptomarkets. Below, I will provide some background information as to what darknet cryptomarkets are, and how the existing literature informed our choice of social identities and thus forum-based data.

Cryptomarkets are hidden websites which allow for the anonymous sale of illicit goods, predominantly drugs, using cryptocurrencies such as bitcoin. Cryptomarkets have been referred to as an 'eBay for drugs' (Maddox et al., 2016) – they host multiple sellers, selling multiple products and display customer reviews for each product available (Barratt & Aldridge, 2016). The Silk Road was the first cryptomarket to exist on the darknet and was active from February 2011 until the arrest of its creator in October 2013. Since its demise, many other cryptomarkets have risen in its place, with authorities estimating that up to $300 million worth of drugs were sold online in 2016 (World Drug Report, 2018).

More recently, research has turned towards the users of these cryptomarkets in an attempt to understand the cryptomarket community on a more social level (Maddox et al., 2016; Masson & Bancroft, 2018). Masson and Bancroft interviewed nine cryptomarket users in an attempt to understand how these users constructed the morality of their illegal activity and what motivated them to visit cryptomarkets. Interestingly, the primary conclusion of Masson and Bancroft's research was that 'Cryptomarket exchange is a form of social action that is not restricted to its economic value for participants.' (2018, p.79). Therefore, whilst previous research focussed on the financial motives behind using cryptomarkets (Aldridge & Décary-Hétu, 2016), Masson and Bancroft identified a more social motive.

The idea that social action is important to cryptomarket users shows synthesis with the interview-based research of Maddox et al. (2016). Maddox and colleagues found that individuals use cryptomarkets as a form of 'constructive activism'. Constructive activism is explained by Maddox et al. as the enactment of libertarian social ideals via a more permissive digital reality. In this way, individuals are able to discuss and share ideas and drug-taking behaviours that are generally stigmatised by the wider community, whilst directly acting out these libertarian ideas through the buying and selling of drugs. This notion of the importance of libertarian ideals is further supported by qualitative research undertaken using data taken directly from Silk Road forums by Munksgaard and Demant (2016). Munksgaard and Demant found that the discussion of libertarian ideals was highly prevalent on the Silk Road prior to the arrest of Ross Ulbricht (the creator of the Silk Road). Following his arrest, the majority of communication on cryptomarkets became more business oriented and pragmatic. Based on the above literature, it appears that there are two primary motives for users of cryptomarkets; a libertarian motive and a financial motive.

For the purpose of this research, we are interested in understanding the possible identities that are present in online cryptomarkets. Based on the above research, it appears that both libertarians and businesspeople operate on cryptomarkets. Further, whilst there are no doubt differences between those who believe libertarian ideals and those who enact their libertarian ideals through drug dealing and the use of cryptomarkets, or similarly between those who are businesspeople versus those who are criminal drug vendors, we nonetheless hypothesize that we can utilise a broader understanding of the non-criminal social identities in order to study the more niche cryptomarket identities.

2.2   Present Research

We therefore first test the hypothesis that we can measure group-level prototypicality through analysing differences in linguistic style between libertarians and entrepreneurs. Our first research question is:

Can we detect a globally prototypical style of communication in both libertarians and entrepreneurs?

In order to address this question, we perform a between-groups analysis using online forum data. Using this dataset, we create a classifier that is able to differentiate between the prototypical linguistic style of libertarians and entrepreneurs. In the remaining studies in this chapter, we test this classifier on a variety of different datasets in order to validate that it is measuring group-level identity prototypicality by eliminating other potential explanations. Finally, in the last two studies of the chapter we assess the link between individual-level prototypicality and self-reported measures.

In Study 1b, we perform a within-individual analysis by tracking the same users across forums in order to rule out demographics as a potential explanation for our initial result, and to examine intra-individual changes in prototypical behaviour when a different identity become salient. In the third study (Study 1c), we confirm that our classifier is capturing identity prototypicality which transcends the immediate local context by testing our classifier on data taken from the Silk Road darknet forums. In our fourth study (Study 1d), we use topic modelling in an attempt to control for the content of online forum posts; we test our classifier on posts from the libertarian and entrepreneur forums which are most similar in content, finding that the

classifier is still able to detect an identity prototypical communication style even when individuals from different groups are discussing similar topics. In the fifth study (Study 1e), we test our classifier on data taken from survey responses in order to both control more tightly for topic, as well as to further demonstrate that our classifier is able to identify an identity-prototypical communication style when identity salience is manipulated using traditional psychological methodologies. Finally, in Study 1f, we repeat the experimental manipulation used in Study 1e to explore the link between behavioural prototypicality and self-reported prototypicality.

2.3   Study 1a: Training a Linguistic Classifier to Assess Identity Prototypicality

2.3.1   Hypothesis

In this first study, our aim is to ascertain whether libertarians and entrepreneurs communicate in distinguishable linguistic styles. Our first hypothesis therefore states that we will be able to detect from which forum (either entrepreneur or libertarian) a text originated using only linguistic style features.

2.3.2   Method

2.3.2.1   Data Collection

We first collected data from individuals writing with either an entrepreneur or a libertarian identity salient. In line with Koschate et al. (2021), we use forum name as a proxy for identity salience. We assume that individuals posting in either an entrepreneur forum or a libertarian forum have the respective identities salient at the time of writing, thus making this data suitable for studying identity-prototypical behaviour. Next, we train our classifier to distinguish whether a contribution

originated in either of these two forums, thereby creating a model that assesses a group prototypical linguistic style.

After receiving ethical approval from the University of Exeter Psychology Ethics Board, data were collected from the Reddit "Libertarian" and "Entrepreneur" forums, known as "subreddits". Google BigQuery was used to gather the data.

We collected one year's worth of posts and comments for both subreddits. We collected the title, text, URL and author of all posts and comments submitted to the Libertarian and Entrepreneur subreddits in 2018. In total, we collected 1,932,334 contributions to the Entrepreneur and Libertarian subreddits. This comprised 41,933 posts and 334,001 comments to the Entrepreneur subreddit ($n = 375,934$) and 65,048 posts and 1,491,352 comments to the Libertarian subreddit ($n = 1,556,400$).

2.3.2.2 Quantification of Linguistic Style

To linguistically analyse the data, we used Linguistic Inquiry and Word Count software (LIWC 2015) (Pennebaker et al., 2015). LIWC uses a bag-of-words language model, so that word order is ignored. It counts the number of words classified into particular linguistic categories, for example affective words, adverbs, future tense words (see Pennebaker et al., 2015, for further detail) and computes a percentage value for each document, reflecting the proportion of a particular feature in a document.

LIWC is a commonly used software package amongst computational social scientists and is favoured amongst researchers seeking to understand psychological processes or differences between individuals. It has been suggested that by looking at the categories of words that individuals use we can learn more about an individuals': attentional focus (Boyd, 2017; Rude et al., 2004), status (Kacewicz et al., 2014), cultures (Michel et al., 2011), attempts to deceive (Newman et al., 2003),

close relationships (Slatcher et al., 2008), thinking styles (Pennebaker et al., 2014) and individual differences such as personality (Mairesse et al., 2007) (for detailed reviews see Boyd & Schwartz, 2021; Pennebaker et al., 2015; Tausczik & Pennebaker, 2010). Based on the range of psychological studies that LIWC software has been used for, we suggest that it is also well-suited to understanding how different social identities are enacted and expressed online.

For our analysis, we were interested in using only the LIWC categories that could be considered as linguistic style. We define style as the part-of-speech categories that are used widely across different contexts and domains regardless of topic (e.g., pronouns and articles) (Schwartz et al., 2013). For this reason, we omitted all LIWC categories that refer to topical or content-based categories such as the 'family', 'power' and 'risk' categories. We also omitted the summary categories provided by the 2015 LIWC software that were an amalgamation of individual word categories such as 'Clout' and 'Authenticity' (Pennebaker et al., 2015). Resultantly, the textual data from each Reddit post was converted into a vector with 41 stylistic features (see Appendix A).

2.3.2.3   Data Preparation

In order to train our model to predict identity-prototypical behaviour, it was necessary to exclude any data that may adversely impact our ability to draw robust psychological conclusions. Firstly, we removed posts containing only URLs. Entrepreneur forum moderators aim to remove any posts that contain only links themselves, and so we only removed one post. In the libertarian subreddit, URLs are frequent and so we removed 52,767 posts at this stage.

We then omitted posts and authors that had been deleted or removed by moderators. Next, we removed posts made by self-identified bots. Bots are

automated scripts used to provide information to users of a subreddit (Massanari, 2013). We removed all submissions containing the word 'bot', as bots often identify themselves using phrases such as 'I am a bot'. We also removed authors with 'bot' in their name and the 'AutoModerator'.

Next, we omitted submissions with fewer than 50 words. As outlined above, the linguistic analysis software used in this research is a word count software which uses percentages to determine the proportion of words in a text that belong to a specific linguistic category. Consequently, in texts with low word counts, particular categories may be highly over-weighted; for example, in a text with only 10 words, each word is weighted at 10%. As a result of this, it is understood that texts with higher word counts allows researchers to draw more robust and reliable psychological conclusions (Boyd, 2017). Further, Chung and Pennebaker (2019), the developers of LIWC software, advise using a minimum cut-off of 100 words where possible. In practice, word count cut-offs are often lower than this, especially when using sparser social media data (50 words – Bäck et al, 2018; 25 words - Koschate et al., 2020; 45 words - Nelson et al., 2017; 50 words – Petrie et al., 2008; 50 words - Wilson, 2019). Our choice of 50 words was made in order to keep as much data as possible in our analysis, whilst ensuring that the data could be used to draw psychologically meaningful conclusions (Boyd, 2017; Pennebaker Conglomerates, 2017).

After data exclusion, we had a sample size of $N = 373,825$ posts ($n = 286,940$ from the Libertarian subreddit and $n = 86,885$ from the Entrepreneur subreddit; see Analytic Strategy below for how we deal with this disparity). In total, there were $N = 27,225$ individuals posting in the Libertarian forum and $N = 25,824$ individuals posting in the Entrepreneur forum. The mean number of words per contribution was 116 ($Med = 85$, $SD = 102$) in the Libertarian forum, and 135 ($Med = 91$, $SD = 194$) in

the Entrepreneur forum. Information pertaining to the number of posts removed at

each stage of the data preparation process are outlined below in Table 2.1.

Table 2.1

Detail of Data Preparation Process

Excluded Data	Libertarian Contributions		Entrepreneur Contributions	
	Removed	Remaining	Removed	Remaining
Initial		1,556,400		375,934
URL removal	52,767	1,503,633	1	375,933
Deleted/removed submissions	22,680	1,480,953	27,164	348,769
Deleted/removed authors	229,383	1,251,570	56,620	295,149
Bot removal in text	39,878	1,211,692	20,653	274,496
Bot removal in author	5,818	1,205,874	1,172	273,324
Submissions > 49 words	918,934	286,940	186,439	86,885
Total posts		4,436		16,203
Total comments		282,504		70,682
Total		**286,940**		**86,885**
Total authors		**27,225**		**25,824**

2.3.2.4 Analytic Strategy

A Random Forest algorithm was used to classify posts as either libertarian or

entrepreneur based on patterns in linguistic features. Random Forests are non-

parametric, supervised learning methods, comprised of multiple decision trees each

voting on which class (in our case, forum) a particular datapoint belongs to. Decision

trees are trained through repeatedly splitting the dataset into subsets consisting of similar datapoints until each datapoint is classified into pre-specified categories. At each split, the dataset is divided based on the value of one of its features. The resultant model is then tested on previously unseen data in order to gauge how successful the model is at distinguishing between the two classes (forums) based on the features (linguistic style variables).

In traditional Random Forest classifiers, the feature and threshold chosen to split the data is mathematically optimal (see Cutler et al., 2012 for more information). For this research, however, we used an Extremely Randomised Trees ("Extra Trees") classifier which chooses the best feature-threshold combination for each split from a small randomly-chosen set (Geurts et al., 2006). In this way, the Extra Trees model is less likely to overfit the training data through a more efficient method of reducing variance and bias within the dataset. Furthermore, due to the randomised procedure of splitting the data, Extra Trees are less computationally expensive.

Imbalanced class sizes can adversely impact a classifier's ability as merely choosing to classify every post as one of the majority class can still achieve an apparently high accuracy. In order to deal with the imbalanced class sizes of our dataset, we undertook random under-sampling of the majority class whereby we ran the classifier on 75,118 randomly chosen posts taken from the Libertarian subreddit in order to match the 75,118 submissions in the Entrepreneur subreddit. We used the same number of comments ($n = 70,682$) and posts ($n = 4,436$) in each sample. We repeated this process 10 times to ensure there was no significant variance between each subset of 75,118 Libertarian forum submissions.

For the initial analysis, we trained and tested our Extra Trees model using 75,118 posts from both the Libertarian and Entrepreneur subreddits. We entered all

41 LIWC style features into the model (see Appendix A). To prevent overfitting and lower the bias of our model, we used k-fold cross-validation with $k = 10$. In k-fold cross validation, a subset (fold) of the data (1/10 in 10-fold cross validation) is held out of the training set and is used to validate the model. In this way, the data is trained on 9/10 of the data, and then validated on the 1/10 that has been held out. This cross-validation is completed 10 (k) times, until all folds have been held out of the training set and used in validation. Empirical evidence has shown k-fold cross-validation using $k = 5$ or $k = 10$ yields test error rate estimates that exhibit neither high estimates of bias nor inflated variance (Kuhn & Johnson, 2013). For more information about how k-fold cross validation see Appendix B.

2.3.2.5 Classifier Feature Importances

Further, we also extracted the importance of each of the 41 features entered into our model in order to ascertain which features may be irrelevant or redundant to the classification process. Using redundant features can lead to models that overfit the training data and therefore perform worse during testing. It is therefore important to select the most predictive features.

The importance of each feature can be calculated through understanding how decision trees use features to split the data. Decision trees split the data into smaller and smaller subsets in an attempt to group similar datapoints together. Every node in a decision tree represents a condition on how to split the data based on a single feature. Features that are located near to the root of the tree are more important as they divide a greater proportion of the data. Feature importance calculations also take into account how successful the node is at splitting the data cleanly into the two target classes. This is referred to as the *impurity* decrease as the splits are chosen so that the dataset becomes more cleanly separated (less impure) at each split.

Therefore, features that neatly slice the data into the two distinct classes enable a greater decrease in the total impurity. Feature importance is then assessed by the impurity decrease that is due to splitting on a particular feature, weighted by the proportion of the data that is split. The importance of feature *j* at a particular node is thus:

$$h_j(i) = w_j H_j - w_{j,left} H_{j,left} - w_{j,right} H_{j,right}$$

where *hⱼ(i)* is the importance of node *j*, *wⱼ* is the proportion of samples reaching node *j*, *Hⱼ* is the impurity value of node *j*, and left and right refer to the two 'child' nodes of node *j* following the split. The overall importance *I(k)* of feature *k* is then calculated as:

$$I(k) = \frac{\sum_{j \sim k} h_j(k)}{\sum_j h_j(k)}$$

where the sum in the numerator runs over all the nodes that split on feature *k* and the sum in the denominator runs over all the nodes.

These are then normalised by dividing over the sum of all feature importances. Finally, in random forests such as Extra Trees, the feature importances are averaged out over all of the trees. In Figure 2.1 below, we can see that there are 9 variables (shaded in grey) that have higher importance than the other features entered into the model.

Figure 2.1

Normalized Impurity-based Feature Importances in Classifying Forum Posts.
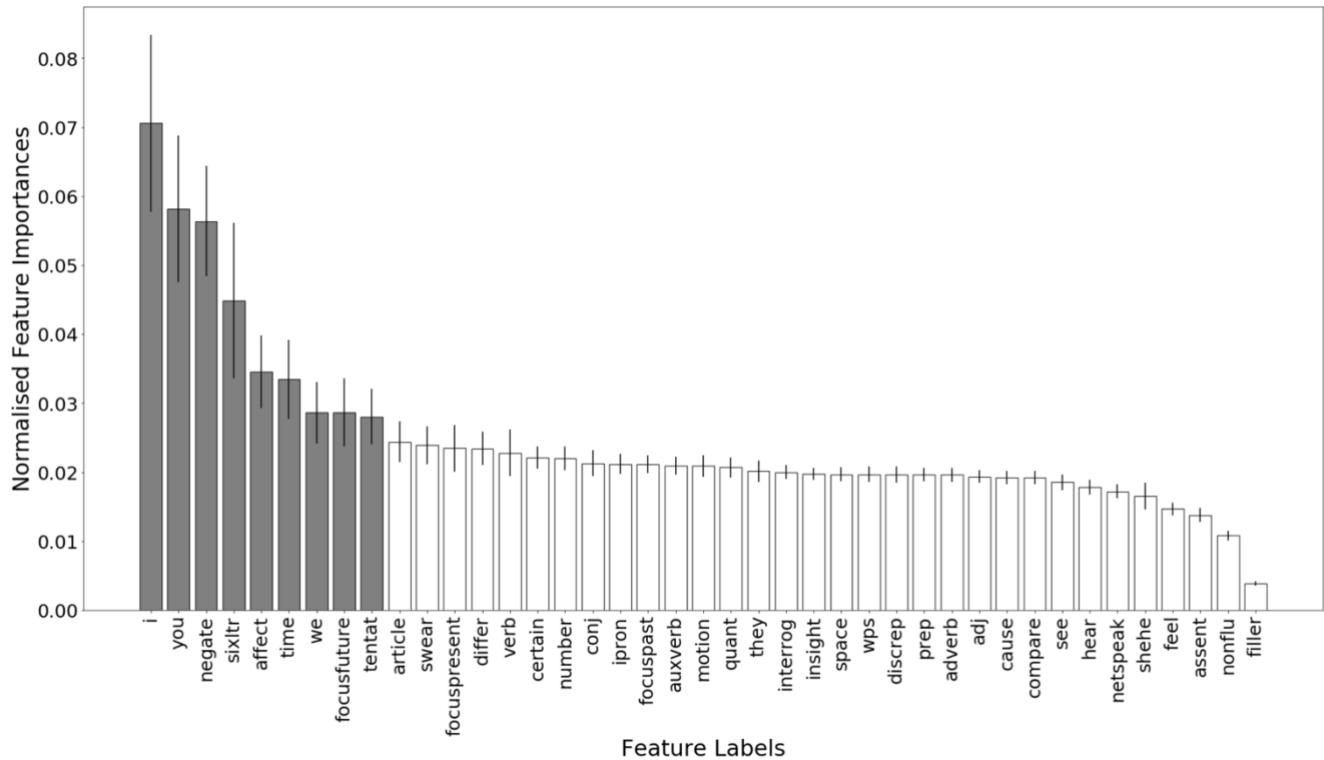
Note. Error bars represent standard deviation (inter-trees variability).

However, one possible concern with using impurity-based measures for feature selection is that they can place too much importance on features that have high cardinality (are more commonly used). For example, it is possible that in our model, the singular first-person pronoun ('I') is being viewed as important as singular first-person pronouns are very common across the dataset and therefore provide more information to the model than less common features such as 'netspeak'.

One way to test this is through performing random permutations. Random permutations randomise the values of each feature in order to ascertain the decrease in the model score when individual features are randomised. This procedure removes the relationship between the feature and the target variable

(whilst keeping the overall distribution of values equal), and therefore enables us to understand how strongly the model depends on each individual feature.

We ran this approach to further verify our 9 selected features. The results of this additional analysis are displayed below in Figure 2.2. We find that the same 9 features are still the most predictive, although we also find that swear words are important too. As both approaches – the impurity decrease approach and random permutation approach – identified first-person singular pronouns (I), second-person singular pronouns (you), negations, six letter words, affective language, time words, first-person plural personal pronouns (we), future focussed language and tentative language as the most important features, we used these nine features in our analysis.

Figure 2.2

Importance of Linguistic Features Calculated Through Random Permutations
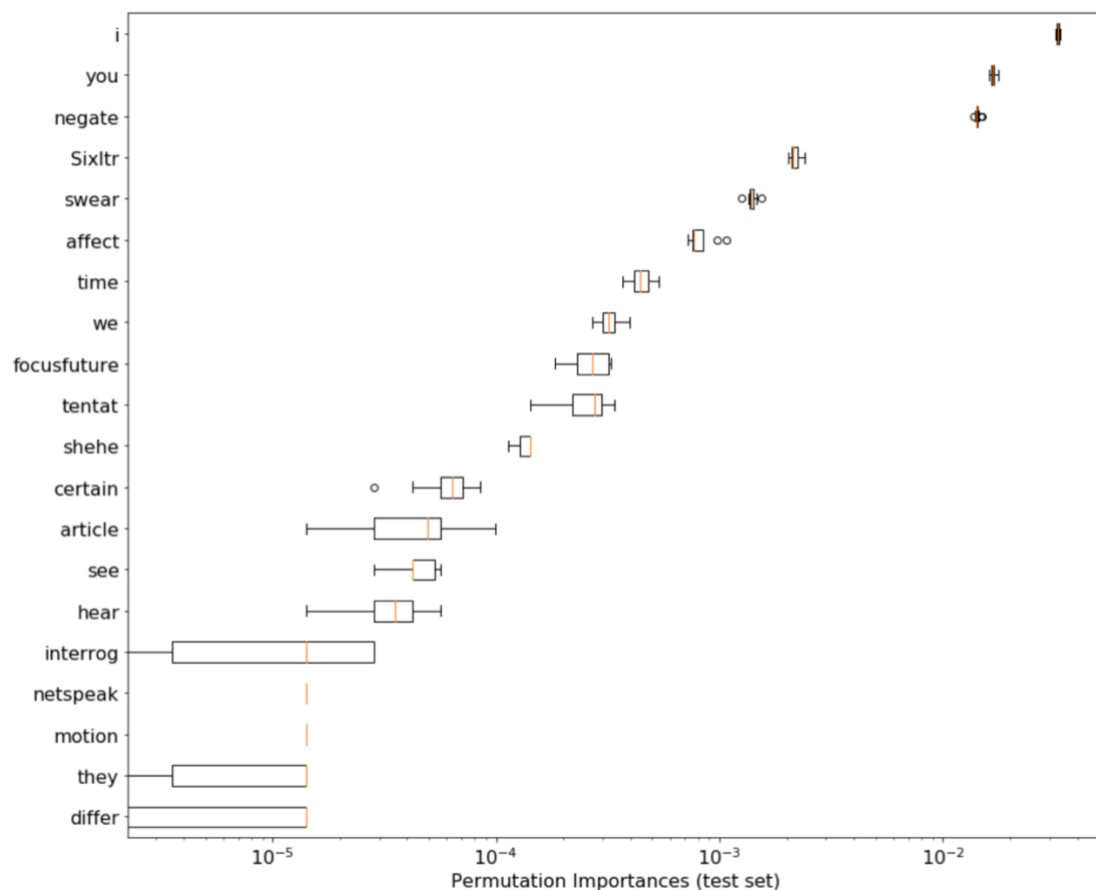
Note. The X-axis corresponds to how important the features were with more important features on the right-hand side of the plot.

The orange lines denote the median importance of each feature when the relationship between the target variable and the feature value is removed.

The outliers correspond to anomalous calculations of the feature importances when the analysis was repeated 100 times.

All other linguistic features not included on the plot had a median lower than 10^{-5}

2.3.3   Results

To assess whether random under-sampling was a robust technique to address the imbalanced class sizes, we ascertained how the model performed when a random subset of 75,118 posts was taken from the Libertarian forum to match the 75,118 submissions in the Entrepreneur forum. Table 2.2 below indicates the Area Under the (receiver operating characteristic) Curve (AUC) and accuracy scores for the classifier and associated SEMs on each trial. As evidenced in Table 2.2, the AUC and standard error of each trial showed little variance, verifying the use of random under-sampling to correct for the class imbalance.

Table 2.2

Comparison of the Mean AUCs and Accuracies with SEMs for Each Subset of 75,118 Posts Taken from the Libertarian Forum Used in Training the Model

Trial	AUC	SEM	Accuracy	SEM
1	.857	.004	.778	.005
2	.857	.004	.779	.003
3	.857	.004	.778	.004
4	.856	.005	.778	.004

5	.856	.002	.777	.003
6	.857	.003	.778	.004
7	.856	.005	.777	.006
8	.857	.004	.779	.001
9	.857	.004	.777	.001
10	.857	.002	.778	.004
Mean	**.857**	**.004**	**.778**	**.004**

For the main analysis, when training our model on the top nine most important features, the model achieved an AUC of .83 and an accuracy of .76. The confusion matrix below (Figure 2.3) illustrates the percentage of posts that were correctly and incorrectly classified.

Figure 2.3

Confusion Rate Matrix for Nine-feature Classifier (Between-Groups Classification)

	Predicted Entrepreneur	**Predicted Libertarian**
Actual Entrepreneur	77% (28,943)	23% (8,616)
Actual Libertarian	25% (9,560)	75% (27,999)

2.3.4 Discussion

Our results demonstrate that linguistic style between libertarians and entrepreneurs differs sufficiently to detect which group a text stems from. This provides support for our first hypothesis which suggests that libertarians and

entrepreneurs communicate in distinguishable communication styles, thus showing synthesis with the results of Koschate et al. (2021).

Moreover, the confusion matrix (Figure 2.3) indicates that by using only nine linguistic features we are able to correctly classify 77% of entrepreneur posts and 75% of libertarian posts. Thus, the variation in linguistic style between these two identities can be captured ~76% of the time using only nine linguistic style features. In other words, if a random post were picked from either the libertarian or entrepreneur forum, the classifier could correctly identify the forum of origin 76% of the time. By comparison, we know that a classifier choosing at chance would get an accuracy of 50%, whereas a classifier choosing perfectly would achieve an accuracy of 100%. In turn then, an accuracy of 76% speaks to the idea that there is a substantial amount of predictable variation between the two groups.

Additionally, the AUC measure takes into consideration the confidence with which the classifier is able to predict the class of each datapoint (in this case forum). In this way, the AUC differs from the accuracy measure as accuracy only computes whether the classifier has identified the correct class (i.e., is above or below the 0.5 threshold). By comparison, the AUC is computed through changing the threshold and assessing whether the classifier is still able to predict the correct class. In this way, it is able to differentiate between posts that are hard to classify (posts close to the 0.5 threshold), and posts that are easier to classify (posts that are distinctively higher or lower than the 0.5 threshold).

Further, several papers have made attempts to compare AUCs to more well-known effect size measures such as Cohen's *d*. In Rice and Harris (2005), the authors suggest that an AUC of .556 is comparable with a small effect size (Cohen's $d = 0.20$), an AUC of .639 corresponds with a medium effect size (Cohen's $d = 0.50$),

and an AUC of .714 corresponds with a large effect size (Cohen's d = 0.80). In turn then, an AUC of .83 (as achieved in the present study) corresponds with a Cohen's d of 1.36. Here then, we can see that the linguistic style differences between libertarians and entrepreneurs are notably large.

However, one possible explanation for this finding is that individuals posting in the Libertarian subreddit are demographically different to those posting in the Entrepreneur subreddit. Previous research has highlighted that demographic factors, such as age and gender, may impact linguistic style (Lockenhoff et al., 2008; Newman et al., 2008). In Study 1b, we validate our model by testing it on individuals who have posted in both the Libertarian and the Entrepreneur forums. By completing a within-person analysis, we are able to exclude stable differences such as demographic and personality differences as an explanation for the results of Study 1a.

2.4 Study 1b: Excluding Confounds Based on Differences in Stable Characteristics

2.4.1 Hypothesis

As outlined above, it is possible that the main driver of linguistic style differences between the Libertarian and Entrepreneur forums is in fact due to demographic and personality differences between the individuals who post in both of these forums. In Study 1b, we hypothesize that our classifier will still be able to differentiate between the prototypical linguistic style of the two social identities even when the text is written by the same individual (H2). This will enable us to both exclude demographic and personality differences as an explanation for our results, as well as to demonstrate that individuals shift their linguistic style in line with salient

social identities. This would therefore contrast with the personality-based linguistic style research which assumes a level of stability over time and context (Boyd, 2019).

2.4.2 Method

2.4.2.1 Test Dataset

To exclude demographic and stable personality factors as alternative explanations, we performed a within-person analysis; we used data from individuals who had posted in both the Entrepreneur forum and the Libertarian forum. For each individual, we calculated their average linguistic scores from posts in both forums. In total, 441 users contributed submissions of over 50 words to both forums; this gave us a test dataset consisting of $N = 882$ (441 individuals with two scores each). For the training dataset, we used the remaining submissions from both forums from authors not included in our test set. After random under-sampling, our training dataset consisted of $N = 142,320$ posts, with $n = 71,160$ from each forum. We used the nine-feature model outlined in Study 1a on our within-person test dataset.

2.4.3 Results

When training the classifier using nine data-driven features outlined in Study 1a and testing the classifier on posts from individuals who post in both the Entrepreneur and the Libertarian forum, our classifier achieved an AUC of .78 (Cohen's $d = 1.10$; Rice & Harris, 2005) and an accuracy of .72. The confusion rate matrix (Figure 2.4), indicates that the classifier was approximately equally successful at classifying libertarian and entrepreneur identities.

Figure 2.4

Confusion Rate Matrix for Nine-feature Classifier (Within-Individual Classification)

	Predicted Entrepreneur	**Predicted Libertarian**
Actual Entrepreneur	75% *(329)*	25% *(112)*
Actual Libertarian	30% *(131)*	70% *(310)*

2.4.4   Discussion

The results of Study 1b indicate that when we exclude demographics and other stable characteristics as a possible explanation for the findings of Study 1a, our model is still able to correctly identify which of our two identities was salient at the time of writing based on a group prototypical style.

The results from the within-person analysis suggest it is possible to detect intra-individualistic style shifts using only nine linguistic features. This result validates the model as it illustrates that the classifier can distinguish between the linguistic style of libertarians and entrepreneurs even when we exclude demographics as an explanation. In line with findings from Koschate et al. (2021), the within-participant analysis shows that individuals shift their linguistic style according to the social identity that is salient in the social context. This result serves to underline the dynamism and fluidity of linguistic style that is often overlooked.

However, another possible explanation for these findings is that individuals may be accommodating to the style of the forum and thus these results may be platform dependent. Communication Accommodation Theory (Gallois et al., 2005; Giles et al.,

2010) suggests individuals linguistically converge with those they are in discussion with, and therefore our results may instead be explained by this individual accommodation to local forum norms. Instead, we propose that individuals have an idea of what is socially normative for their identity (an identity prototype [Hogg, 2001]) and thus behave in line with this prototype regardless of the online platform on which they are posting. In this way, we are interested in a context-independent prototype as opposed to the locally derived prototype as measured in previous studies such as Postmes et al. (2000). In order to rule out local accommodation and local prototypes as an explanation for these results then, we examine a platform-based explanation in Study 1c.

2.5   Study 1c: Excluding Local Norms as an Explanation

2.5.1   Hypothesis

In Study 1c, our primary interest is in understanding whether the results from the previous studies can be explained purely by accommodation to local forum norms. Here, we argue that individuals have a cognitive representation of their social identity which prescribes identity congruent behaviour. More specifically, we hypothesize that our model will still be able to detect prototypical linguistic styles of libertarians and entrepreneurs on a different platform (Silk Road) (H3).

2.5.2   Method

2.5.2.1   Test Dataset

To exclude local norms as an explanation for the results of Studies 1a and 1b, we used data from the Silk Road which has been collated and made publicly available by Branwen and colleagues (2015). Two forums were identified that linked strongly to the libertarian and entrepreneur social identities; the 'Vendor Roundtable'

forum and the 'Philosophy, Economics and Justice' forum (Munksgaard & Demant, 2016). The 'Vendor Roundtable' forum consisted of vendors discussing ideas to improve their business models whilst the 'Philosophy, Economics and Justice' forum consisted of political and philosophical discussion related to libertarian ideas. The initial sample consisted of $N = 20,836$ posts, with 10,780 originating in the 'Economics, Philosophy and Justice' forum and 10,056 originating in the 'Vendor Roundtable' forum. After removing posts with fewer than 50 words, this left $N = 10,494$ posts with $n = 4,746$ posts from 553 users in the Vendor Roundtable forum and $n = 5,748$ posts from 1,131 users in the Philosophy, Economics and Justice forum. We used the Reddit-trained nine-feature model outlined in Study 1a and tested it on the Silk Road dataset.

2.5.3   Results

When using the model trained on nine features, the classifier achieved an AUC of .65 (Cohen's $d = 0.55$), and an accuracy of .61. Further inspection of the confusion rate matrix (Figure 2.5) suggests that the classifier is better at classifying entrepreneurs than libertarians; the classifier correctly classifies 69% of entrepreneurs, but only 55% of libertarians.

Figure 2.5

Confusion Matrix for Nine-feature Classifier Tested on Silk Road Data

	Predicted Entrepreneur	Predicted Libertarian
Actual Entrepreneur	69% (3,259)	31% (1,487)
Actual Libertarian	45% (2,614)	55% (3,134)

2.5.4   Discussion

By testing our model on data taken from the Silk Road, we were able to explore whether the results of Studies 1a and 1b could be explained by accommodation to local norms. The results of Study 1c indicate that when using a Reddit-trained classifier on Silk Road data, the classifier was still able to distinguish between our two identities 65% of the time. According to Rice and Harris (2005) an AUC of .65 corresponds to a Cohen's *d* effect size of 0.552 thus demonstrating a medium effect. However, further inspection of the confusion matrix revealed that a significant proportion of the libertarian posts were being misclassified as entrepreneur.

A likely reason that the AUC was still showing a medium effect size despite the confusion matrix demonstrating misclassification of libertarian posts is that the threshold to determining libertarians from entrepreneurs may be different on the Silk Road as to compared to Reddit. This is a common issue within classification research and points to the difficulty of testing a classifier on a dataset that is different from the training set (i.e., testing outside of the specific Reddit forum on which the data were trained). However, whilst the classifier is misclassifying libertarian posts, the AUC of .65 gives us an idea of how confidently the classifier is making predictions. That is, with the classifier classifying everything under 0.5 as entrepreneur, and everything over 0.5 as libertarian, the medium effect size AUC coupled with the results from the confusion matrix suggests that perhaps the libertarian posts are close to the boundary of classification. In other words, there are likely to be many libertarian posts that are between 0.4 and 0.5, which are being incorrectly classified as entrepreneur as they are below 0.5 (hence the results in the confusion matrix), but this is because the Reddit classification threshold is set at a different position to the Silk Road threshold. Thus, due to the differences in the

distribution of the data across platforms, this leads to a lower overall accuracy. In turn, this is why the AUC is a better measurement for learning about the differences between the groups than accuracy. With the AUC still being 0.65, we can see that there are definitive differences between the libertarian's communication style and entrepreneur's communication style on the Silk Road, it is just that these differences may not align perfectly with the threshold used on the training dataset (e.g., Reddit platform).

Additionally, the aim of this study was to indicate that the results of Study 1a and 1b are explained by more than just local norms. We appreciate that other factors may impact communication style, for example conversational partners as evidenced in the Communication Accommodation Theory literature (see Giles et al., 1991), however, we argue that despite the influence of domain specific factors, the fact that the classifier is able to classify both identities at a higher rate than chance (50%), indicates that local norms alone cannot explain the findings in Study 1a and 1b. It may well be possible that similar local norms exist on both the Silk Road and Reddit platforms, however the fact that they are similar for each identity speaks to the idea that these are not forum-specific local norms that arise out of interactions. Instead, these are norms that transcend the local group context and speak more to the idea that they are identity-specific, context-independent norms. This provides evidence to support the notion of a cognitive psychological identity construct that influences communication style regardless of the immediate local context.

Further, the difficulty in predicting libertarian identities on the Silk Road may be because Silk Road libertarians are influenced by a different prototype than the libertarians on Reddit. In this research, we make the assumption that we can utilise a broader understanding of identity prototypes in order to classify identities on the Silk

Road. However, the findings of this analysis suggest that there may be something unique about libertarian identities on the Silk Road that our classifier is not currently capturing. For example, it could be argued that libertarians on the Silk Road are more extreme in their identities than those using Reddit. This explanation is in line with the interview-based research of Maddox et al. (2016) who find that darknet cryptomarkets 'facilitate a shared experience of personal freedom within a libertarian philosophical framework' (p.111). In this way, being present on the Silk Road is a libertarian act of rebellion against the rules of society and thus likely attracts a more 'extreme' libertarian than those debating ideas on Reddit.

Nonetheless, the possible discrepancy between Silk Road and Reddit libertarians points to a novel way of comparing the similarity of identity prototypes through language analysis. In order to further understand how libertarians operating on the Silk Road may perceive their social identities, we could use classifiers trained on other similar identities, such as anarchists or intellectuals, in order to understand which identity-prototypical communication style is closest to that of the Silk Road libertarian. This idea is explored further in Section 3 of the thesis.

A further confound that may impact the results of our analysis is the content of the communication. It is highly likely that communities of entrepreneurs and libertarians are not discussing similar topics, and thus the choice of topic content may be driving the classification rate. In Studies 1d and 1e, we seek to control for the possible impact of content using both naturally-occurring forum data (Study 1d) as well as in controlled experimental conditions (Study 1e).

2.6 Study 1d: Excluding Content of Communication as an Explanation

2.6.1 Hypothesis

In Study 1d, we control for the role of content in the classification process in a forum setting by testing our classifier on posts that are similar in content across the forums. Whilst the classifier only uses stylistic features, it is possible that content may be having an impact on the way in which individuals are communicating. Therefore, in order to rule out topic as a possible confound, we tested our model on posts that were deemed to be similar in their semantic content. This enabled us to ascertain whether content was in fact playing a key role in the classification process.

Our hypothesis remains that we will still be able to predict whether the texts were written in a prototypically libertarian or an entrepreneur style, even if the semantic content of the posts is highly similar. We therefore still expect to achieve an AUC greater than chance level (.50).

2.6.2 Method

2.6.2.1 Analytic Strategy

In order to understand whether the linguistic style classifier is impacted by the content of the communication, we isolated posts that were most similar in content so that we could test our classifier on this subset. By testing on the most similar posts, we were able to understand whether content is having an impact on the results of the classifier.

In order to isolate the posts that were most similar in content for use in a test set, we employed a word embedding approach called doc2vec (Mikolov et al., 2013). Generally speaking, word embedding approaches aim to understand the particular meaning of a word based on other words in the same sentence. This is best

summarised by an oft repeated quote from John Firth which states 'You shall know a word by the company it keeps' (Firth, 1957, p.11).

Word2vec, from which doc2vec was developed, works by predicting words from the other words around it and is considered one of the foundations of Natural Language Processing (NLP) (Ganegedara, 2018). Word2vec is a shallow, two-layer neural network which aims to represent words as numerical vectors wherein the similarities of the words can be calculated through ascertaining the distance between two word-vectors. The input to the model is a large corpus of words, and the output is a vector space whereby each word is assigned a unique vector made up of numbers. All vectors have a specific orientation and these orientations can be seen to encapsulate the relations between the words. For example, word2vec can capture relations such as verb tenses, capital cities and gender (see Figure 2.6).

Figure 2.6

Graphical Visualisation of Word Embedding Relations as Captured Through Spatial Relations

Male-Female   Verb Tense   Country-Capital
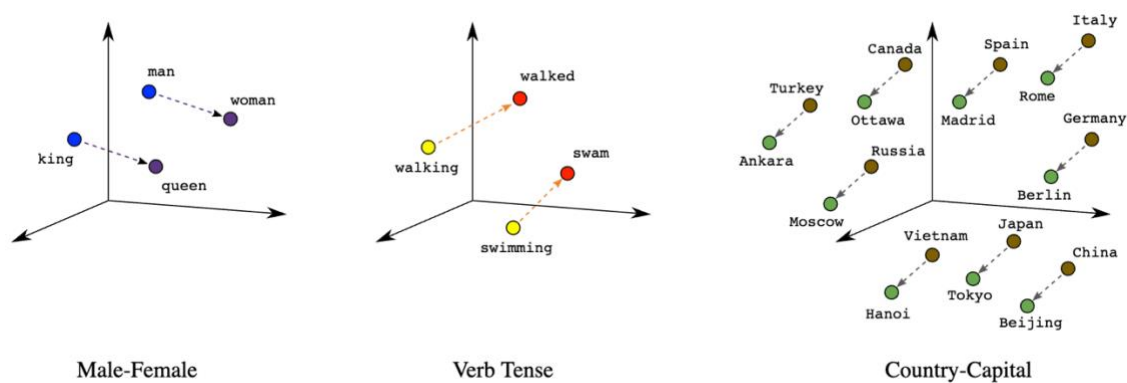
Note. Adapted from *Embeddings can produce remarkable analogies*. (https://developers.google.com/machine-learning/crash-course/embeddings/translating-to-a-lower-dimensional-space) by Google Developers is licensed under CC BY 4.0.

By extension, doc2vec (document to vector) aims to create a numeric vector representation for each document as opposed to for each individual word. Doc2vec works by adding another vector which captures the paragraph (document) ID. Whilst word vectors aim to capture the concept associated with a particular word, document vectors aim to capture the concept underlying a whole document. Previous research has found doc2vec to be a state-of-the-art method for understanding similarities between texts (Lau & Baldwin, 2016). For more information about word2vec and doc2vec approaches see Mikolov et al. (2013) and Le and Mikolov (2014).

In order to isolate posts from both Reddit forums that are most semantically similar, we trained a doc2vec model on the 150,236 posts used in Study 1a (75,118 posts from each forum). We pre-processed the data by lowercasing all sentences and removing all punctuation. After training our model, which gave us unique document vectors for each of the 150,236 posts included in the original analysis, we computed the pairwise cosine similarity for each pair of posts where one had originated in the Libertarian forum and one in the Entrepreneur forum. Cosine similarity refers to the cosine of the angle between two vectors in an n-dimensional space. It is calculated through taking the dot product of the two vectors and dividing this by the product of the length of the vectors. The formula is:

$$similarity(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}}$$

where A and B refer to the two vectors, and i refers to each number within the vector.

By calculating the cosine similarity between each libertarian post and each entrepreneur post, this enabled us to identify those posts which were most semantically similar (those which had the closest vectors). We isolated the posts that had a cosine similarity of higher than 0.8. This gave us a test dataset of 27,547 posts (n = 13,779 from the Libertarian forum and n = 13,768 from the Entrepreneur forum). We then used the remaining 122,689 posts as a training dataset (61,339 from the Libertarian forum and 61,350 from the Entrepreneur forum). This training set consisted of Libertarian and Entrepreneur posts that were not similar in semantic content. Thus, using the nine features outlined in the previous studies, we trained the classifier on the 122,689 posts least similar in content, and tested our classifier on the 27,547 posts that were most similar in content.

2.6.3 Results

When using the model trained on nine features, the classifier achieved an AUC of .83 (Cohen's d = 1.36; Rice & Harris, 2005), and an accuracy of .76. Inspection of the confusion rate matrix (Figure 2.7) suggests that the classifier is equally able to classify both libertarian and entrepreneur posts.

Figure 2.7

Confusion Rate Matrix for Nine-feature Classifier Tested on Semantically Similar Data

	Predicted Entrepreneur	Predicted Libertarian
Actual Entrepreneur	74% (10,226)	26% (3,542)
Actual Libertarian	22% (3,088)	78% (10,691)

2.6.4 Discussion

In Study 1d, we aimed to rule out the possibility that differences in content were the driving force behind the differences in communication style picked up by our classifier. Whilst we do not dispute that our two groups are very likely to be discussing different topics in their respective forums, we nonetheless find that when there is overlap in the content discussed by our two groups, we are still able to detect identity prototypical behaviour using only our nine linguistic style features. Moreover, we show that the AUC when using only similar content in testing is identical to that of Study 1a, thus supporting the hypothesis that content of communication does not impact our stylistic classifier.

This finding provides an important justification for using stylistic indicators as opposed to content-based indicators. As argued in the literature review, the aim of this research is to create a classifier that can work across domains and across different contexts. This is particularly relevant for studying criminal groups wherein the content of communication may be masked or 'off-topic', but the style of communication may still be indicative of one's social identity. In this study, therefore, we have coherently demonstrated that the classifier is able to detect whether the style of textual data is prototypical of either a libertarian or entrepreneurial identity, regardless of the topics being discussed.

The doc2vec approach is a robust methodology for assessing the similarity of texts in naturally-occurring datasets. However, one possible downside to this approach is the inability to observe whether the model, using a neural network approach to understand statistical regularities in the data, is truly picking up on semantic similarities between texts or whether there are artefacts of the data that are driving the analysis. For example, previous researchers have stressed the

importance of 'clean' data when using word embedding based approaches, and thus it is possible that there are artefacts within the Reddit dataset such as unidentified URLs that may influence the vector representation of particular documents. Whilst we did clean the dataset thoroughly prior to analysis (see Table 2.1, Study 1a), it is important to acknowledge possible limitations associated with black-box neural network approaches.

Therefore, in order to more transparently observe the impact of content on our stylistic classifier, the next study (Study 1e) uses experimental data wherein content is controlled via the design of the study. Therefore, in Study 1e we return to traditional experimental methodologies in order to directly control for the influence of content on stylistic classification and to validate that we are studying a cognitive notion of identity prototypicality.

2.7 Study 1e: Experimental Validation Controlling for Topic, Audience, and Demographic Differences

In the next study, we aimed to cross-validate our approach using traditional offline psychological methods. Thus far, in line with the research of Koschate et al. (2021), we have made the assumption that an individual communicating in an online forum is likely to have the corresponding social identity salient. However, one way to further verify this assumption is to manipulate identity salience under controlled conditions (Verkuyten & Hagendoorn, 1998). This enables us to demonstrate that we are detecting behaviours that are the direct result of a cognitive change in salience of a global identity that transcends the immediate local context. Therefore, in this study, we aim to further validate the finding of Study 1c by demonstrating that the classifier is not merely detecting the adoption of local norms or local group prototypical

behaviour, but is instead detecting an identity prototypical linguistic style which is independent of the local context.

In this study, we also aim to provide concurrent validity to our classifier by testing the relationship between self-reported prototypicality and prototypical behaviour. Whilst no such research has attempted to assess whether self-reported judgments of one's own prototypicality link directly to one's behavioural prototypicality, in this study we test this idea using the classifier developed in the previous studies.

2.7.1 Hypothesis

The first hypothesis of Study 1e is that we will still be able to use our classifier trained on naturally occurring online data to detect the identity-prototypical style of texts that have been written by individuals with the respective identities salient in controlled, experimental conditions. Further, our second hypothesis states that there will be a positive correlation between self-reported prototypicality and behavioural prototypicality.

2.7.2 Method

2.7.2.1 Participants

After receiving ethics from the University of Exeter's Psychology Ethics Committee, we recruited a total of 110 participants who self-identified as both entrepreneurs and libertarians. Participants were recruited using the online platform Prolific Academic. First, we used Prolific to conduct a pilot study in order to identify individuals who self-identified as both libertarian and entrepreneurs. In this pilot study, we asked individuals whether they considered themselves to be i) an entrepreneur and ii) a libertarian. The pilot study was sent to 600 participants, 110 of whom self-identified as entrepreneurs and libertarians. According to Faul et al.

(2007), with a sample of 90, we have power of 0.8 to detect a correlation of the size 0.25 or higher.

It was important that all participants identified as both entrepreneurs and libertarians as we wished to control for possible demographic and personality differences that may exist between the two groups. A total of 110 participants were recruited for the study, however many of our respondents did not write enough text in order to be included in the final analysis. As our classifier is trained on texts which are over 50 words (in line with the recommendations outlined in Study 1a), we needed to ensure that only individuals who had written more than 50 words were included in the analysis. In total, this left us with 45 respondents of whom 24 (53%) were female and 21 were male participants. Their ages ranged from 18 to 64, with the majority of respondents (38%) reporting that they were in the 45-54 age bracket. Participants came from a diverse range of countries, although the majority were from the UK (23 participants, 51%) or North America (11 participants, 24%). With a sample of 45 respondents, we had power of 0.8 to detect a correlation 0.36 (Faul et al., 2007).

The study follows a 2 (identity salience: libertarian v entrepreneur) x 2 (topic: identity-relevant, identity-neutral) design. We used the same topics in both conditions in order to control for the content of communication and to understand how the relevance of the topic to the identity impacted the classifier's ability to detect an identity-prototypical style. In this way, topic is a within-subject factor and identity is a between-subjects condition. We used the automatic randomization feature in the Qualtrics software to randomize our participants to either the libertarian or entrepreneur condition. Independent t-tests revealed no significant demographic differences between the two groups.

2.7.2.2   Procedure

Participants were informed that we were interested in studying the communication style of libertarians and entrepreneurs. We provided definitions for both identities, stating that 'By 'libertarian' we are referring to individuals who subscribe to the political philosophy that takes individual liberty to be the primary political value' and 'By 'entrepreneur' we are referring to individuals who have set up a business (or businesses), are self-employed or earn money buying and selling items online for a living.' Participants were also informed that they would be reimbursed £2.00 for the time taken to participate in the study. All participants consented to take part in the study.

In the first stage of the study, we verified that all participants identified as both libertarian and entrepreneur. As we had already conducted a pre-screening study used to identify individuals with both of these identities, the vast majority of individuals confirmed this to be the case. The six participants who reported one or neither of the chosen identities were thanked for their interest and debriefed.

The remaining 104 participants were assigned to either the entrepreneur or the libertarian stream. They were then asked questions about both their social identities such as their level of identification using four items taken from Doosje et al. (1997) with Cronbach's alpha = .924 (Libertarian), and .945 (Entrepreneur), and four items intended to measure the degree of self-reported prototypicality (adapted from van Knippenberg & van Knippenberg, 2005) with Cronbach's alpha = .885 (Libertarian) and .949 (Entrepreneur).

The next part of the study involved manipulating the salience of their respective identities using a commonly used manipulation also employed by Koschate et al (2021). This manipulation asks individuals to write three things that

group members do often, rarely, well and badly (Haslam et al., 1999). It is intended to make an identity salient by encouraging individuals to consider what behaviours are prototypical of the identity without introducing the comparison to an outgroup (Haslam, 2004; Koschate et al., 2021).

After this salience manipulation, individuals were asked to write a minimum of 3 sentences about two topics. These topics were the same for both groups. One topic was an identity-relevant topic and asked participants to give their opinion on the 45% tax bracket of individuals earning over £150,000 in Britain, and the other was an identity-neutral topic which asked individuals to comment on the idea that wisdom comes with age. The topics were randomised such that some participants saw the identity relevant topic first whilst others saw it last.

Next, we performed a manipulation check by asking individuals how easy it was to think of themselves as either libertarian or entrepreneur (Haslam, 2004).

2.7.3 Results

Using the nine-feature classifier from Study 1a, we tested the classifier on written submissions that exceeded 50 words. In total, 62 posts contained more than 50 words with 29 from individuals with a libertarian identity salient and 33 from individuals with an entrepreneur identity salient. Using the classifier trained on Reddit data, we found that we were able to detect an identity-prototypical style with an AUC of .71 (Cohen's $d = 0.78$; Rice & Harris, 2005). We also noted that the classifier was slightly better at detecting libertarians than entrepreneurs (see Figure 2.8).

Figure 2.8

Confusion Rate Matrix for Nine-feature Classifier Tested on Experimental Data

(Study 1e)

	Predicted Entrepreneur	**Predicted Libertarian**
Actual Entrepreneur	61% *(20)*	39% *(13)*
Actual Libertarian	31% *(9)*	69% *(20)*

Interestingly, when we separated between the identity relevant (tax question) and identity neutral (wisdom question) topics, we found that the classifier performed better when individuals were writing about identity neutral topics (AUC = .76, Cohen's d = 1.00; Rice & Harris, 2005) and dropped when they were writing about identity-relevant topics (AUC = .66, Cohen's d = 0.58; Rice & Harris, 2005). It was also evident that when we divided the data down into identity relevant and neutral questions, the classifier became better at detecting the prototypical libertarian style as opposed to the prototypical entrepreneur style (see Figures 2.9 and 2.10 below).

Figure 2.9

Confusion Rate Matrix for Nine-feature Classifier Tested on Identity Neutral

Experimental Data (Study 1e)

	Predicted Entrepreneur	**Predicted Libertarian**
Actual Entrepreneur	67% *(10)*	33% *(5)*
Actual Libertarian	29% *(4)*	71% *(10)*

Figure 2.10

Confusion Rate Matrix for Nine-feature Classifier Tested on Identity Relevant

Experimental Data (Study 1e)

	Predicted Entrepreneur	**Predicted Libertarian**
Actual Entrepreneur	56% *(10)*	44% *(8)*
Actual Libertarian	33% *(5)*	67% *(10)*

In addition to the results of the classifier, the means and standard deviations for

identification, self-reported prototypicality and the salience manipulation check are

detailed in Table 2.3. No significant differences between the groups were observed

for the self-reported variables.

Table 2.3

Mean and Standard Deviation for Self-Reported Variables

	Libertarian		Entrepreneur	
	M	*SD*	*M*	*SD*
Identification	5.53	1.27	5.28	1.35
Self-reported prototypicality	3.72	0.88	3.30	1.19
Salience manipulation check	3.83	1.13	4.0	0.93

However, due to the low number of participants whose data exceeded 50 words ($n = 33$ entrepreneurs, $n = 29$ libertarians), we decided that the correlational analysis between self-reported prototypicality and behavioural prototypicality per identity would not have enough power to draw any robust conclusions from. According to Faul et al. (2007), with a sample size of 29 and a power of 0.8, we would only be able to find an effect size of 0.44 or higher. It therefore seemed appropriate to rerun Study 1e and to complete this correlational analysis once we had collected more data (Study 1f).

2.7.4 Discussion

The results of Study 1e demonstrate that we can predict an identity-prototypical style even when individuals are writing outside of group contexts (e.g., in forums). This further supports the argument made in Study 1c, suggesting that our behavioural measure is detecting a context-independent prototypical style as opposed to only a locally prototypical style or the desire to to manage one's impressions within a group (Klein et al., 2007). This finding provides further evidence to suggest that the psychological group membership exists within the individual

(Turner et al., 1987), and can be cognitively activated using salience manipulations. Thus, even when acting in isolation individuals are still influenced by a cognitive representation of the group. This work shows synthesis with the findings of Koschate et al. (2021) who also found that the salience of particular social identities could be predicted in short, isolated texts taken from online surveys.

Also in synthesis with the findings of Koschate et al. (2021), we noted that the identity-neutral topic elicited a linguistic style that was more identity-prototypical than the identity-relevant topic. In this research, for our identity relevant topic we used a question that was relevant to both identities concerning individual taxation. We also ensured that all the participants identified both as libertarians and as entrepreneurs. It is however possible that this question led to the salience of their other social identity – i.e., an individual in the entrepreneur condition may have automatically considered this question more in light of their libertarian identity as opposed to their entrepreneurial identity. It is well-known that the issue of taxation is of particular significance to libertarians (Long, 1998), and thus it is possible that this may have impacted the responses of those in the entrepreneur condition (see Figure 2.10).

An alternative explanation for this finding is that individuals may have felt greater need to demonstrate their social identity through the way in which the identity-neutral topic was answered. Asking individuals to answer an identity-neutral topic regarding whether wisdom comes with age from their libertarian or entrepreneurial identity may have led them to feel greater pressure to demonstrate their libertarian or entrepreneurial prototypicality through the way in which the message was communicated, as there is not an identity-prototypical semantic answer for this question. This is in contrast to the identity-relevant question whereby individuals commenting on tax from a libertarian or entrepreneurial perspective would likely

have been aware of the identity-prototypical answer for their respective identities. Herein lies the inherent value of using linguistic style to understand identity prototypicality; in instances where the content of communication may not be relevant to a particular social identity, we are still able to assess the prototypicality of individuals based on their style alone. This also enables us to transcend topics or platforms as demonstrated in Study 1c and Study 1d.

One limitation of this study pertains to the low amount of data included in the analysis. As our classifier was trained on data over 50 words (see Study 1a methodology for detailed rationale), we deemed it suitable to only test our classifier on data with a similar word count. This is in part due to the proportional nature of LIWC word category calculations. Nonetheless, Goksuluk et al. (2016) indicate that a total sample size of 42 (with 21 for each class) is sufficient to test the classifier with power of .78, an AUC of .71 and a Type I error rate of 0.05. With sample sizes of 29 and 33 for our libertarian and entrepreneur identities respectively, we can therefore rely on the finding that independent of topic, our classifier was able to correctly identify the prototypical styles of our identities with sufficient power. Similarly, when we disaggregated the data and analysed the written text for each question separately, our analysis was still suitably powered. Using the identity-neutral texts, Goksuluk et al. (2016) suggest that with a sample size of 14 in each class, an AUC of .76 gives us power of .8. Using the identity-relevant textual data however, our sample size of 33 (with 15 and 18 in each class) and AUC of .66 only gives us power of .48. It is therefore important that we are careful how we interpret this finding.

As outlined in the results section, we did not have enough data in order to run a well-powered correlation analysis. Based therefore on the insufficient amount of data collected, we decided to rerun this study (Study 1f), this time including a minimum

character limit on the textual questions in order to force participants to write a minimum of 50 words.

2.8 Study 1f: Assessing the concurrent validity of the classifier using self-reported measures of prototypicality

2.8.1 Hypotheses

The final study of the chapter aimed to provide concurrent validity to our measure of prototypicality by comparing the results of the classifier with self-reported measures of prototypicality. The method of this study thus follows the method of Study 1e, with a small adaptation used to increase the word count of the responses received. We also included more measures of prototypicality in order to gain greater understanding into the relationship between self-reported prototypicality and behavioural prototypicality. Our hypothesis remains the same as Study 1e, predicting that there will be a positive correlation between self-reported prototypicality and behavioural prototypicality.

2.8.2 Method

2.8.2.1 Participants

After receiving ethics from the University's Departmental Ethics Board, we recruited a total of N=185 participants who had previously completed a pre-screening study wherein they identified themselves as both entrepreneurs and libertarians. Participants were recruited using Prolific's online platform.

As before, it was important that all participants identified as both entrepreneurs and libertarians as we wished to continue to control for possible demographic differences that exist between the two groups. Whilst we conducted a pre-screening survey to identify our target group, we noted that a significant number (~8%) of

participants reported contradictory answers regarding their group identification. Of the 185 participants recruited who stated that they identified both as an entrepreneur and a libertarian in the pre-screening study, 15 participants later reported that they did not identify with both identities when asked again three days later (the time between the pre-screening study and main survey). These participants were excluded from the study at this point.

Of the remaining 170 participants, eight participants' data were excluded for not filling out both the questions necessary for analysis, one participant's data was excluded as the survey was completed in Spanish, and a further two participants did not write enough text to be included in the analysis. This left us with a total of 159 participants (99 (62%) male, 59 (37%) female, one unreported gender). Their ages ranged from 18 to 64, with the majority of respondents (44%) reporting that they were in the 25-34 age bracket. With a sample size of 159, we had power of 0.8 to detect a correlation 0.20 (Faul et al., 2007).

As before, the study follows a 2 (identity salience: libertarian v entrepreneur) x 2 (topic: identity-relevant, identity-neutral) design. We used the same topics as had been used in Study 1e; an identity-relevant topic concerning the UK's highest tax bracket, and an identity-neutral topic concerning the idea that wisdom comes with age. The topic remained a within-participants factor and identity salience a between-participants factor. The automatic randomization feature in the Qualtrics software was again used to randomize our participants to either the libertarian or entrepreneur condition. Independent t-tests revealed no significant demographic differences between the two groups.

2.8.2.2 Procedure

As in Study 1e, participants were informed that we were interested in studying the communication style of libertarians and entrepreneurs. We also provided definitions for both identities, stating that 'By 'libertarian' we are referring to individuals who subscribe to the political philosophy that takes individual liberty to be the primary political value' and 'By 'entrepreneur' we are referring to individuals who have set up a business (or businesses)'. Participants were informed that they would be reimbursed £2.00 for the time taken to participate in the study. All participants consented to take part in the study.

We then verified that all participants identified as both libertarian and entrepreneur. As we had already conducted a pre-screening analysis to identify individuals with both of these identities, the majority of individuals confirmed this to be the case. The 15 participants who reported one or neither of the chosen identities were thanked for their interest and debriefed.

The remaining participants were first asked about their demographic information and then about the business that they had set up. They were then assigned to either the entrepreneur or the libertarian condition. In each condition, the salience of either a libertarian or entrepreneur social identity was manipulated as in Study 1e.

Participants were then asked about the level of identification they had with the given salient social identity using four items taken from Doosje et al. (1997), with a Cronbach's alpha of .918 (Libertarian), and .842 (Entrepreneur), and several items intended to measure the degree of self-reported prototypicality. To assess prototypicality, we used four items to assess an 'ideal' notion of prototypicality, with a Cronbach's alpha of .869 (Libertarian), and .885 (Entrepreneur), and five items to

assess an 'average' notion of prototypicality, with a Cronbach's alpha of .873 (Libertarian), and .863 (Entrepreneur). These measures were taken from Steffens et al. (2021) meta-analysis and are listed in Table 2.4 below.

Table 2.4

Prototypicality Measures Included in Study 1f

Question	Adapted from	Ideal or average
Do you personify the positive values of [the group]?	Leon et al. (2009)	Ideal
Do you exemplify what it means to be a [group member]?	Janson et al. (2008); Steffens et al. (2014)	Ideal
Would you consider yourself as a model [group member]?	Steffens et al. (2014)	Ideal
Do you act according to the goals and values of [group members]?	Sluss (2006)	Ideal
Do you represent what is characteristic about [group members]?	Giessner et al. (2013); Leon et al., (2009); Platow & van Knippenberg, (2001); Steffens et al., (2014); van Knippenberg & van Knippenberg, (2005)	Average

How similar are you to other [group members]?	Platow & van Knippenberg, (2001); van Knippenberg & van Knippenberg, (2005)	Average
Do you have a lot in common with other [group members]?	Janson et al. (2008); Platow & van Knippenberg, (2001); van Knippenberg & van Knippenberg, (2005)	Average
Do you think you are a good example of a [group member]?	Platow & van Knippenberg, (2001); van Knippenberg & van Knippenberg, (2005)	Average
Do you embody what it means to be an [group member]?	Steffens et al. (2014); van Knippenberg & van Knippenberg, (2005)	Average

In the next part of the study, individuals were asked to write about the two topics. The main change between Study 1e and this study was the introduction of a minimum character limit of 300 characters (approximately 50 words). This was added in an attempt to ensure that participants wrote enough text to be analysed by our classifier, an issue that we had flagged up previously during Study 1e.

Finally, we again performed the manipulation check by asking individuals how easy it was to think of themselves as a libertarian or entrepreneur (Haslam, 2004).

2.8.3   Results

2.8.3.1   Classification analysis

As before, we ran the textual data through the LIWC 2015 software (Pennebaker et al., 2015) in order to derive the linguistic style of each post. We used

the nine most important style features outlined in the previous sections in our

analysis.

Firstly, using the nine-feature classifier from Study 1a, we tested the classifier

on our participants' written data. We found that our nine-feature classifier trained on

Reddit data was marginally better than chance at detecting an identity-prototypical

style with an AUC of .56 (Cohen's $d = 0.21$; Rice & Harris, 2005). However,

inspection of the confusion matrix revealed that the classifier was unable to correctly

classify the majority of the entrepreneur posts (see Figure 2.11).

Figure 2.11

Confusion Rate Matrix for Nine-feature Classifier Tested on Experimental Data

(Study 1f)

	Predicted Entrepreneur	**Predicted Libertarian**
Actual Entrepreneur	41% *(58)*	59% *(82)*
Actual Libertarian	29% *(51)*	71% *(127)*

Again however, when we analysed the two topics separately (identity-relevant

and identity-neutral), we found that the classifier performed better when individuals

were writing about an identity neutral topic (AUC = .58, Cohen's $d = 0.28$; Rice &

Harris, 2005) and dropped slightly when they were writing about an identity-relevant

topic (AUC= .55, Cohen's $d = 0.17$; Rice & Harris, 2005). Through inspecting the

confusion matrices, it was evident that when we divided the analysis down into

identity-relevant and identity-neutral questions, we could see the classifier was notably worse at classifying entrepreneur posts in the identity-relevant condition.

Figure 2.12

Confusion Rate Matrix for Nine-feature Classifier Tested on Identity Neutral Experimental Data (Study 1f)

	Predicted Entrepreneur	Predicted Libertarian
Actual Entrepreneur	51% (36)	49% (34)
Actual Libertarian	34% (30)	66% (59)

Figure 2.13

Confusion Rate Matrix for Nine-feature Classifier Tested on Identity Relevant Experimental Data (Study 1f)

	Predicted Entrepreneur	Predicted Libertarian
Actual Entrepreneur	31% (22)	69% (48)
Actual Libertarian	24% (21)	76% (68)

2.8.3.2 Analysis of the Link Between Behavioural Prototypicality and Self-reported Prototypicality

Outlined in Table 2.5 below are the means and standard deviations for identification, self-reported prototypicality and the salience manipulation check.

Table 2.5

Mean and Standard Deviation for Variables in Study 1f

	Libertarian		Entrepreneur	
	M	*SD*	*M*	*SD*
Identification	5.48	1.26	6.00	1.07
Self-reported prototypicality	3.36	0.80	3.42	0.91
Salience manipulation	3.64	0.81	3.63	1.08

To assess the relationship between self-reported prototypicality and behavioural prototypicality, individual level behavioural prototypicality was calculated using the probabilistic output of the classifier outlined above. The probabilistic output of the classifier refers to the likelihood that the classifier predicts a post as being libertarian or entrepreneur. A score closer to 0 indicates a more prototypically libertarian linguistic style, whereas a score closer to 1 indicates a more prototypically entrepreneur linguistic style. The libertarian scores were reversed coded for intuitiveness; consequently, higher behavioural prototypicality scores relate to higher prototypicality for both identities.

For each identity, we used a Spearman's rank correlational analysis to assess the association between self-reported prototypicality, identification and behavioural prototypicality.

2.8.3.2.2 Libertarian Condition

The results from the correlation analysis are displayed in Table 2.6 below. As can be seen, no significant correlations were found between self-reported prototypicality and behavioural prototypicality in either the identity-neutral condition, or the identity-relevant condition.

Table 2.6

Correlations Between Self-reported Prototypicality and Behavioural Prototypicality in the Libertarian Condition

Variable	*n*	*M*	*SD*	1	2	3	4
1. Identity relevant prototypicality score	89	0.35	0.22	-			
2. Identity neutral prototypicality score	89	0.45	0.23	.21*	-		
3. Self-reported prototypicality	89	3.4	0.80	.16	-.12	-	-
4. Self-reported identification	89	3.0	1.26	-.08	-.03	.48***	-

Note. *p < .05, ** p < .01, *** p < .001

However, due to the high correlation between identification and self-reported prototypicality, we reran the analysis, this time controlling for identification. As noted in Table 2.7 below, a significant positive correlation was observed between self-reported prototypicality and behavioural prototypicality for the identity-relevant question. However, it must be noted that with a sample of 89 and an effect size of 0.22, this only gives us power of 0.67 (Faul et al., 2007) and thus these results should be interpreted cautiously.

Table 2.7

Correlations Between Self-reported Prototypicality and Behavioural Prototypicality in the Libertarian Condition When Controlling for Identification

Variable	n	M	SD	1	2	3
1. Identity relevant prototypicality score	89	0.35	0.22	-		
2. Identity neutral prototypicality score	89	0.45	0.23	.21	-	
3. Self-reported prototypicality score	89	3.4	0.80	.22*	-.12	-

Note. Controlling for identification

Note. * $p < .05$

2.8.3.2.3 Entrepreneur condition

The results from the entrepreneur analysis are displayed in Table 2.8 below. In the entrepreneur condition, we also observed no relationship between self-reported prototypicality and linguistic style prototypicality in either the identity neutral or identity relevant condition.

Table 2.8

Correlations Between Self-reported Prototypicality and Behavioural Prototypicality in the Entrepreneur Condition

Variable	*n*	*M*	*SD*	1	2	3	4
1. Identity relevant prototypicality score	69	0.40	0.22	-			
2. Identity neutral prototypicality score	69	0.51	0.25	.16	-		
3. Self-reported prototypicality	69	3.4	0.91	.00	-.23	-	-
4. Self-reported identification	69	4.0	1.08	-.04	-11	.76***	-

Note. $*p < .05$, $** p < .01$, $*** p < .001$

However, once again we decided to repeat the analysis and control for identification due to the high correlation between identification and self-reported prototypicality. As noted in Table 2.9 below, we still observed no significant correlations between self-reported and behavioural prototypicality. The relationship between self-reported prototypicality and the prototypicality of responses to the identity neutral question were close to significance ($p = .056$), although in the opposite direction to our predictions. That is, the more prototypical an individual reported being, the less prototypical their response was when answering the identity neutral question. However again, it must be noted that with a sample size of 69 and an effect size of -.23, this gives us a power of 0.61 (Faul et al., 2007). Again then, it is important to be cautious when interpreting these results.

Table 2.9

Correlations Between Self-reported Prototypicality and Behavioural Prototypicality in

the Entrepreneur Condition When Controlling for Identification

Variable	*n*	*M*	*SD*	1	2	3
1. Identity relevant prototypicality score	69	0.40	0.22	-		
2. Identity neutral prototypicality score	69	0.51	0.25	.16	-	
3. Self-reported prototypicality	69	3.4	0.91	.05	-.23[†]	-

Note. Controlling for identification

Note. *p < .05, ** p < .01, *** p < .001

Note. † denotes *p* = .056

2.8.4   Discussion

This study aimed to replicate the results demonstrated in Study 1e and provide

concurrent validity to the measure of prototypicality through comparing it to existing

prototypicality measures. In the classification analysis, we noted surprisingly lower

AUCs across all three analyses (identity-neutral, identity-relevant and combined)

when compared to the findings of Study 1e. However, despite these low AUCs, we

observed a significant positive association between self-reported prototypicality and

behavioural prototypicality in the libertarian condition when participants were writing

about an identity-relevant topic. We also noted a close to significant negative

correlation between self-reported prototypicality and behavioural prototypicality in the

entrepreneur condition when participants were writing about an identity-neutral topic.

However, the power of both these results are low and thus must be interpreted

cautiously. This discussion will first explore possible reasons for the lower AUCs in Study 1f when compared to Study 1e, and then go onto to consider the results of the correlational analysis.

2.8.4.1 Classification analysis

One possible explanation for significantly lower AUCs in this study, as compared to Study 1e and the results of Koschate et al. (2021), pertains to the motivation for writing. Whilst in both Study 1e and in Koschate et al., (2021) participants were simply asked to write a minimum of 2-3 sentences, in this replication we used a minimum character limit to force a longer response from the participants. As a result of this minimum character limit, it is plausible that participants wrote longer texts simply in order to finish the task and earn the monetary reward. Conversely, in Study 1e it could be argued that individuals who wrote longer texts were genuinely more engaged with the task and more internally motivated. This change in the locus of motivation may therefore be key to understanding when prototypical behaviour arises and when it does not. This is an important area for future research to elucidate when salience leads to prototypical behaviour and when it does not.

Whilst this explanation is only speculatory, Abrams (1994) has also suggested that salience may not always lead directly to prototypical behaviour. According to his social self-regulation model, Abrams (1994) argues that salience leads to a perceptual shift in the way circumstances may be interpreted, but this does not directly translate to behavioural prototypicality. Specifically, Abrams suggests that group members may regulate their behaviour – to be more or less prototypical – based on their public or private awareness of self. This shows synthesis with the strategic aspect of the SIDE model wherein in some instances, the salience of an identity does not lead to identity-prototypical behaviour. For example, Reicher et al.

(1998) observed that students strategically endorsed a less student-prototypical position when they were visible to staff (outgroup members), although were more likely to endorse a student prototypical position when visible to other students as well as staff (both in and outgroup members). However, Abrams' notion of public-awareness points to a generalised hyperawareness of others' opinions (regardless of group membership) as being important to understanding behavioural regulation. In the absence of a clear in- or outgroup audience, Abrams finds that individuals with high public awareness (and lower private awareness) tend to show interpersonally desirable forms of behaviour regardless of the salience of an identity (Abrams, 1992; Greenberg, 1983). For example, Abrams (1992) finds that those with higher public-awareness tend to display lower ingroup bias than those with lower public-awareness. Whilst in the current study we did not measure private or public self-awareness and thus cannot draw robust conclusions about its role regarding prototypical behaviour here, our findings support Abrams' idea that salience may not always lead directly to prototypical behaviour (Abrams, 1994).

Alternatively, it is also possible that the minimum character limit served to remind individuals that they were participants in a study. By insisting that participants continue writing until they have surpassed a word limit, this may have reinstated the salience of a 'participant identity' by reminding the individual that they were there to complete a task. In order to assess the credibility of this argument, we compared whether there was a statistical difference in the responses to the salience manipulation check at the end of Study 1e compared to Study 1f. Using an independent t-test, we found that individuals in Study 1e reported statistically higher responses to the manipulation check following the written task than the participants within this study, $t(249)=1.973$, $p = 0.043$, $d = .268$. This significant difference

provides weight to the argument that the identity salience of individuals within this study may have been compromised by forcing participants to write a certain number of characters. Here, we see the difficulty in assessing subtle prototypical behaviours within controlled experimental conditions.

2.8.4.2 Correlational Analysis

Whilst above we have noted possible explanations for the lower AUCs observed in the first part of the analysis, it was interesting to observe that we were still able to find a small positive correlation between self-reported and behavioural prototypicality in the libertarian condition. One possible explanation for this pertains to understanding how the spread of prototypical behaviour relates to the limitations outlined above. For example, whilst we noted that participants in Study 1f reported lower responses to the salience manipulation check than participants in Study 1e, we also observed a significant positive correlation between participants' responses to the manipulation check and their self-reported prototypicality in both the libertarian condition, $r(89) = .23$, $p = .03$, and the entrepreneur condition, $r(69) = .45$, $p < .001$; individuals who reported being less prototypical of their identity were also more likely to report that they found it difficult to think of themselves as a libertarian or entrepreneur at the end of the study (Haslam, 2004). This therefore suggests that these results may still be suited to measuring the within-group variation in prototypicality. Whereas, classification methods look to model the homogeneity between two groups, correlational models look to understand the variation within each group. Thus, just because the AUCs were lower than in the previous studies, this does not necessarily mean that the variation of the data cannot be accurately modelled or understood at the within-group level.

In the libertarian condition, we observed a significant positive correlation between behavioural prototypicality (in the identity-relevant condition) and self-reported prototypicality. However, at the same time, in the entrepreneur condition we observed a close to significant *negative* correlation between self-reported prototypicality and behavioural prototypicality (in the identity-neutral condition). Thus, this suggests that libertarians who report being more prototypical, do tend write in a more prototypically libertarian style when writing about a topic that is directly relevant to their identity, however entrepreneurs who report being more prototypical write in a more prototypically libertarian style when writing about an identity neutral topic. Notably however, the power of both these correlations was below 0.8 (Faul et al., 2007).

The relationship between self-reported and behavioural prototypicality is therefore quite complex and challenging to understand. One of the key difficulties in interpreting these results pertains to where individual's sense of their own prototypicality comes from. Previous research has indicated that the relationship between group members' perception of their position within a group and their resulting group behaviour is not straightforward. For example, whilst some note that marginal members' loyalty to the group is less reliable over time (Doosje et al., 1999), thus leading marginal members to be less motivated to act for the benefit of the group (Jackson & Saltzstein, 1958) and lower ingroup bias when the distinctiveness of the group is threatened (Jetten et al., 1997), others have indicated that marginal members may demonstrate more loyalty to their group than prototypical members and thus go to greater extremes for their group (Jetten et al., 2002; Jetten et al., 2006; Schmitt & Branscombe, 2001). Similarly paradoxically, at the other end of the scale, researchers have noted that individuals who are

perceived as more prototypical are in turn given greater leeway to behave less prototypically (Giessner & van Knippenberg, 2008; Giessner et al., 2009; Ullrich et al., 2009). Thus, there has been a great deal of research that has thus considered the relationship between perceived-prototypicality (or marginality, Ellemers & Jetten, 2013) and group-related outcomes, although this research has ascertained that the relationship between prototypical perceptions and behaviours is notably complex (Ellemers & Jetten, 2013; Gomez et al., 2014).

In turn then, perhaps it is not surprising that the results of our analysis are not clear cut. Here I will offer some speculatory explanations for the disparate results between self-reported prototypicality and behavioural prototypicality for libertarians and entrepreneurs, although it must be noted that at this time, there is not enough evidence to fully be able to interpret or understand this relationship.

Firstly then, it could logically be argued that in order for an individual to accurately perceive themselves as a prototypical group member, they must also be aware of and have had a significant amount of experience with other group members. Longitudinal research from Ryan and Bogart (2001) demonstrated that individuals have a more accurate perception of their ingroup's prototype the longer that they have been identified with the group. Similarly, as pointed out by Ellemers and Jetten (2013), whilst an individual's identification with a group is largely under their own control, judgments of prototypicality are more influenced by the awareness of other ingroup members. For example, I could identify as an academic but until I have seen or communicated with a variety of other academics, I may find it challenging to gauge how similar I was to the other academics, regardless of my own level of identification with the academic label. Consequently, it could be argued that

the individuals who report higher levels of prototypicality, are those who have interacted with many other ingroup members.

Further, another possible reason for the disparity between the libertarian and entrepreneur results relates to the type of exposure that individuals have had to other group members. For example, if an individual has a lot of interaction with 'extreme' group members, then one may perceive one's position to be less prototypical than is 'true'. This may be especially true for entrepreneurs, where the identity label is such that one is less likely to regularly congregate or interact with other entrepreneurs. In comparison to libertarians, where the identity is by definition politically collective (Deaux et al., 1995), entrepreneurs, by their very nature, are more individually focused (Deaux et al., 1995). That is, one could identify as an entrepreneur having only seen examples of famous entrepreneurs in the media, however, for an individual to identify with a political orientation it is more likely that the individual has discussed politics with other similarly-minded individuals due to the collective nature of the identity. This idea therefore opens the possibility that different types of identity may impact perceptions of prototypicality in different ways. This shows synthesis with the work of Lickel et al. (2000), who find that different types of group are perceived as more entitative than others. Specifically, the researchers note that groups who have more interaction with each other – intimacy groups, task groups and weak social relationships – are perceived as more similar to each other than 'social categories'. Of course, it could be argued that both entrepreneurs and libertarians are social category groups, but nonetheless this research introduces the idea of variability within the perceived similarity of group members which may thus impact prototypical judgments.

To circle back to the argument made previously by Ellemers and Jetten (2013), judging prototypicality requires an understanding of other ingroup members positions, yet behaving prototypically may not rely on comparing oneself to specific others, only to a generally abstract prototype. Thus, in the same way that identification does not require an understanding of other group members, perhaps behavioural prototypicality also does not require direct knowledge of specific others. For this reason then, perhaps it is not surprising that self-reported prototypicality and behavioural prototypicality do not *always* align well.

Finally, this difference between self-reported prototypicality and behavioural prototypicality brings into light the lack of research which has been completed to understand the link between self-perception and group processes (see Postmes & Jetten, 2006). Whilst previously others have noted the need to understand the link between self-perceptions, group perceptions and group memberships - with particular reference paid to situations where one's perception of one's group does not align with others' perceptions (Barreto et al., 2003; Barreto, et al., 2006; Jetten & Hornsey, 2011) – here we suggest that greater attention also needs to be paid to the interactive process between group-perception and behaviour. At present, it is not clear whether individuals behave prototypically because they think of themselves as similar to other group members, or whether they think themselves of themselves as similar because they happen to behave prototypically. In fact, it is entirely plausible that the relationship between self-perception and behaviour is not unidirectional but is in fact reciprocal. Therefore, in this chapter, by providing a possible methodology of assessing prototypical behaviour we can now begin to explore this link in a way that has not previously been possible due to the lack of behavioural measures.

2.8.4.4   Concluding Remarks

In sum, the results provided in this study suggest that prototypical behaviour is impacted heavily by the motivation of the individual as well as small changes to the experimental design. Further, we note that the relationship between self-evaluations of prototypicality and behavioural prototypicality may not be straight forward. This is in synthesis with a substantial body of work which has demonstrated that judgments of low prototypicality may both enhance the desire to demonstrate one's group membership, or alternatively lead to less group-related behaviour (Ellemers & Jetten, 2013). As noted, this similarly complex relationship between perception and behaviour occurs at the other end of the prototypicality scale, where more prototypical members are given more leeway with regards to group normative behaviour (Giessner et al., 2009). Consequently then, whilst this study set out to provide concurrent validity to the behavioural measure of prototypicality, instead the complex results suggest that this relationship may not be clear cut. Thus, current measures of prototypicality may be measuring something conceptually distinct from our behavioural measure of prototypicality. This presents an exciting avenue for future research to explore.

2.9   Chapter Discussion

Taken together, the six studies outlined in this chapter provide fairly robust evidence to support the idea that behavioural prototypicality can be measured in written communication using linguistic style, even if this does not align directly with perceptions of prototypicality. We therefore provide a methodology for measuring behavioural prototypicality which is independent of context. In this chapter, we have shown that when individuals are motivated to behave in line with their salient social identity, we are able to detect an identity-prototypical linguistic style regardless of

content (Studies 1d and 1e), domain (Studies 1c and 1e) and individual factors such as demographics and personality (Studies 1b and 1e). However, concurrent evidence in the form of a correspondence between self-reported prototypicality and behavioural prototypicality measured through linguistic style is currently missing (Study 1f). We outline the significance of these results both theoretically and methodologically below.

2.9.1 Theoretical Conclusions

The studies outlined in this chapter enable us to observe how individuals behave in prototypical ways which are independent of the local group context. Whilst the majority of online research has focused on local prototypicality as derived through specific contexts and interactions, (e.g., Postmes et al., 2000), the present research suggests that individuals have a more enduring sense of what it means to be a member of a particular group and that this impacts behaviour when the identity is salient. Thus, whilst individuals may form relationships with those in their immediate local context and report high identification with particular local groups (Pendry & Salvatore, 2015), the current research suggests that individuals also act in line with their global identity prototypes. More specifically, this finding provides evidence for the self-categorisation theory premise which suggests that individuals do not have to directly communicate with other group members in order to identify with, and categorise themselves as, a particular social identity label (Turner et al., 1987). In this chapter, we saw how individuals who have likely never interacted directly with one another (i.e., those from the Silk Road, Reddit and Prolific participants), still communicate using similar linguistic styles. This suggests that group members have a cognitive group representation which impacts behaviour regardless of the local context. Thus, whilst norms may arise within local contexts,

here we are finding that those with a shared salient social identity communicate in similar ways regardless of domain.

In addition to providing a methodology for assessing behavioural prototypicality, this chapter also drew attention to the disparity between the way individuals may perceive their own prototypicality, and the way in which they behave. This finding was not wholly unexpected given the complex relationship that has previously been observed between prototypical perceptions and resulting group related behaviours and outcomes (Ellemers & Jetten, 2013). However, the results of Study 1f suggest a new avenue for exploring the relation between prototypical self-perceptions and prototypical behaviour. In fact, through using the behavioural methodology outlined in this chapter, it may be possible to draw more robust conclusions as to when individuals' perceptions and behaviours align, and when they do not. One suggestion that was outlined in the discussion of Study 1f, pertained to the significance of the level of interaction that individuals had previously had with other group members. This is an exciting new avenue for research to explore now that we have a behavioural measure of prototypicality.

2.9.2   Methodological remarks

Methodologically speaking, this research also helps us to understand how individuals portray their prototypicality in online environments. More specifically, the dearth of social cues available in computer mediated interactions limits the ways in which individuals can convey this prototypical knowledge. In this research, we have demonstrated that linguistic style provides a suitable medium for studying identity prototypicality due to its ubiquity and also malleability in all communications. In Study 1b where we focussed on intra-individual style shifts, we observed the ease with which individuals adapt their style given the change in social context and the change

in salient social identity. It must be remembered that throughout the studies within this chapter, we have used only nine stylistic features (first-person singular pronouns, first-person plural pronouns, second-person pronouns, negations, long words, affective language, time-based language, future focussed language and tentative language) to demonstrate this shift. This focus on linguistic style to demonstrate social identity prototypicality shows some synthesis with the qualitative research of Postmes et al. (2000).

In their initial study, Postmes et al. (2000) demonstrated how individuals can use the style of their communication to create new local identities. In their research, they analysed the communication style of students communicating in small groups within their university courses. In contrast to the research outlined here, the authors of this study manually coded the communications with labels such as 'humour' or 'insults'. Here however, we demonstrated how automated textual analysis can also provide insight into the way in which individuals communicate and enact their group identities. Further, in contrast to Postmes and colleagues (2000) study, we have illustrated that not only can identity prototypes be induced through locally normative behaviour, but they can also impact behaviour in a more top-down, global manner. It is therefore notable that linguistic style may be used to achieve both prototypicality at the local level as well as prototypicality at the supra-individual, global level.

Further, as outlined in Study 1d and 1e, identity-congruent stylistic differences are detectable even when content is kept constant. In fact, in the final two studies of the chapter, we noted that style may be more important to determining prototypicality when a text is written about a topic that is not considered to have an identity-congruent response. This finding was in line with the research of Koschate et al. (2021) who similarly found that when two social identities (parents and feminists)

were asked to comment on a topic previously rated as neutral to both identities, they also noted that individuals conformed more to their identity-prototypical style than when asked to discuss an identity-relevant topic. This finding is important especially in relation to groups who may be discussing similar topics, however, have different prototypes guiding their behaviour. For example, law enforcement agencies are interested in distinguishing between drug-related forums which are devoted to helping individuals with addictions, and those which are instead masquerading as help but are in fact more concerned with selling drugs illegally. In this instance, both forums may present similarly when considering the content of the forum, but the prototype of the identities on these forums may be considerably different. In this way, stylistic analysis enables us to transcend the content of communication and instead access an idea of who the individual considers themselves to be in a given moment. This is particularly relevant when thinking in terms of criminal or stigmatised identities; an idea further explored in Chapter 4 of this thesis.

In this chapter, we have developed a measure of behavioural prototypicality that allows us to study social identities in naturally occurring online environments. This measure allows researchers to explore social identity hypotheses within real world environments in order to understand more about how social identities impact behaviour and cognition outside of the laboratory. In Chapter 4 of this thesis, we indicate how this measure could be used to understand how real-world events may lead to changes in a group's prototype. Overall, having a behavioural measure of prototypicality enables us to both extend the theoretical concept of prototypicality (i.e., by looking at processes of development) as well as to use behaviour to infer social identity characteristics (i.e., to understand how social identities change in relation to key events).

2.9.3 Limitations

Nonetheless, despite the potential of this methodology for furthering our understanding of prototypes and behavioural prototypicality, there are some limitations. Firstly, it must be noted that it is difficult to disentangle globally prototypical behaviour, and locally prototypical behaviour. Thus, in our initial classifier (Study 1a) we identified nine variables using a bottom-up approach that appeared to be informative at distinguishing between identities. However, at this stage, we were unsure whether these nine variables were just distinguishing between local group norms or possible demographic differences in individuals from both groups. In the following studies, we ruled out these possible explanations, demonstrating the robustness of the classifier in a variety of conditions. In fact, the success of the classifier across the first five studies pointed to the idea that these nine variables were particularly informative for distinguishing between group prototypes. However, it must be said that whether these nine variables are the optimal features to determine prototypicality remains unknown.

An alternative approach to determine important prototypical variables may instead be to predict top-down which variables are associated with identity prototypes. For example, we could use knowledge about the identity prototypes to predict which linguistic features would be most informative. However, a primary criticism of this approach, as noted by peer reviewers, is the subjectivity of the researcher's judgment. Whilst one can coherently argue that the libertarian prototype is reportedly intellectual and politically passionate (Iyer et al., 2012) whereas the entrepreneur prototype tends to be more direct and pragmatic (Gibson, 2000), it is exceedingly difficult to predict how this will convert into linguistic style features. Whilst there may be some features such as long words and emotion words that

would presumably reflect the intellectual passion reflected in the libertarian prototype

for example, defending more nuanced positions on the use of personal pronouns or

negations remains challenging. However, as experimental linguistic style analysis

develops, this area may be an interesting avenue for future research.

Further, previous research has suggested that it is not necessarily individual

linguistic features that enable us to predict psychological characteristics, it is instead

the combination of features (Boyd, 2017). In this way, there may be a variety of

different linguistic patterns that are captured using only a small number of linguistic

style features. Therefore, whilst a libertarian prototype defined by intellectuality may

lead libertarians to regularly communicate in an intellectual style with long, complex

sentences, by using machine learning techniques we can also understand their

linguistic style when they do not adhere to this particular pattern. In other words,

libertarians may also have a distinct non-intellectual pattern of communication that is

just as prototypical which relates to, for example, their utilitarian directness (Iyer et

al., 2012). In this way, by using more complex statistical models we are able to

capture the different traits and attributes that may be accentuated by social identities

at different points in time. Therefore, the higher level of dimensionality by which a

group's linguistic fingerprint can be classified, corresponds to the high number of

dimensions on which a group may define itself. Similar arguments regarding

individual dimensions and traits have been made in favour of using machine learning

within the realm of authorship attribution (Boyd & Pennebaker, 2017).

Nonetheless, there are important criticisms regarding the use of non-linear,

less transparent statistical processes within psychological research. Boyd (2017)

noted that being able to directly understand what a statistical process is doing is

integral to understanding the result of the analysis. Specifically, Boyd mentions the

trade-off between using complex models to predict future outcomes and using transparent models in order to ensure that one can directly explain and understand psychological phenomena. Whilst I agree with Boyd's acknowledgement that this is an important trade-off, I would contend that by repeatedly testing a less transparent machine learning model on new and different datasets, this can enable researchers to understand precisely what their model can and cannot do. For example, in this chapter we have used various different study designs and test sets in order to validate our interpretations. Thus, whilst Boyd's point stands regarding our inability to draw robust explanatory conclusions from our first study alone – at this stage we cannot be certain what the classifier is detecting – by testing the classifier on five different test sets, we can gain a greater understanding as to what the classifier is measuring and also what its limitations are. Thus, in this way, the lack of transparency of the machine learning model is explored through further testing, as opposed to by reverting to more simplistic yet transparent statistical processes. Arguably, this process of testing a 'black-box' model on a variety of different datasets to see how it performs is analogous to the approach taken by cognitive psychologists in trying to understand the capabilities of the human brain. With the human brain as a kind of 'black-box' model, we can understand cognitive processes by using a variety of experimental designs to learn more about the brain's performance. In synthesis, by repeatedly testing our classifier on diverse new test sets we are able to draw increasingly robust conclusions as to the model's capabilities.

In sum, this chapter has illustrated how we can use naturally-occurring online linguistic data to study behavioural prototypicality in real world environments. We have developed a methodological approach for measuring global prototypicality in real-world environments using naturally-occurring data. Further, we have also

demonstrated how we can combine numerous study designs and complex machine learning models in a way that does not lead to a compromise in psychological explainability.

3 Studying the Link Between Social Influence and Prototypical Behaviour Online

In the previous chapter, we built a classifier that was able to detect identity-prototypical behaviour regardless of content (Study 1d and 1e), domain (Study 1c and 1e) and individual factors such as demographics and personality (Study 1b and 1e). In this chapter, we use this model to measure the prototypicality of actors operating in online forums, and to assess the relationship between prototypicality and influence online. Specifically, we are interested in understanding whether individuals who display more globally-prototypical behaviour may receive more attention from or have more influence on other group members. This chapter therefore focuses on the applied value of this behavioural approach with regards to studying influential actors online. This may be useful particularly for studying criminal identities, but also the link between global prototypicality and influence more generally. From a theoretical perspective, it is important to understand whether global prototypicality is still related to social influence, or whether instead it is only the local/inductive aspects of prototypicality that make it predictive in a particular social context (e.g., Johnson et al., 2015).

3.1 Study 2a: Applying the Model Outlined in Chapter 2

3.1.1 Literature Overview and Hypotheses

3.1.1.1 Social Psychological Study of Social Influence

In Study 1f, the results of our analysis suggested that individuals may not have an accurate idea of their own level of prototypicality, however, we know from previous research that group members tend to agree on the prototypicality of other members of their group (Giessner et al., 2009). Whilst it must be stated that the research of Giessner et al. (2009; also van Knippenberg et al., 1994), involves

comparing the vignettes of a highly prototypical leader to a non-prototypical leader, the authors nonetheless find that ingroup members are correctly able to identify the highly-prototypical leader. What is unclear, however, is whether highly-prototypical individuals are recognised as such by group members in real-word contexts. The vignettes used in Giessner et al.'s (2009) study may have been successful at outlining traits that ingroup members view as prototypical, but whether ingroup members would have selectively attended to these attributes in a real-world setting remains unstudied. This chapter therefore aims to understand whether other group members appear to recognise the prototypical style of individuals who we have behaviourally classified as prototypical. This will provide concurrent validity to our classifier if we can identify the link between prototypical behaviour and the responses of others to that behaviour.

Using naturally occurring data, one way to understand whether group members are recognising the prototypical behaviour of others in naturally occurring contexts is to ascertain whether they are more influenced by prototypical individuals compared to non-prototypical individuals. Offline, research has indicated that leaders derive influence as other group members perceive them to embody and represent what is group prototypical (Hogg, 2001; Hogg et al., 2012; Turner, 1991; van Knippenberg & Hogg, 2003) and thus it makes sense to assume that a similar relationship will be observed between prototypicality and influence online. Research has demonstrated that group members are more trusting of those perceived as prototypical due to the assumption that more prototypical leaders have greater motivation to pursue the interests of the group (Giessner & van Knippenberg, 2008; van Knippenberg & van Knippenberg, 2005). Additionally, leader emergence research has suggested that in groups with no defined leader, individuals who are

perceived as more prototypical have a greater likelihood of emerging as leaders (Fielding & Hogg, 1997).

However, it must be noted that in the wealth of literature that has sought to understand the link between the perception of prototypicality and leadership effectiveness (for meta-analyses see Barreto & Hogg, 2017; Steffens et al., 2021), perceptions of prototypicality have been measured through self-report methodologies. In some studies, participants are asked to rate how prototypical they perceive a group leader to be and then to rate leadership qualities such as effectiveness and trust (Hirst et al., 2009; van Knippenberg et al., 2008). In others, individuals are simply told how prototypical a leader is and then asked to rate their leadership qualities (Platow & van Knippenberg, 2001). Whilst these studies demonstrate the link between perceived prototypicality and leadership abilities, it is not clear exactly how an individual comes to their perception of another's prototypicality.

Due to the richness of offline interactions, there could be any number of social cues that influence perceptions of prototypicality in offline situations (e.g., clothing, gait, age, language, appearance and so on). In contrast, social cues in online interactions are much more limited. We have demonstrated in the previous chapter that individuals use linguistic style to convey their prototypicality, but here we assess whether linguistic style may be used in naturally-occurring environments to evaluate others' prototypicality. Without having direct access to the perceptions of individuals, we use behavioural measures as a proxy for recognition of prototypicality.

Previously, research has indicated that linguistic style plays a key role in general impression formation online. For example, Larrimore et al. (2011) noted that in online Peer-to-Peer lending environments, use of concrete language (articles and

quantifiers) predicted the likelihood of receiving a loan. Additionally, Toma and Angelo (2015) found that in online medical communities, forum-users were more likely to be perceived as experts if they used more words in their posts, fewer I-pronouns and anxiety-related words as well as more long words and negations. Finally, Hallier-Willi et al. (2014) also stress that communication style plays an integral role in the formation of corporate impressions in online communities; suggesting that corporate online community members form more favourable impressions of companies who are more direct in their communication.

Therefore, combining the finding that individuals communicate in group prototypical ways, and that linguistic style plays a key role in impression formation, we predict that prototypical group members are recognised by other group members from their linguistic style. Further, we hypothesize that there will be a positive relationship between linguistic prototypicality and influence within the forum, in line with the findings from offline research (Baretto & Hogg, 2017) and referent informational influence theory (Turner et al., 1987).

3.1.1.2 Measures of Online Influence from a Computer Science Perspective

Within the computer sciences literature, influence research has focused on understanding how connections and interactions between individuals enable us to gauge which users are most influential (see Peng et al., 2018 for a review); this approach falls into the domain of social network analysis. Metrics included in the computation of influence within this literature are: the number of interactions between users (Yang et al., 2010), novelty of information that a user provides to the network (Song et al., 2007), density of connections between individuals (Zhou & Liu, 2015), and topical expertise (Munger & Zhao, 2015). A large proportion of this literature is

also devoted to understanding how the structure and topology of social networks impacts diffusion of information throughout the network (e.g., Wang et al., 2013).

However, social network analysis identifies influential individuals from a purely mathematical perspective. Through analysis of the structure of networks and the connections between users, the network-based approach is able to identify key nodes that may impact the diffusion of information through a network (see Razaque et al., 2019 for a review). However, measuring influence using social network metrics is quite tautologous; individuals are labelled as influential if they, for example, have a higher number of connections than others within the network, however this measure of influence does not allow us to understand *why* particular members have more connections for example. By taking the sociopsychological reality of individuals into account, we suggest that we are able to gain greater understanding into the social dynamics of online groups, and thus extend beyond labelling individuals as influential because of their influence alone.

Thus, this research aims to assess whether the relationship between group prototypicality and social influence transfers to the online realm. Further, we also illustrate the value of psychological theory in domains primarily focussed on using advanced data-driven mathematical techniques to model social processes.

3.1.1.3 Hypotheses

For this study, we will quantify influence using basic network-theory measures in order to understand whether this idea is worth further exploration. Previous research looking to identify leaders within online communities has suggested that the number of responses that an individual is able to generate is a key indicator of their influence (or at least influential potential) within a forum-based environment (Huffaker, 2010). Based on this quantification of influence, we hypothesize that

prototypical individuals will be able to generate more responses from others compared to those who are less prototypical. Specifically, we hypothesize that there will be a positive correlation between the prototypicality of an individual and the number of responses their posts generate (H1).

3.1.2 Method

3.1.2.1 Data Collection

We used the Reddit classifier outlined in the previous chapter to generate prototypicality scores for each individual and then correlated this with measures of influence online.

We used the nine-feature classifier from Study 1a to assess the prototypicality of previously unseen Reddit data from March 2019. For the March dataset, we collected the title and text of each post, as well as the author, URL and number of comments. For the comments, we collected the author, text, comment id, parent post id and link id. The link id referred to the id of the top-level post and the parent post id referred to a comment within a thread that someone may be responding to.

The total sample consisted of $N = 172,966$ submissions, with $n = 33,679$ from the entrepreneur subreddit, and $n = 139,287$ from the libertarian subreddit. For the linguistic analysis, we applied the same data pre-processing procedure outlined in Study 1a (see Table 2.1). This left us with a sample size of 23,720 posts from the libertarian forum (4,917 authors), and 7,479 posts in the entrepreneur forum (3,342 authors). The mean number of words per contribution was 217 ($Med = 136$, $SD = 264$) in the libertarian forum, and 151 ($Med = 101$, $SD = 186$) in the entrepreneur forum. Further details of the data preparation process are displayed in Table 3.1 below.

Table 3.1

Detail of Data Preparation Process (Study 2a)

Omitted Data	Libertarian Contributions		Entrepreneur Submissions	
	Removed	Remaining	Removed	Remaining
Initial		139,287		33,679
URL removal	3,916	135,371	0	33,679
Deleted/removed submissions	2,325	133,046	3,561	30,118
Deleted/removed authors	21,211	111,835	4,161	25,957
Bot removal in text	3605	108,230	2,069	23,888
Bot removal in author	328	107,904	49	23,839
Submissions > 50 words	84,634	23,720	16,360	7,479
Total posts		*461*		*1,304*
Total comments		*23,259*		*6,175*
Total		**23,720**		**7,479**
Total authors		**4,917**		**3,342**

3.1.2.2 Procedure

We determined the prototypicality of each individual in the forum using the output of the Extra Trees classifier (as in Study 1f). Having determined a prototypicality score for each forum post, we calculated the average prototypicality score for each author per forum.

In order to determine how many responses an individual generates, we used the number of comments per post metric provided by Google BiqQuery, as well as

centrality measures. To calculate the centrality measures, we derived a network of connections based on who responded to who. In this network, each node represented a user in the forum and an edge between two nodes represented an interaction. We instantiated an edge between two nodes if one user had replied to another user's post or comment. More specifically, we used a directed network approach which meant that if User A commented on User B's post, an edge would exist from A to B, but not from B to A. We decided a directed network approach was more appropriate than an undirected network approach (where the same edges would exist between A and B as between B and A regardless of sender and receiver), in order to distinguish between individuals who have more connections because they receive more incoming comments, and those who have more connections because they comment more on other's posts.

By constructing this network of connected users, we can determine which nodes are most central in the network. An individual's indegree centrality is the number of replies a user receives on all posts and comments. Whilst it is common to use indegree centrality alone as a measure of response generation (Huffaker, 2010) for the current research, we use an individual's indegree centrality divided by their total number of contributions to understand how many responses an individual generates whilst taking into consideration their own posting behaviour (Mislove, 2009); we refer to this as a 'centrality ratio'. For example, a person who has posted once and received 4 responses would have a centrality ratio of 4, whereas a person who has posted 4 times and received 1 response would have a centrality ratio of 0.25. Based on this metric, the higher the centrality ratio the greater the number of responses an individual is able to generate per contribution.

The primary difference between the centrality ratio and number of comments is that the number of comments per post focuses on an individual's ability to generate responses based only on their posts, whereas centrality measures consider whether an individual's comments are also likely to generate responses. In forums with longer discussions, indegree centrality may prove more important in understanding whether an individual's comments are likely to generate further responses. On the other hand, in forums with more of a question-answer format, the number of comments variable may be sufficient for understanding response generation.

To generate centrality measures we used popular software package UCINET (Borgatti et al., 2002). UCINET generates centrality scores by constructing a network of connections and quantifying these connections. Finally, we calculated the centrality ratio by dividing each author's indegree centrality by their total contributions.

3.1.3 Results

To assess the relationship between prototypicality and influence, we conducted several Spearman's rank correlations. First, we reverse coded the libertarian prototypicality score as scores closer to zero unintuitively referred to higher libertarian prototypicality.

We then correlated prototypicality scores with centrality ratios and number of comments per post. Using the prototypicality score determined from the nine-feature classifier, in the libertarian forum a statistically significant relationship was observed between linguistic prototypicality and centrality ratio, $r_s(4915) = .07$, $p < .001$. The

relationship between prototypicality and number of comments was in the predicted direction, although did not reach significance, $r_s(357) = .10$, $p = .061$.[2]

In the entrepreneur forum, we found no relationship between centrality ratio and prototypicality, $r_s(3340) = -.006$, $p = .744$, however we did find that prototypical individuals received more comments on their posts, $r_s(1096) = .11$, $p < .001$.

To explore the difference in significance between prototypicality and centrality ratio and number of comments, we conducted an independent samples t-test to ascertain whether there was a difference in the number of comments per post between the two forums. We found that the number of comments on libertarian threads ($M = 22.5$, $SD = 48.2$) is significantly higher than those on entrepreneur threads (M = 13.0, $SD = 33.3$), $t(1763) = 5.79$, $p < .001$.[3]

3.1.4   Discussion

The results begin to indicate that individuals who are more linguistically prototypical tend to generate more responses from other group members, although this effect size is markedly small.

The results showed that prototypical individuals in the libertarian forum tended to have a higher centrality ratio than those who were less prototypical; relative to the number of contributions they made, prototypical individuals received more responses. Whilst there was no relationship between prototypicality and the centrality ratio in the entrepreneur forum, there was a positive relationship between prototypicality and the number of comments generated on posts. Whereas the centrality ratio measures how many responses an individual receives on all

---

[2] The difference in degrees of freedom is due to the fact that we are only considering authors who have made posts here, whereas the centrality ratio correlation included all authors (including those who only commented on other's posts).

[3] The degrees of freedom here respond to the total number of posts as opposed to authors.

contributions (including comments), the number of comments variable consists only of the number of responses an individual receives on their posts. We therefore explored this difference in finding by comparing the average length of discussion thread in each forum and found that the libertarian forum had longer threads (more comments) than the entrepreneur forum. As a result, it can be argued that the centrality ratio becomes a more informative measure of response generation in forums with greater discussion such as the libertarian forum, whereas the number of comments variable is an apt measure of response generation in forums with more of a question-answer format. This is because in forums with more of a question-answer format, the comments (answers) are less likely to be aimed at generating further responses. The findings outlined here therefore make sense in light of the differing forum structures.

However, this brings us to acknowledge the limitation in using response generation as a measure of social influence within a forum. From a social identity perspective, influential individuals are those who are able to define the group's identity as well as the normative attributes and behaviours associated with it (Hogg et al., 2012). Unfortunately, in online forums we do not have direct access to the perceptions of individuals who may have been influenced by a post or person yet decide not to respond. Based on this limitation, we can only consider the explicit responses in order to understand the reaction of other group members. In this study however, we chose to measure influence via response generation without considering the content of the responses that were generated. This is in line with much of the social network-based approaches to the study of influence and gives us an approximate understanding of whether identity prototypicality as quantified using linguistic style is an area worth further exploration. Based on the small but significant

results within this study, future work should examine this relationship with more robust and psychologically motivated measures of influence. One such measure of influence may be to content analyse the posts (either qualitatively or computationally) in order to assess the level of agreement that individuals have with prototypical posters. This would help us to understand whether individuals were receiving more comments on their posts through agreement, or whether the higher number of comments actually relates to controversy.

Another key limitation of this study pertains to the training data that were used to generate prototypicality measures. In the previous chapter, we argued for the distinction between local group prototypicality, and context-independent prototypicality. However, in this study, we trained our classifier on the original Reddit data used in Chapter 2 and then applied the model to data from the same platform. That is, whilst the classifier could tell us how prototypical the Reddit posts were in light of previous Reddit posts from the same forums, this methodological design means that we cannot distinguish between content-independent prototypicality, and locally-derived prototypicality. This methodological decision was undertaken in line with the previous studies and aimed to show the concurrent validity of the Reddit classifier specifically. However, in choosing to use the same source of data in both training and application, it is not possible to ascertain whether we are finding a relationship between local prototypicality and influence or whether this relationship is between global prototypicality and influence.

One way to address this limitation is to repeat the analysis whilst training the classifier on a separate dataset. By training on a different set of data, this allows us to transcend any specific forum norms. Therefore, an individual who is determined as highly prototypical by a forum-independent classifier, can be said to be showing a

more globally prototypical style. Whilst this methodology is not perfect – we will still be capturing the platform specific forum norms of the training data – we contend that it will give us a better idea of whether the relationship between prototypicality and influence as uncovered here can be explained solely by locally normative behaviour.

3.2   Study 2b: Studying the Impact of Global Prototypicality on Influence in Forums

In the next study, we sought to control for the impact of a Reddit-specific forum style by using a classifier trained on another platform. In Chapter 2, we demonstrated that a Reddit-trained classifier was able to detect an identity-prototypical style across a wide range of contexts, however, this classifier is likely also picking up Reddit-specific behaviour. Therefore, in order to directly test the hypothesis that individuals who are globally prototypical are more influential than others, we have decided to train a new classifier in order to control for the influence of forum norms on our previous result.

However, it must be said now that a classifier trained on another platform is likely also to be picking up platform-specific norms of the independent platform. Unfortunately, this is very challenging factor to mitigate or control for; all forums will have their own local norms. However, this analysis attempts to understand whether a relationship between prototypicality and influence can still be observed *despite* the forum-specific norms of the training platform. Whilst we do not expect the relationship between influence and prototypicality to be strong – due to the inherent error in the 'global' prototypicality measure – we feel that this is the best way that we can attempt to demonstrate that the findings of Study 2a are not merely local accommodation norms.

This distinction is an important one, as it forms the argument for the entire thesis. As we have argued elsewhere, various studies have shown a link between local prototypicality and one's ability to influence others. For example, in their 2015 study, Johnson et al. note a link between how similar an individual's language is to the local group collective, and how influential others perceive them to be.

Therefore, as this research is directly concerned with understanding context-independent prototypicality which is not influenced by local group norms, Study 2b will use a classifier trained on data from the Silk Road in order to assess whether individuals are influential at the global-prototypicality level, and not just the local level. Whilst, we have demonstrated in Study 2a that individuals may also be influential at a more local level, Study 2b aims to control directly for the influence of local group prototypicality.

3.2.1   Hypotheses

In the current study then, we use data from the Silk Road forums (outlined in Study 1c) to train our classifier. We then apply this classifier to the same Reddit data used in the previous study (Study 2a). By performing this analysis, we are able to verify whether individuals are influenced by a context-independent prototype, or whether the findings of Study 2a are forum-specific. We hypothesize, in line with the findings of Study 2a, that there will be an association between the prototypicality of an individual's communication, and the responses that they receive from others in their forums. Further, we suggest that libertarian prototypicality will correlate with the total centrality ratio due to the discussion-based nature of the forum, and that entrepreneur prototypicality will correlate with the number of comments received on each post.

3.2.2 Method

3.2.2.1 Data

In order to train our classifier, we used the data from the Silk Road outlined in Chapter 2, Study 1c. Previously, we have ascertained that the individuals posting in these data have comparable social identities to the individuals posting on the Reddit forum.

As before, the initial sample of Silk Road data consisted of $N = 20,836$ posts, with 10,780 originating in the libertarian forum titled 'Economics, Philosophy and Justice' and 10,056 originating in the entrepreneur forum titled the 'Vendor Roundtable'. In line with the previous studies, we removed the posts with fewer than 50 words; this left $N = 10,494$ posts with $n = 4,746$ posts from 553 users in the Vendor Roundtable forum and $n = 5,748$ posts from 1,131 users in the Philosophy, Economics and Justice forum.

3.2.2.2 Procedure

In order to ascertain which features to include in the model, we first trained a Silk Road classifier using all 41 stylistic variables and all 10,494 posts. After visually inspecting the feature importances (see Figure 3.1 below), we decided to use the top nine stylistic variables as these held greater importance in classifying the entrepreneur and the libertarian identities. Further, due to the previous success in using nine linguistic indicators in Chapter 2, we decided that nine variables provides a robust number of indicators to capture the differences between the identities. These are: numbers, six letter words, first-person singular pronouns such as 'I', insight-based language, affective language, articles, motion-based language, past tense focus and time. We can see that four of these variables (first-person singular pronouns, six letter words, affective language and time) overlap with the top nine

variables that we had been using to train the Reddit classifier. In turn, it could be argued that the overlap between these two forums' top feature importances speaks to the 'global'-ness of the prototypicality measure already. Nonetheless, it was important to derive variables from the Silk Road classifier that were not going to be influenced by possible local forum norms from the Reddit forums.

Figure 3.1

Feature Importances of Classifier Trained on Silk Road Data.
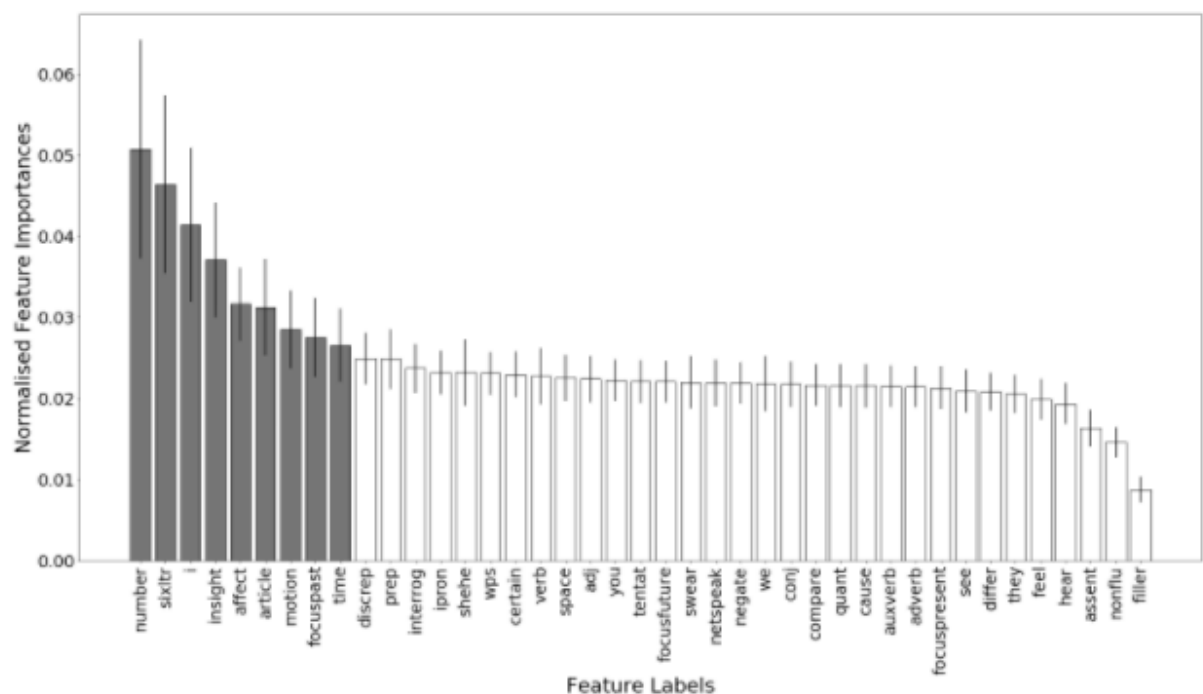
Note. Error bars correspond to standard deviations.

The first step was to ascertain whether the nine Silk Road variables were at all able to predict identity-prototypical behaviour at the group level in the Reddit forum. By training the classifier using the nine stylistic variables and testing the classifier on the March 2019 Reddit posts detailed in Study 2a, our classifier achieved an AUC of .66 (using equal sample sizes of 7,479 posts per forum). Based on this result, it was clear that the classifier was suitably able to detect an identity-

prototypical style on the Reddit platform, and thus it would be appropriate for calculating prototypicality scores at the individual level.

We then trained our classifier on the Silk Road data, using the nine stylistic variables that we had predetermined as being important to conveying prototypicality between our two social identities. In line with the previous study in this chapter, we calculated a prototypicality score for each post within the forum and then calculated the mean prototypicality score per author. We used the influence metrics – the centrality ratio and the number of comments per post – which had been generated in Study 2a to assess the relationship between influence and prototypicality.

3.2.3   Results

After reverse coding the libertarian scores to ensure all prototypicality metrics were coded the same way, we repeated the four Spearman's rank correlations outlined in Study 2a.

In the libertarian forum, we noted a very small yet significant correlation between prototypicality and the centrality ratio, $r_s(4915) = .029$, $p = .040$. The relationship between prototypicality and number of comments was marginally significant, $r_s(357) = .10$, $p = .057$.

In the entrepreneur forum, we found no relationship between the centrality ratio and prototypicality, $r_s(3340) = -.025$, $p = .145$, however, again we found a small positive relationship between how prototypical an individual is and the number of comments received on their posts, $r_s(1096) = .099$, $p = .001$.

3.2.4   Discussion

In this second analysis, we sought to understand whether the relationship between prototypicality and influence could be explained by more than just forum-specific local norms or a local level of prototypicality. By training a new classifier on

data taken from a different platform and then applying this model on Reddit data, we observed a very small association between prototypicality and response generation. These small associations were found between libertarian prototypicality and the centrality ratio, in line with the results of Study 2a, as well as between entrepreneur prototypicality and the number of comments received on posts, also in line with Study 2a. It must be noted however that the effect sizes observed within this study were small, verging on trivial.

Due to the effect sizes noted, it is difficult to convincingly conclude that there is a positive association between global, behavioural prototypicality and influence. It appears instead that this research topic warrants further investigation, ideally with more reliable measures of both influence and global prototypicality. In this study, we demonstrated the inherent difficulty in measuring behavioural prototypicality, especially when using a data-driven bottom-up approach. Perhaps in future, the optimal method for building a context-independent classifier may be to use data from a variety of different sources, as opposed to only one independent source as attempted in this study. By using a larger dataset to train a classifier on, it is more likely that the variability in prototypical behaviour, as impacted by topic or platform for example, will be ironed out.

One other approach to building a classifier is to use a top-down approach which relies on knowledge of the identity prototype to predict which features will be most important at distinguishing between identities. In this way, the classifier is able to focus on the features that are directly relevant to the portrayal of an identity, whilst limiting the impact of local forum norms as captured in bottom-up approaches. However, as outlined in Chapter 2, the difficulty associated with this approach lies in being able to correctly justify the choices one makes as to the prototypical features.

Whilst we may have access to traits or characteristics of entrepreneurs and libertarians from previous research, how these directly relate to their communication style is less clear. For example, whilst we know that entrepreneurs define their entrepreneur identities with reference to strategic orientation, commitment to opportunity and a control of resources (Brown et al., 2001), how this translates to the style of their communication is less clear and more challenging to justify.

Another possible improvement to this study methodology concerns how influence is measured and conceptualised. In order to further this line of inquiry, one avenue may be to directly analyse the communications of others towards individuals that are more prototypical. This could be completed both qualitatively, by directly assessing how individuals respond to prototypical individuals, but also computationally using Linguistic Style Matching techniques (LSM; Niederhoffer & Pennebaker, 2002). Previous research has highlighted that individuals with a low reputation in forums accommodate towards those with higher reputations (Jones et al., 2014). In Jones et al.'s study, reputation is a forum-specific variable and encapsulates the number of upvotes, length in community and number of posts a user has made. By combining LSM techniques with the measure of prototypicality outlined in this research, we could further test whether influence when quantified as accommodation, is related to behavioural prototypicality.

3.3   Chapter Discussion

The findings from this chapter suggest that there may be a link between behavioural prototypicality and influence, although the small effect sizes in both studies mean that this hypothesis requires further exploration. Whilst a plethora of research has utilised self-report survey methodology to study prototypicality and social influence in offline contexts, the primary approach to studying social influence

in online environments has focused on the structure and relationships of social groups and the position of members within these groups rather than socio-cognitive processes. In the present research therefore, we attempted to combine sociopsychological theory and computational methodologies to demonstrate the value in using behavioural data to measure socio-cognitive processes. However, whilst we found statistically significant associations in both studies, the small effect sizes demonstrate the challenge of using large naturally-occurring datasets to measure socio-cognitive processes indirectly.

Despite the small effect sizes, this chapter made the argument that through using indirect measures of socio-cognitive processes, this can enable us to better understand the characteristics that make individuals influential in changing social contexts. It provides us with a means to describe influential individuals online without the circularity of describing them as influential because of their influence. For example, influential users or bloggers online have been defined by their ability to generate responses (Huffaker, 2010), the immediacy with which others respond to them (Yang et al., 2010), content similarity between responses (Huffaker, 2010), contribution of novel content (Song et al., 2007) or the number times the user's message is shared (Cha et al., 2010). However, these metrics serve only to describe the behaviour of influential users and their followership rather than provide an explanation of their influence based on a testable model or theory.

Interestingly, it is possible to draw parallels between the network approach to influence with how leadership and influence were originally studied within psychology. Both early leadership researchers and current network analysts approach influence research through focusing on the traits of the individuals (Judge et al., 2002) and social exchange relationships between leaders and followers

(Graen & Uhl-Bien,1995; Hollander, 1980). Where social-exchange theorists study the interpersonal relationship between team members and leaders (Hollander, 1992), network analysts similarly study how influencers are responded to by those around them through shares, likes, retweets and comments. Thus, in a similar way to how psychologists came to note the importance of psychological group processes underlying influence, we propose that this understanding of group psychology is needed to extend the current approaches to social influence online.

As noted earlier, future research in this area would benefit from developing models that are less likely to be influenced strongly by the domain of communication. As we outlined in the discussion of Study 2b, there is a difficulty in using data-driven classifiers to access a global prototypicality. However, one further approach that may help to achieve the aim of an optimised global prototypicality measure, would be to use combined data from many forums who share the social identity of interest in order to minimise the impact of each individual forum's local norms. By training a model on an amalgamation of different forums, this would likely serve to increase the variability in the training set but simultaneously lead to the development of a classifier that was less impacted by local norms. In turn, this classifier may be better at capturing the global identity prototypicality of forum users at the individual level.

In conclusion, this chapter has demonstrated both the limitations and strengths of using data-driven behavioural classifiers to test social psychological theory online. The effect sizes of both studies in this chapter were small, albeit were statistically significant in the predicted directions pointing to the need for more research in this area. Further, we have argued that indirect measures of sociopsychological processes allow us to escape the circularity of describing individuals as influential because of their influence. By developing this methodology

by using more robust measures of both influence and prototypicality, this may allow us to understand at a theoretical level why some individuals have more influence than others.

4   Developing a Typology of Social Identities based on Group Prototypical Behaviour

So far in this thesis, we have explored how linguistic style can be used to measure both group-level prototypical behaviour on aggregate (Chapter 2) and individual-level prototypicality (Chapter 3). In this chapter, we look to extend this by exploring how linguistic style analysis can be used to create a typology of social identities based on the similarities in identity-prototypical behaviour. Previously, researchers have developed group typologies that use self-reported data or structural group factors (i.e., face-to-face small groups or large categorical groups) which have proved valuable in demonstrating that not all group memberships are psychologically equivalent to their members. Specifically, different types of group membership have been found to serve different functions for individuals (Deaux et al., 1999) as well as differentially impacting group processes such as identification and perceived entitativity (Lickel et al., 2000). However, whilst the importance of social identity content has been recognised in individual research papers (e.g., Livingstone & Haslam, 2008), no group typologies have been developed which group together identities based on group prototypical attributes or the content of identities.

In this chapter, we demonstrate the value of a prototypical behaviour-based typology which categorises groups based on how they behave in real world contexts. Using automated text-analysis, this approach allows us to understand the prototypical behaviour of groups outside of laboratory conditions and can therefore be used to develop a better understanding of elusive naturally occurring groups such as those involved in criminal activity. In Study 3a we develop a typology of identities based on similarities in naturally occurring linguistic style data. In Study 3b, we use within-individual data to demonstrate that this typology is not impacted by individual-

level factors such as demographics and personality traits. In Study 3c, we use the content of forum posts to understand more about how these stylistic indicators relate to the perceived values and emotions of the groups included within our analysis. In Study 3d, we add more identities to our analysis in order to understand the impact of platform on our behavioural typology. We also include the data from online groups which are argued to have no pre-existing sense of identity (e.g., Smith et al., 2015a), in order to test the idea that even newly-formed groups can be categorised using our behavioural typology. And finally, in Study 3e, we demonstrate that global-identity prototypical behaviour can change and evolve over time; whilst the enduring nature of the global-prototype may be less fluid than the locally-induced and context-dependent prototype, this does not mean that it does not change as the wider societal structure changes. In this way, globally prototypical behaviour can be said to be independent to the local context, but not to changes at the macro or societal level (Pehrson & Reicher, 2012). In sum, this chapter serves to illustrate the value in using naturally-occurring behavioural data to explore similarities in group prototypical behaviour. We further demonstrate that this approach can enable us to study the evolution of real-world groups over long time frames, thus providing an exciting application for this novel psychological methodology.

4.1   Study 3a: Developing a Behavioural Typology of Identities

Various researchers have argued that group typologies are important for understanding and explaining social group processes due to the acknowledgement that not all group memberships are perceived in the same way by individuals (Brown, 2000; Brown & Williams, 1984; Deaux, 2000; Prentice et al., 1994). For example, Deaux et al. (1999) provide evidence to suggest that different groups have different social functions for their members: for members of religious groups, their religious

identity functions to enhance self-esteem, self-insight and cohesion, whereas members of sports teams tend to stress the intergroup comparison, social interaction and group-based esteem functions of their social identities. Similarly, Lickel et al. (2000) demonstrate that individuals perceive varying levels of entitativity between different types of group. They find that participants' perceptions of ingroup entitativity are positively correlated with how strongly the participants valued their membership within said group, as well as the perceived size of the group. It has therefore been argued that based on the differing meanings, perceptions and functions that groups hold for their members, it is an 'important step for [Social Identity Theory] to [..] incorporate these central dimensions of group diversity and no longer to assume that a group is a group is a group as far as key social psychological mechanisms are concerned' (Brown, 2000, p.761).

This idea that there are a variety of different types of group to which social identity researchers should pay heed was keenly discussed at the turn of the century. Research attempted to distinguish between similar groups based on stereotypical similarity (Deaux et al., 1995), perceived entitativity (Lickel et al., 2000), the self-reported social motivations and functions of the group membership (Deaux et al., 1999), social identity components such as individualism and collectivism (Luhtanen & Crocker, 1992; Prentice et al., 1994), social roles and categories (Thoits, 1992), ascribed or achieved identities (Baumeister, 1986; Giddens, 1991; Nagel, 1995) as well as identification strength (Branscombe & Wann, 1994; Doosje et al., 1995, Ellemers et al., 1997; Spears et al., 1997). These various different streams of research all pointed towards the idea that groups are not psychologically equivalent for their members and that these variations in group types and classifications can lead to differential outcomes of group-related processes.

Specifically, it was argued that 'such variations are crucial when the investigator moves into the realm of naturally-occurring identity categories' (Deaux, 2000, p.4). However, despite this acknowledgment of the importance of categorising different types of identity in naturally-occurring settings, it must be noted that no such attempt has been made towards understanding identity categories using only naturally occurring data; all existing typologies still rely on self-reported data.

One of the most well-known pieces of research which sought to create a typology of groups was that of Deaux et al. (1995). Deaux and colleagues asked participants to sort 64 different social identities into coherent groupings. The 64 identities used in the analysis were social identities that had been reported in a previous study as being important to participants (Deaux, 1991), as well as some more stigmatised identities that were added to increase representativeness of the set. The participants were laypeople; the researchers did not collect information on whether they classed themselves as in- or outgroup members of any of the 64 identities. Cluster analysis and multidimensional scaling of these data suggested that individuals perceived five different types of identity: relational, vocational/avocational, political, stigmatised and ethnic or religious.

In light of this finding, researchers have looked to understand the value of Deaux et al.'s classification system (1995) for predicting group processes and outcomes. For example, Ethier and Deaux (1990) find that in ethnic identities (with impermeable boundaries), status is a key predictor of identification. Through studying Hispanic students at university, they noted that those who perceived their ethnic identity as lower status (due to the perceived threat of the university environment) were more likely to report lower identification with their identity. Conversely, Jackson et al. (1996) noted that in stigmatised low-status groups such

as cigarette smokers, when identity boundaries were perceived as impermeable and the group was provided with negative information about the group (vs no information), group members rated their group as more likeable and competent. The findings of these two studies suggested that permeability of identity boundaries may impact different types of identity (e.g., stigmatised or ethnic), in different ways.

More aligned with the ideas of the present research, Huddy (2001; 2004) suggests that political identities tend to be perceived as less variable and more stable than has been suggested by social identity theory (Haslam et al., 1992; Turner et al., 1987). Whilst she contends that political identities, such as feminists, may show some movement over time, the prototypical view of a feminist is a more enduring representation than social identity theory suggests (Huddy, 2001). Huddy suggests that by relying on the minimal group paradigm to understand group-based outcomes, social identity theorists tend to downplay the significance of the history and lived reality of certain types of group. Whilst this idea has not been studied in detail as of yet, we can see how the type of group to which an individual belongs, i.e., stigmatised, political or ethnic, may be found to have a measurable impact on group processes such as identification, stability and ingroup favourability (Ethier & Deaux, 1990; Jackson et al., 1996; Lickel et al., 2000).

Whilst Deaux et al.'s (1995) classification system has proved valuable in understanding the impact of different types of group on group related processes, this typology fails to consider how individuals perceive their own group identities. For example, whilst a layperson in Deaux et al.'s study may consider the transgender identity as a stigmatised identity, it is not clear whether transgender individuals would agree with this classification. Here we return to the argument made earlier with regards to the difference in the way an outgroup may perceive an identity

(stereotype) and the way an individual perceives their own identity. As the people in Deaux et al.'s (1995) study did not clarify their group status – i.e., in- or outgroup – their typology only allows us to assess layperson or societal-level stereotypes about groups, but not prototypes.

In addition to this, several researchers have demonstrated how the content of particular social identities is valuable to understanding and predicting group-related outcomes. For example, Livingstone and Haslam (2008) illustrate the importance of antagonistic identity content in understanding intergroup conflict in real-world groups. Similarly, Jetten et al. (1996) show that groups who perceive themselves as fairer are less likely to discriminate against outgroups. Additionally, a plethora of lab-based research has indicated that the priming of specified social norms in minimal group paradigms can impact how group members respond to later tasks (e.g., Postmes et al., 2001b). Finally, crowd-based research by Reicher (1984) demonstrates the value of paying attention to the content and prototypes of particular groups in order to understand collective group behaviour; using interview data from individuals involved in the Bristol riots of 1980, Reicher noted that the individuals behaved in line with their group prototypes, refusing to engage in actions they deemed to be in antithesis to their group's values. Therefore, despite the value in understanding how groups perceive the content and prototypes of their own social identities, no such research has sought to create a typology or framework of social identities based on ingroup perceptions or behaviour.

However, now we have evidence to suggest that some aspects of group prototypes may be more stable and less robust to the influence of local group norms (Chapter 2), it is possible to ascertain whether the more stable and less context-dependent aspects of the prototype can be used to create a meaningful typology of

social identities. Herein lies the value of the behavioural prototypicality measure that has been developed in the previous chapters of this thesis. Using this measure, we contend that we can use the similarity of linguistic styles to infer similarity between group prototypes. Thus, we propose that two groups with similar linguistic styles will also have similar group prototypes. We therefore suggest that this linguistic approach can enable us to create a typology of identity prototypes based on naturally occurring online data.

In order to test the idea that similar group prototypes will have similar linguistic styles, it is necessary to have an approximate ground truth as to which identities are similar in their prototype. Whilst, as outlined previously, no research has directly compared the similarity of different identity prototypes using ingroup members, we suggest that the five categories outlined by Deaux et al. (1995) gives us a starting point for understanding prototype similarity. Whilst the participants in Deaux et al.'s study would have simply been accessing a stereotypical view of the social identities – as it is unlikely that they were ingroup members of all 64 identities – we suggest that the findings of this study provide the closest estimate of prototype-based similarity between social identities. Further, Jussim et al. (2015) suggest that stereotypes tend to be relatively accurate representations of groups, at least for some types of groups, such that the beliefs about an outgroup are often accurate when compared against the actual behaviour of a group. Whilst this 'accuracy' does not assess whether a group's perception of itself (the prototype) matches an outgroup's perception (the stereotype), it nonetheless gives weight to the idea that stereotype similarities may be a strong starting point for considering prototype similarities.

4.1.1   Hypotheses

Therefore, based on the findings of Deaux et al. (1995), our hypothesis states that social identities from the same group type (vocational/avocational, relational, ethnic and religious, stigmatised and political), will have more similar linguistic styles to other identities within their group type than between group types. This pre-registered hypothesis and methodology can be found at https://osf.io/jk6na/.

4.1.2   Method & Analysis

4.1.2.1   Data Collection

To develop our typology of prototypes, we chose three identities from each of the five group types outlined by Deaux et al. (1995). This allowed us to ensure that our choice of groups was both wide-ranging as well as meaningful to individuals (Deaux, 1991). We assessed which of the groups used in Deaux and colleagues' (1995) research had suitable Reddit forums from which we could collect data and chose those with substantial Reddit communities – these choices were preregistered at https://osf.io/jk6na/. The subreddits from which we collected our data are outlined below in Table 4.1.

Table 4.1.

Information Pertaining to the 15 Subreddits used in the Analysis

Social Identity and Type	Subreddit	Total users	Subreddit description
Partner (relational)	r/relationships	~9m	A community built around helping people and the goal of providing a platform for interpersonal relationship advice between redditors. We seek posts from users

			who have specific and personal relationship quandaries that other redditors can help them try to solve.
Mother (relational)	r/breakingmom	~61k	Moms only. Just say what's going on. No judgments, no nastiness.
Father (relational)	r/daddit	~180k	Subreddit for Dads. Single Dads, new Dads, Step-dads, tall Dads, short Dads and any other kind of dad.
Teacher (vocational)	r/Teacher	~7.1k	n/a
Entrepreneur (vocational)	r/Entrepreneur	~750k	A community of individuals who seek to solve problems, network professionally, collaborate on projects and make the word[sic] a better place. Be professional, humble, and open to new ideas.
Salesperson (vocational)	r/sales	~135k	Everything you need to know about sales, selling, business development, lead generation, prospecting, closing and more!
Muslim (ethnic/religious)	r/islam	~132k	r/islam is the place to discuss any topics related to Islam and Muslims
Asian American (ethnic/religious)	r/asianamerican	~43k	Anything related to Asian and Pacific Islander Americans, as well

as other Asians who grew up outside of Asia. This includes news, discussions, pictures or videos. Whilst members of all races and nationalities are welcome, our purpose is to foster a sense of community among Asian Americans and their respective counterparts in the Asian diaspora. Topics do not necessarily need to be related to race as long as they contribute to the community.

Christians (ethnic/religious)	r/Christianity	~262k	A subreddit to discuss Christianity and aspects of Christian life. All are welcome to participate.
Libertarian (political)	r/Libertarian	~442k	A place to discuss libertarianism, politics, related topics, and to share things that would be of interest to libertarians.
Conservative (political)	r/Conservative	~568k	The place for Conservatives on Reddit.
Feminist (political)	r/Feminism	~195k	Discuss and promote awareness of issues related equality for women
Alcoholic (stigmatised)	r/stopdrinking	~261k	This subreddit is a place to motivate each other to control or stop

			drinking. We welcome anyone who wishes to join in by asking for advice, sharing our experiences and stories or just encouraging someone who is trying to quit or cut down. Please post only when sober; you're welcome to read in the meantime.
Transgender (stigmatised)	r/asktransgender	~149k	Transgender questions; transgender answers.
Homeless person (stigmatised)	r/homeless	~27.9k	This is a discussion and advice group. Do not beg or soft-beg for cash, donations etc.

After receiving ethical approval from the University of Exeter Psychology Ethics Board, and pre-registering the methodology and hypotheses, we collected one year's worth of posts and comments from the 15 subreddits listed above. Using Google BigQuery, we were able to collect posts or comments that had been posted to the aforementioned forums between January 2018 and January 2019. As per the methodology outlined in Chapter 2 of this thesis, we collected the title, text, URL and author of all posts and comments.

Following data collection, we quantified the linguistic data using LIWC 2015 software (Pennebaker et al., 2015). In line with earlier chapters, we used the 41 stylistic categories outlined in Appendix A. We then performed the same data cleaning tasks outlined in Study 2a in order to ensure that we had a high-quality

dataset from which we could draw robust conclusions. We excluded all posts made by bots, all authors who have deleted their accounts or have had their accounts removed, all posts which have been deleted or removed, all posts that contain only a URL and all posts with less than 50 words in line with common practice in computational psycholinguistic research using LIWC software, as well as in line with the methodology outlined earlier in this thesis. Table 4.2 below indicates how many posts and comments remained in our dataset after the data had been cleaned.

Table 4.2

Data Remaining After Excluding Low-Quality Posts and Comments

Subreddit	Number of Posts	Number of Comments
r/Entrepreneur	16,466	91,035
r/Teachers	13,613	98,806
r/sales	5,590	25,260
r/asianamerican	521	17,520
r/islam	3,616	54,511
r/Christianity	19,086	356,778
r/Conservative	1,500	134,804
r/Libertarian	4,580	357,292
r/Feminism	627	15,390
r/relationships	71,292	103,382
r/daddit	2,484	17,393
r/breakingmom	12,590	87,326
r/asktransgender	30,695	207,587
r/homeless	1,033	7,750
r/stopdrinking	49,730	204,934

As outlined in Chapter 2, imbalanced class sizes can adversely impact the results of machine learning algorithms, and so it was necessary to select an equal sample of each identity's data. We therefore selected 521 posts from each forum (as this was the number of posts from the forum with the fewest posts), and 7,750 comments (as this was the number of comments from the forum with the fewest comments) from each forum. We selected a total of 8,271 contributions from each forum to include in our analysis.

4.1.2.2   Analytic Strategy

4.1.2.2.1   Creating the Dissimilarity Matrix

In order to quantify the similarity between group prototypical linguistic styles, we compared each identity with each other identity using an Extra Trees model; in total, there were 105 pairwise comparisons. We included all 41 linguistic style features in the analysis (see Appendix A)[4]. We divided the 8,271 datapoints for each identity into a training set and a test set; we used 4,136 to train each model, and 4,135 to test each model.

We therefore ran 105 Extra Trees algorithms to determine how close two groups were in their linguistic style, using the resultant AUCs as a measure of dissimilarity between groups. Where two social identities had a similar linguistic style, the AUC would be closer to 0.5 (= random guess) demonstrating that the classifier was less able to distinguish between the two identities. Conversely, in social identities with particularly distinct stylistic differences, the AUC of the Extra Trees classifier would be closer to 1.0 (= perfect separation). By using the AUC output of the Extra Trees classifier in all 105 pairwise comparisons, we constructed a

---

[4] We also completed the analysis using only 'function word' categories. Instead of 41 generally stylistic linguistic categories, function words refer to indicators which are known to be very challenging to consciously control. The results from this analysis are very similar and are outlined in Appendix C.

dissimilarity matrix which illustrated how dissimilar each identity was from each other identity (see Figure 4.1 below).

Figure 4.1

Shaded Dissimilarity Matrix Consisting of AUC Results (Between-Groups)

Category		entrepreneur	sales	teachers	asianamerican	islam	christianity	libertarian	feminism	conservative	breakingmom	daddit	relationships	stopdrinking	asktransgender	homeless
Vocational	entrepreneur	0.5	0.68	0.76	0.87	0.86	0.87	0.86	0.86	0.88	0.92	0.9	0.95	0.9	0.83	0.79
	sales	0.68	0.5	0.72	0.88	0.88	0.89	0.9	0.89	0.91	0.9	0.88	0.94	0.89	0.84	0.76
	teachers	0.76	0.72	0.5	0.82	0.84	0.84	0.87	0.82	0.85	0.85	0.84	0.91	0.85	0.79	0.76
Ethnic/Religious	asianamerican	0.87	0.88	0.82	0.5	0.81	0.83	0.83	0.73	0.77	0.89	0.88	0.94	0.92	0.83	0.87
	islam	0.86	0.88	0.84	0.81	0.5	0.65	0.79	0.75	0.77	0.92	0.9	0.92	0.94	0.85	0.87
	christianity	0.87	0.89	0.84	0.83	0.65	0.5	0.78	0.75	0.77	0.91	0.89	0.92	0.93	0.85	0.89
Political	libertarian	0.86	0.9	0.87	0.83	0.79	0.78	0.5	0.74	0.68	0.95	0.94	0.96	0.96	0.9	0.89
	feminism	0.86	0.89	0.82	0.73	0.75	0.75	0.74	0.5	0.73	0.9	0.9	0.92	0.93	0.82	0.87
	conservative	0.88	0.91	0.85	0.77	0.77	0.77	0.68	0.73	0.5	0.93	0.91	0.95	0.96	0.9	0.89
Relational	breakingmom	0.92	0.9	0.85	0.89	0.92	0.91	0.95	0.9	0.93	0.5	0.7	0.83	0.88	0.86	0.86
	daddit	0.9	0.88	0.84	0.88	0.9	0.89	0.94	0.9	0.91	0.7	0.5	0.87	0.89	0.87	0.86
	relationships	0.95	0.94	0.91	0.94	0.92	0.92	0.96	0.92	0.95	0.83	0.87	0.5	0.94	0.9	0.91
Stigmatised	stopdrinking	0.9	0.89	0.85	0.92	0.94	0.93	0.96	0.93	0.96	0.88	0.89	0.94	0.5	0.82	0.88
	asktransgender	0.83	0.84	0.79	0.83	0.85	0.85	0.9	0.82	0.9	0.86	0.87	0.9	0.82	0.5	0.81
	homeless	0.79	0.76	0.76	0.87	0.87	0.89	0.89	0.87	0.89	0.86	0.86	0.91	0.88	0.81	0.5

4.1.2.2.2 Multidimensional Scaling

The next step was to run multidimensional scaling on the dissimilarity matrix to understand how the similarities between groups can be conceptualised in n-

dimensional space. Specifying a Euclidean distance model, we computed and

plotted the eigenvalues on a Scree plot, noting that the model that best fit the data

had two dimensions (Kruskal & Wish, 1978).

The two-dimensional MDS values were plotted on a scatterplot (Figure 4.2).

From Figure 4.2, we can identify that our five groups of identity each appear to

cluster together. That is, the social identities that comprise each group type

proposed by Deaux et al. (1995) (vocational, relational, stigmatised, ethnic/religious

and political) cluster more closely together on the two-dimensional plot.

Figure 4.2

Multidimensional Scaling Plot Illustrating the Relations Between the 15 Different

Social Identities (Between-Groups Analysis)
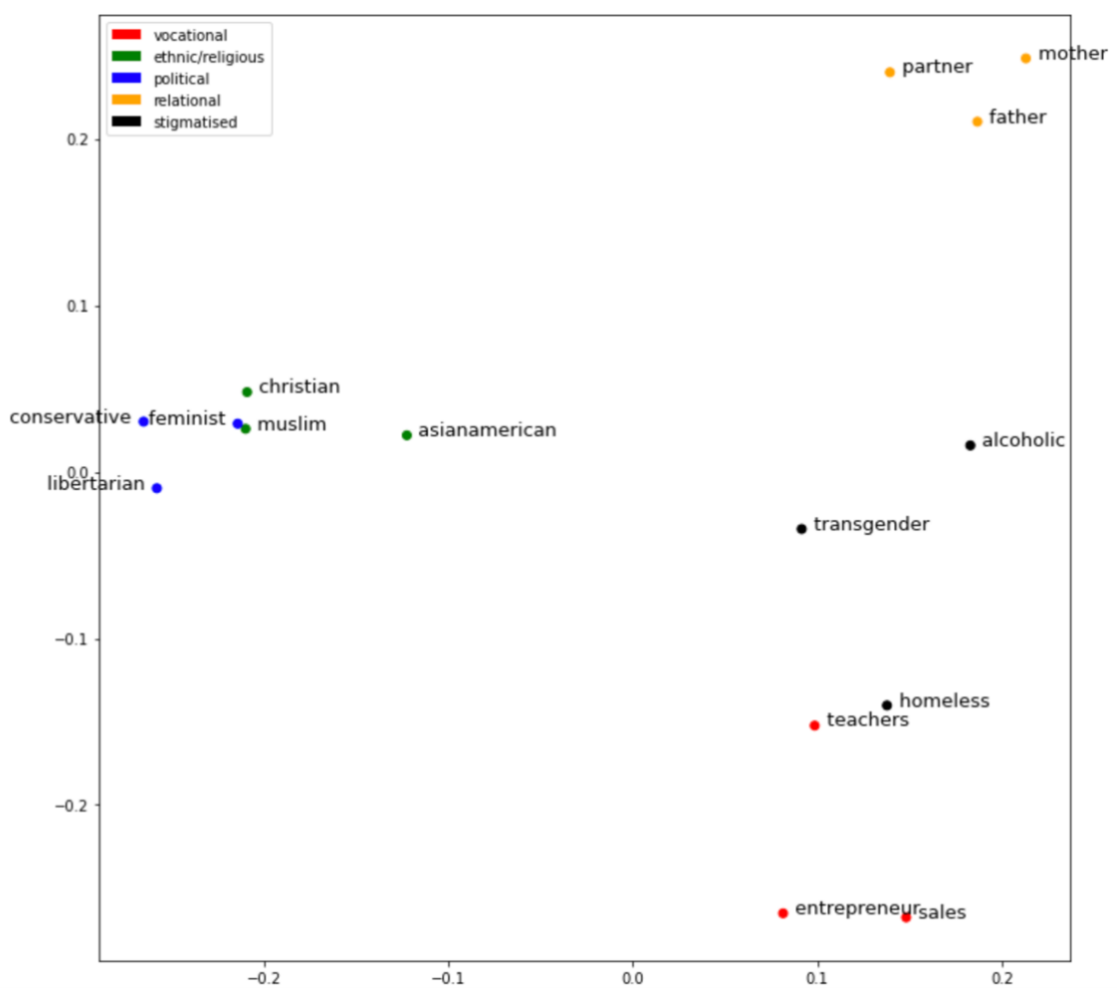
4.1.2.2.3 Cluster analysis

In our pre-registered methodology, we outlined that we planned to run a cluster analysis to further understand how the identities cluster. Prior to undertaking the MDS analysis outlined above, we did not know the number of dimensions that best fit the data. Fortunately, as the MDS analysis showed that the data were best fit to two dimensions, the scatterplot in Figure 4.2 above allowed us to directly visualise each cluster. Comparatively, in Deaux et al.'s (1995) analysis, their 64 identities were best explained using 5 dimensions. In this instance, cluster analysis would have been essential due to the inability to plot the identities in 5 dimensions. Therefore, whilst cluster analysis was not essential to understanding the clustering of identities in our own analysis, we decided to complete it anyway in keeping with the methodology which had been specified in advance.

Hierarchical agglomerative cluster analysis was performed on the dissimilarity matrix outlined in Figure 4.1. Agglomerative clustering starts with each social identity as a singleton cluster and merges clusters successively based on their similarity. Similarity (or distance between clusters) can be calculated in multiple different ways (see Nielsen, 2016), however in order to test the hypothesis that social identities are more similar to their own group type than other group types, we required a similarity measure that computes within-cluster variance. For this reason, we used Ward's method as Ward's method aims to find the pair of clusters that have the lowest increase in within-cluster variance after merging (Nielsen, 2016). Ward's method calculates the distance between two clusters by computing the increase in the sum of squares of two clusters when merged.

Using Ward's method, we note four main clusters as seen in Figure 4.3 below. As is evident in Figure 4.3, the social identities do not cluster perfectly into our five

hypothetical groups, however we can still observe distinct overlap between the groups proposed by Deaux et al. (1995) and the groups uncovered in our bottom-up behavioural analysis here. Possible reasons for the disparity are suggested in the discussion.

Figure 4.3

Dendogram of 15 Social Identities Using Hierarchical Agglomerative *Cluster Analysis (Ward's method)*
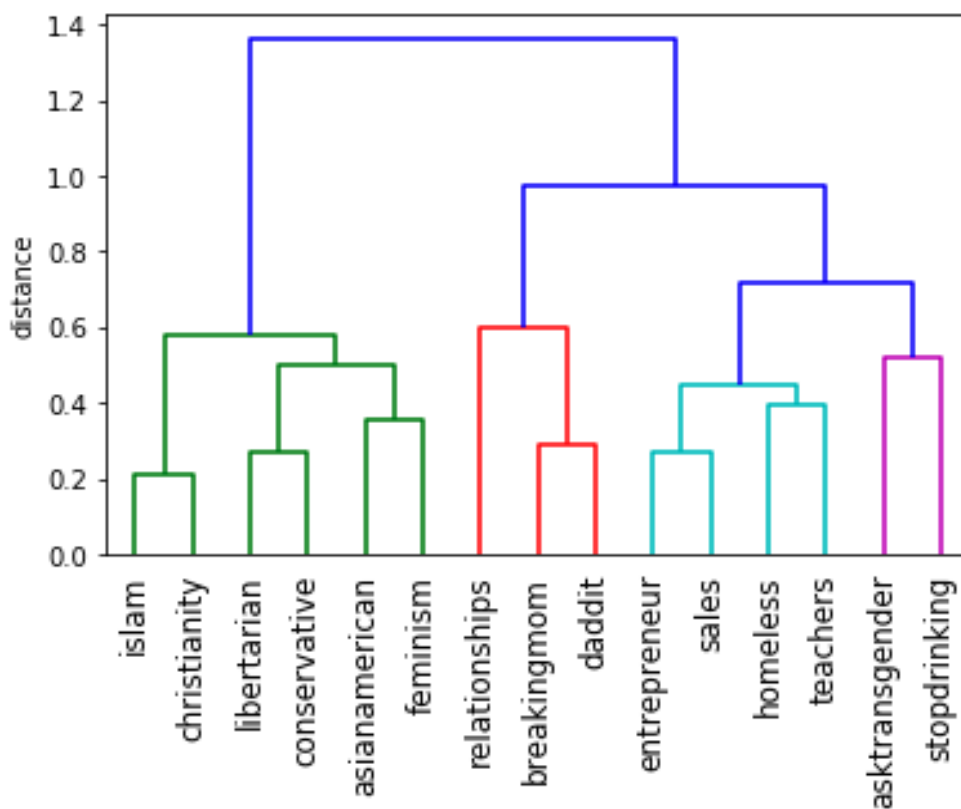
4.1.3   Discussion

In Study 3a, we looked to create a typology of social identities based on similarities in prototypical linguistic style behaviour. Specifically, we explored whether the perceived similarities noted by Deaux et al. (1995) corresponded with prototypical behavioural similarities. We found that when using multidimensional scaling to visually map the linguistic similarities between social identities, we were

able to identify the five types of identity proposed in the perception-based research of Deaux et al. (1995); thus, the three social identities from each of our five types of identity were similar in their linguistic style and tended to cluster together. We also used cluster analysis to further test the hypothesis that similarities within-group type would be greater than between-group type. The results of this analysis demonstrated remarkable similarity to the five groups proposed by Deaux et al. (1995), although we also noted overlap between our ethnic and religious identities with our political identities. Further, contrary to our predictions, we noted that our homeless identity tended to cluster with the vocational/avocational identity rather than with the stigmatised identities.

Whilst previously it could have been argued that each social identity has a way of communicating that is independent of other social identities and this behaviour is an end in itself, here we have demonstrated more clearly that similarities in linguistic style map to similarities in social identity. This provides compelling evidence to support the idea that prototypical behaviour can be used to understand and study social identities 'in the wild'. That is, not only do groups have different linguistic styles (Chapter 2), but these linguistic styles reveal something about the groups themselves. In contrast to content-based norms, whereby different groups may adopt the use of different words (e.g., 'fleek'; Grieve et al., 2017), here we can see that explicit patterns in communication style relate to the type of group displaying those norms. Arguably, it can therefore be suggested that by using linguistic style we are able to understand more about how individuals express their identities prototypically, instead of just how they adopt local norms. In this way then, it can be argued that groups may have an enduring sense of identity that guides and influences their

behaviour above local group norms, locally induced prototypes, or even specific attitudes and beliefs.

As we can see from Figure 4.2, the method of using linguistic style appears to transcend the finer content-based and arguably more locally-fluid aspects of group identities. In Figure 4.2, we see that the feminist identity is unintuitively (from an attitudes perspective) close to the Muslim identity. As argued earlier in the thesis, this may be because specific beliefs and attitudes are more amenable to change in interaction with others. This has been demonstrated consistently within the polarisation literature (Mackie, 1986) even when controlling for superordinate category identities (Koudenburg et al., 2019). Conversely, by using linguistic style indicators we suggest that we are measuring a more generalised idea of what it means to be a member of a particular group. In this way, style captures the way that identities are enacted and expressed as opposed to the direct opinions associated with them.

As can be seen on the MDS plot in Figure 4.2, and was later reiterated by the cluster analysis, the political and ethnic/religious identities appear to cluster closely together. The reason that these results can be seen to differ from the findings of Deaux et al. (1995) pertains to the meaning ascribed to particular ethnic and religious labels. To an outsider, or an individual who may not personally identify with the category label (e.g., the participants in Deaux et al.'s (1995) study), ethnic and religious labels may be used as social labels to divide individuals into distinct groups on the basis of perceived physical and behavioural differences (McGarty, 1999). Thus, whilst an outsider may group social identities together such as 'ethnicities' or 'religions', this approach fails to comprehend what those identities actually *mean* for the individuals and what the identity label stands for *in practice.*

As has been argued elsewhere (Young, 1990), races, and by extensions ethnicities, are inherently political by the fact that they exist. Races are often defined in antithesis to the majority group; an individual may define someone as an Asian American in order to distinguish their ethnic identity from the superordinate American identity label. Conversely, the 'White American' is perceived as the prototypical version of an American (Danbold & Huo, 2014; Devos & Banaji, 2005) and thus does not require an explicit label. In this way then, the label 'Asian American' is used as a political tool to emphasise the boundary between Asian and White Americans. It therefore follows that the Asian American identity is political in nature[5], even if this is not recognised by laypeople such as those in Deaux et al.'s (1995) study. This argument is further supported by the lack of existence of a 'White American' or 'European American' Reddit forum. It is therefore no surprise that the Asian American identity appeared closer to the overtly political identities than when only stereotypical, perceptual judgments were used to create an identity typology. This discrepancy between Deaux and colleagues' (1995) results and our own points to the crucial value of understanding prototypical behaviour as a direct enactment of social identities in the real world, as opposed to relying merely on the stereotypical perception of identities.

In a similar vein, the similarity between religious and political identities can be explained through understanding the fervently agentic nature of both of these identities (Deaux et al., 1995, Study 2). More specifically, both religious and political identities are involved in collective activism with the intent of improving society; both identity types aim to create a lens through which to interpret human action and as a

---

[5] Note. This research was undertaken before the Covid19 pandemic which likely exacerbated the politicisation of the Asian American identity.

result, a blueprint to improve upon society. In turn, our analysis as to how these identities are enacted goes beyond merely categorising individuals as similar (McGarty, 1999; McGarty et al., 2002; Deaux et al., 1995). In fact, our focus on understanding explicitly how identities are enacted and expressed fundamentally disagrees with the viewpoint of researchers such as Smith (2004) who, for example, define political identities as "the collective label for a set of characteristics by which persons are recognized by political actors as members of a political group." (p. 302). By focusing only on social categorisation as opposed to self-categorisation, a crucial element of the purpose and value of social identities is neglected. Whilst we do not doubt that social categorisation produces meaningful understandings of how groups are perceived, in order for self-categorisation theory to advance past a perceptual understanding of group processes, we need to devote more time and research to directly understanding the purpose and behaviours of groups within natural contexts, and from ingroup members' perspectives. Thus, whilst it may be possible to ask individuals to identify traits that relate to their social identity prototypes (e.g., Hains & Hogg, 1996; Haslam et al., 1995), it is argued that this approach fails to grasp the true meaning and purpose of particular social identities. Conversely, we suggest that it is imperative to understand what an identity label means to people *in situ*.

This idea is neither new nor revolutionary, but the evidence presented in this study serves to further underline its importance. A similar argument has been made recently with regards to the study of prejudice, referring to the importance in understanding prejudice as an expression of identity performance (Durrheim et al., 2016). The authors coherently argue that the 'nature of prejudice' has been defined top-down as if prejudice were a cognitive style, personality trait or affective response (e.g., Dovidio et al., 2005; Duckitt, 1992). However, they argue, in line with others

(Reynolds et al., 2012), that these definitions of prejudice are least useful in contexts when attempting to understand how prejudice is mobilised for a specific cause. In turn, they point to the need to understand how and when prejudice is mobilised as an expression of identity, instead of as a trait of the individual.

In the same way then, the present research findings and the disparity between our findings and the results of Deaux et al. (1995), point to the value in understanding identities not simply as category labels imposed upon individuals either by laypeople or in laboratory-based research, but instead as socio-cognitive processes which are used to create and construct individuals' social realities. This is in synthesis with the intended study of social identities as put forward by Tajfel (1979). The meanings ascribed to social identities by those who self-identify serve an explicit purpose; they are not merely category labels that can be captured using traits such as 'happy-go-lucky' and 'pleasure-loving' (Haslam et al., 1995). They are also not simply labels that can be comprehensively understood by asking laypeople's perspectives (Deaux et al., 1995). Social identities, and the meanings individuals ascribe to them, are thus ways of making sense of reality, and to study them outside of the context in which they are mobilised offers little value in understanding their direct influence on behaviour.

At this point, it bears noting that in this research the context is an online public forum. It is likely that this specific context will have an impact on how individuals choose to communicate and the purpose of their communication. Further, it could reasonably be argued that the linguistic style of a religious community in a community newsletter, for example, would be different to the linguistic style of religious individuals in a Reddit forum. However, on the other hand, it must also be noted that in Chapter 2 (Study 1e), we were still able to detect a prototypically

libertarian and prototypically entrepreneurial style even when individuals were posting outside of the online forum context. This finding suggested that there is an aspect of the communication which transcends the explicit purpose or mobilisation of these identities. Having said this, context may indeed play a key role in impacting the purpose of communication and the style of group communication, although I suggest that this is a 'feature' and not a 'bug' of this methodological approach. I do not aim to deny the role of context and purpose, but instead argue that the purpose of communication is intrinsically tied up with the 'purpose' and defining features of the identity. Conversely however, platform-specific effects are not a feature of the identity (e.g., perhaps Reddit is an inherently political space). Therefore, in Study 3d, we aim to control for platform-specific effects which may be unrelated to the purpose and motivations of the identity through using data from other online data sources.

As noted previously in this thesis, between-groups analysis may not allow us to observe directly how individuals change their style in line with their salient social identity. For the next part of the analysis therefore, we will repeat this study using a within-individual design. This will enable us to control for stable individual-level factors such as demographics and personality.

4.2   Study 3b: Excluding Demographic and Personality Differences Using a Within-Individual Design

For the next part of the analysis, in synthesis with Chapter 2, we wanted to rule out demographic or personality differences as a possible explanation for the results of Study 3a. As has been outlined earlier in this thesis, previous research has demonstrated the impact that demographic and personality factors may have on linguistic style (e.g., Newman et al., 2008). By completing a within-individual study

design, this allows us to demonstrate that individuals change their linguistic style based on the salience of their social identity.

4.2.1 Hypotheses

Our hypothesis for this section remains that we will still be able to distinguish between the four or five types of identity using only within-individual data.

4.2.2 Method

For Study 3b, we took the dataset outlined in Study 3a and isolated the individuals who had posted in two of our forums of interest (N=50,536). These 50,536 individuals were individuals who had posted in *any* two forums. The total number of overlapping individuals in each pairwise comparison is shown in Figure 4.4 below. Thus, for example, by examining Figure 4.4 we can see that 743 individuals posted in both the entrepreneur forum as well as the conservative forum.

Figure 4.4

Sample Size Included in Within-Individual Analysis

It must however be noted that there are some identities which are mutually exclusive. For example, it is unlikely that an individual who posts in both the Islam and Christianity forums regards themselves as both a Christian and a Muslim. Instead, it is more likely that they identify with either of these identities or neither. Nonetheless, we include these identities in our analysis as they will allow us to ascertain whether the AUC is driven by something other than an identity prototype. That is, we expect the AUC of our Muslim-Christian classifier to be lower than the AUC of the other classifiers, due to the fact that individuals posting in both forums

are likely not to be adhering to the prototype of both identities. According to theory,

individuals only behave prototypically when an identity with which they identify is

salient (Turner et al., 1987). By leaving these mutually-exclusive identities in the

analysis therefore, we can assess whether this suggestion is true, or whether

individuals behave prototypically regardless of their identification with an identity.

From each individual included in the analysis, we took one of their posts from

each of the two forums in which they commented. We isolated these posts and used

them in our test sets. Our training data were made up from the remaining posts from

both forums that were not written by the individuals in our test set.

We then ran the Extra Trees algorithm as before, testing each model on data

which had come from individuals who had posted in the two forums. In synthesis with

Study 3a, we included all 41 linguistic style features in the analysis. Figure 4.5

(below) illustrates the dissimilarity matrix showing pairwise AUCs.

Figure 4.5

Shaded Dissimilarity Matrix of AUC Results (Within-Individual)
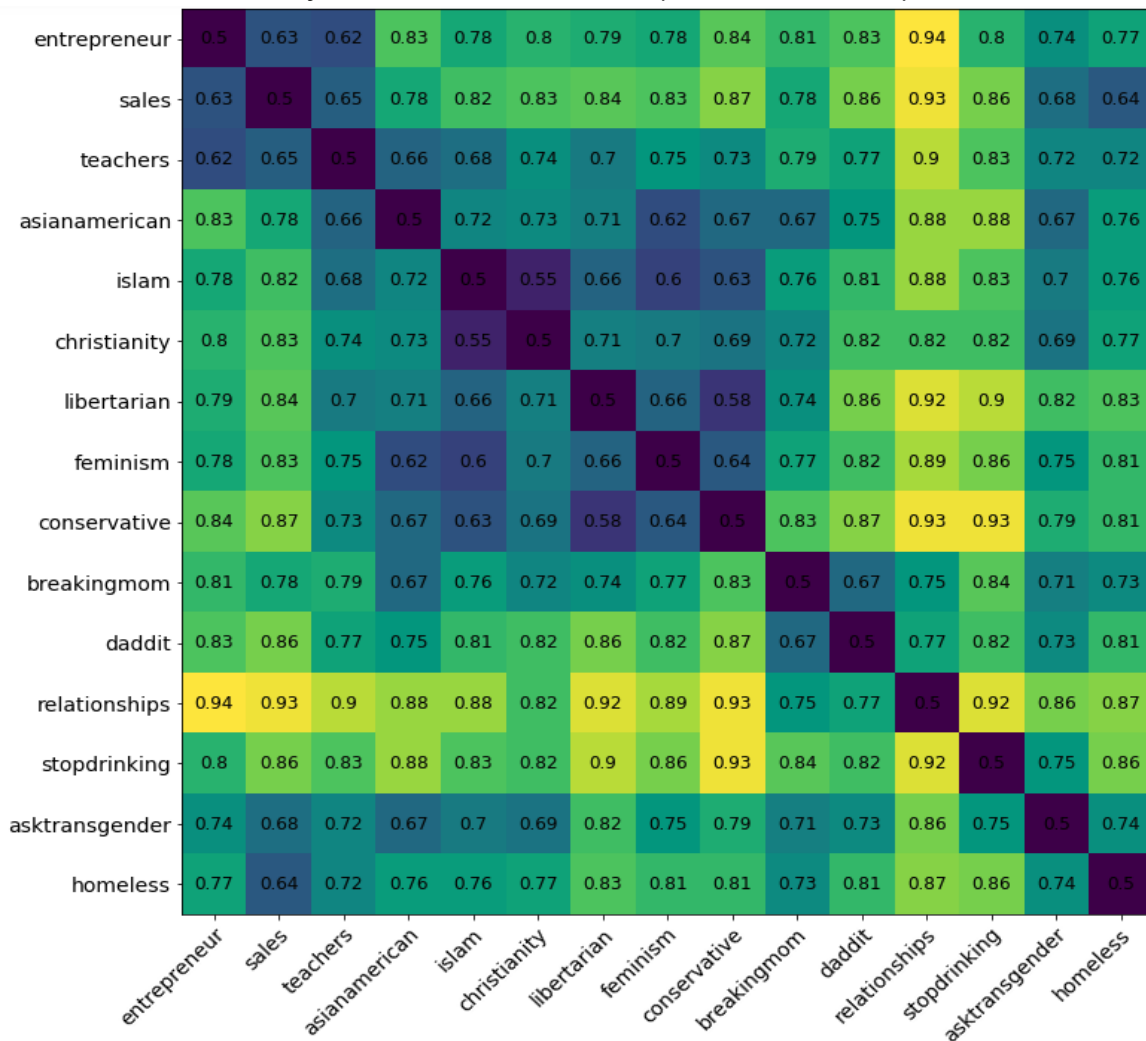
As we can see from Figure 4.5, the AUCs of identities with the same group type are substantially lower than in the between-groups design. This can be seen clearly through the darker squares on the diagonal of the dissimilarity matrix. For example, here we can see that the AUC between the r/Christianity and r/islam was very close to chance at .55 as opposed to .65 in the between-groups design. Similarly, the difference between the r/Libertarian and r/Conservative was .58 as opposed to .68 in the between-groups design. This points to the idea that individuals cross-posting in these forums were likely not identifying with both the identities, thus making it more challenging to detect a prototypical style. Instead, it is more likely that an individual

posting in two religious forums or two political forums identifies with either one of the groups, or neither.

In line with Study 3a, we then ran the MDS analysis as before, using the dissimilarity matrix outlined in Figure 4.5. Again, we visually analysed the Scree plot, finding that two dimensions best fit the data. The results of the MDS are shown in Figure 4.6 below. We found that these results were comparable to the results of Study 3a; we were able to distinguish between relational, (a)vocational, stigmatised, ethnic/religious and political identities. We again observed that the ethnic/religious identities were notably close to the political identities.

Figure 4.6

Multidimensional Scaling Plot Illustrating the Relations Between the 15 Different Social Identities (Within-Individual Analysis)
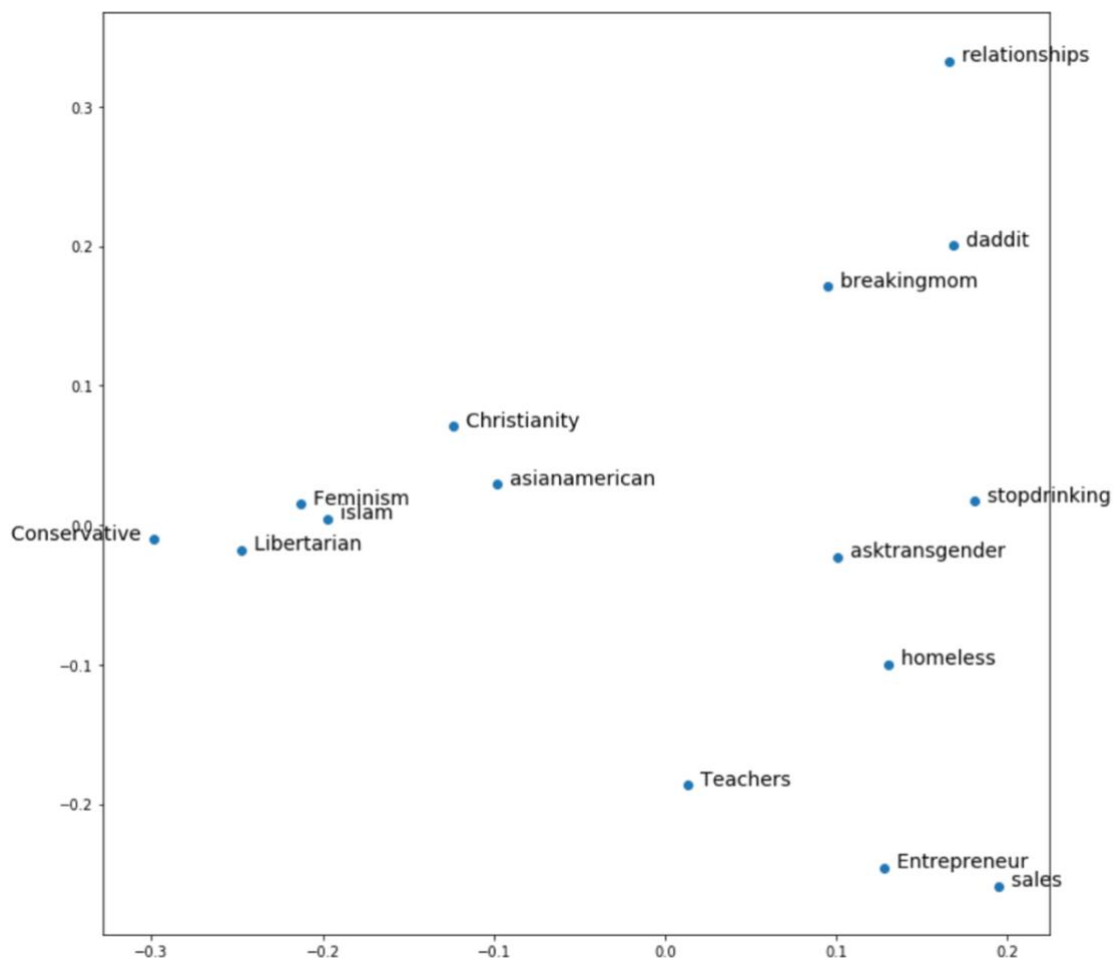
4.2.3 Discussion

In Study 3b, we demonstrated that the findings outlined in Study 3a are not explainable by demographics and personality factors alone. By controlling for these factors using a within-individual design, we were still able to demonstrate the five group types via their distinct linguistic styles. This is in line with the results that we demonstrated in Study 1b, suggesting that the linguistic style differences between our groups are not a product of individual differences.

However, in completing this analysis, there are a couple of limitations worth pointing out. Firstly, this analysis brought to light the fact that individuals posting in two forums may not self-categorise with both identities. For example, it is unlikely that an individual is both a Christian and a Muslim at the same time. However, we noted that in many of the identities whereby two conflicting identities were apparent, the AUC was notably lower (than both the between-groups analysis as well as other within-individual AUCs) suggesting that these individuals did not conform entirely with the prototypical way of communicating for both groups. Instead, it appears evident that the individuals included in this within-individual analysis still demonstrated a group type (e.g., religious) style of communication. This finding points to the value in developing a more superordinate typology of identities.

This acknowledgment of group types is particularly interesting as it gives us another lens at which to consider social identities as they are enacted in real-world environments. In the literature review, I made the argument that most research had focused on the local identities of individuals by defining the group at the forum or online community level. In Chapter 2 of the thesis, I demonstrated that there appears to be a more categorical, global identity prototype which influences linguistic style behaviour. Thus, even when individuals are communicating on different platforms, or

when identity salience is manipulated experimentally, we still tend to observe a group prototypical way of communicating (see also Koschate et al., 2021). However, in this chapter, we point to the idea that there is an even higher superordinate level of identity prototype categorisation that helps to understand how groups behave. This typology in turn can help us to determine the 'type' of an identity, before focusing on the specifics of the globally prototypical behaviour and then the locally induced identity. This idea is explored in more detail in the general discussion of this chapter.

In our research, we find that the social identities are best represented in two dimensions as per the MDS plot in Figures 3.2 and 3.6. However, at present it is not entirely clear what those two dimensions refer to. In the following section, we look to explore whether we can utilise the content of group communication to understand how these groups differ from one another. Whilst we contend that content is domain specific, e.g., an entrepreneur on the Silk Road is likely to be discussing different topics than an entrepreneur on Reddit even if their identities are similar, in order to best understand what distinguishes between these identities we suggest that we may be able to more generalised non-topical semantic indicators such as values and emotions.

4.3 Study 3c: Understanding the Dimensions of the MDS Analysis

Thus far, we have proposed that individuals change and shift their linguistic style based on the salience of a social identity and the prototypical content of that identity. In order to further validate this explanation, this study aims to ascertain whether the position of each social identity on the MDS plot corresponds with possible attributes of their group prototype. Conceptually therefore, we are aiming to

explain each identity's position in multidimensional space by using other features of written communication which may reveal cues to the content of the group's prototype.

From both a theoretical and empirical standpoint, it has been suggested that values and emotions are central to how social identities are cognitively represented by their members. Using a free recall task, Bettencourt and Hume (1999) assessed which features individuals spontaneously recalled in relation to their social identities as compared to their personal identities. By asking participants to finish 20 statements starting 'I am…' with either a personal identity or a specific social identity in mind, they found that individuals were more likely to report values, emotions and affiliative statements when asked about their social identities, as compared to their personal identities. As a result, Bettencourt and Hume (1999) concluded that the cognitive representation of the group, commonly known as the group prototype, tended to consist of values, emotions and affiliations whereas the cognitive representation of the personal identity tended to consist of traits and characteristics. These findings are in line with more theoretical perspectives on social identity content (Brewer, 1991; Deaux, 1993; Turner et al., 1987).

In addition to the research of Bettencourt and Hume (1999), more recent research by Katzarska-Miller and colleagues (Katzarska-Miller et al., 2014; Reysen & Katzarska-Miller, 2013; Snider et al., 2013) indicates the importance of values in understanding identity content. Notably, they demonstrate that individuals who identify with the label 'global citizen' report valuing diversity, endorsements to protest against unethical corporations and intergroup helping more highly than those who do not. In addition to this, Reysen and colleagues (2013) demonstrate that other groups such as Texans, students and Americans also show consensus on the prioritised

values of their social identities. As we can see from the identities included within this analysis, these groups were not mutually exclusive (an individual can identify as all three of the specified identities), thus pointing to the relativity of the identity-defining values. This therefore suggests that the values that are reported as being important to each group, corresponds with the prototype of the group as opposed to just the values of the individuals. Based on this idea then, it appears that values may be key to understanding the dimensions of differentiation between our 15 social identities.

Previously, research has suggested that individual-level values can be reliably identified using automated text analysis (Boyd et al., 2015) and thus we aim to extend this finding to assess values at the group level. Through comparing naturally occurring written communication, self-reported values and value-laden behaviours, Boyd et al. (2015) find strong support for language-based value-behaviour links. In fact, they note that when values are operationalised through employing linguistic analysis on naturally occurring social media posts, they are better able to predict an individual's future behaviour than when using self-reported scales. Building on this research, Ponizovskiy et al. (2020) developed a simple and easy-to-use value lexicon which is based on the Schwartz theory of basic values (Schwartz, 1992). Schwartz's theory of basic values (1992) suggests that there are 10 values which 'form a quasi-circumplex structure based on the inherent conflict or compatibility between their motivational goals' (Schwartz & Boehnke, 2004, p.203). This circumplex is illustrated in Figure 4.7 below.

Figure 4.7

Ten Schwartz Values Coloured to Indicate the Theoretical Model of Relations and

Higher Order Values (Schwartz, 1992).
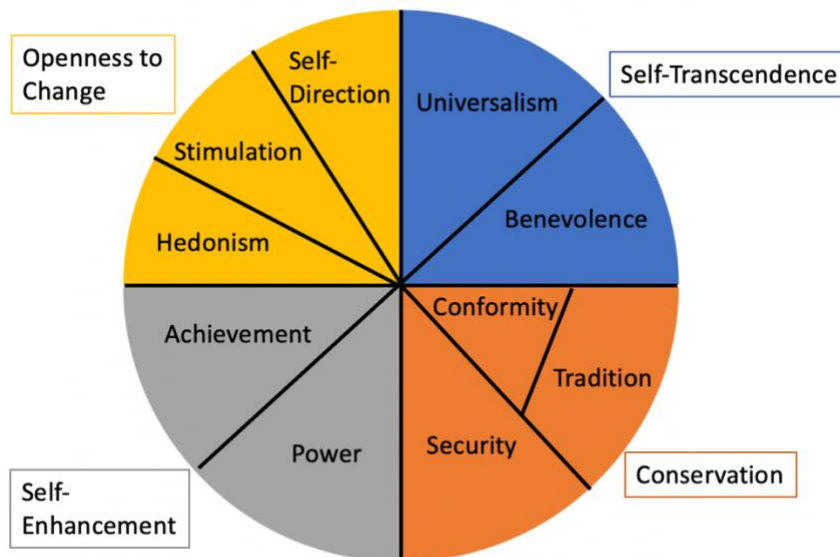
Ponizovskiy et al. (2020) find that their vocabulary-based dictionary has high

reliability and a pattern of correlations between values which shows synthesis with

the circumplex structure of values proposed by Schwartz (1992). Whilst this specific

dictionary-based tool has not yet been used to assess social group values, previous

research has demonstrated the benefit of using Schwartz values to understand and

conceptualise the values of particular social groups (e.g., Saroglou et al., 2004). In

light of this, we suggest that the dictionary-based tool may be well suited to

understanding the values of a plethora of groups.

In relation to the emotional aspects of identity content, Smith et al. (2007) noted

less variability in responses when individuals were asked to report their emotions as

group members as compared to just individuals. Specifically, Smith et al. (2007)

found that when participants answered general questions such as 'to what extent do

you generally feel _____?', they were more likely to give a broad range of answers when a personal identity was salient, than when a group identity was made salient. When individuals were asked their general emotional experience with reference to their group identities, different group identities produced different answers from the participants. That is, when the participant was considering their general emotions as a student, they reported emotions that were different to when they were reporting their emotions as a Democrat. This effect was strongest when participants were highly identified with the groups of interest. Further, the shared emotional experience of participants occurred in the absence of both fellow group members or group-related objects, thus suggesting that it was the context-independent cognitive representation of the identity that led to the change in self-reported emotions. We thus contend that emotional responses may be seen as another aspect of the global identity prototype (Branscombe & Reysen, 2008).

In relation to assessing the emotional aspects of social identities, a multitude of previous textual analysis research has utilised the National Research Council Canada (NRC) emotion lexicon (Mohammad & Turney, 2010; 2013) based on Plutchik's wheel of emotions (1984). Plutchik's theory of emotion supposes that there are eight basic emotions; joy, anger, sadness, trust, anticipation, fear, disgust and surprise (Plutchik, 1984). Using crowd-sourced data, Mohammad and Turney (2010) developed a large, high-quality lexicon that has been used widely to assess the emotional content of text through deriving a vocabulary for each of the eight emotional categories. The NRC emotion lexicon includes over 10,000 emotion-based words, making it more comprehensive than other emotion lexicons such as those included in the LIWC dictionary. The lexicon has been cited over 1,000 times in the past decade and has been used in a diverse array of tasks including: to understand

narrative arcs in stories (Reagan et al., 2016), the spread of fake news online (Vosoughi et al., 2018), personality detection (Majumder et al., 2017) and general sentiment models (Bravo-Marquez et al., 2014). We therefore suggest that the NRC emotion lexicon is well-suited to assessing the emotional content of our 15 social identities.

4.3.1 Hypotheses

In sum, we hypothesize that we can use the value-based dictionary developed by Ponizovskiy et al. (2020), and the NRC emotion lexicon developed by Mohammad & Turney (2010) to assess the values and emotions of our 15 groups using automated text analysis. Further, we hypothesize that the position of our 15 identities on our stylistic MDS plot will be predicted using the value content and emotional content of each group's forum posts. This will further add weight to the idea that the way in which a group communicates corresponds with features of their group's identity prototype such as values and emotions.

4.3.2 Method & Analysis

4.3.2.1 Values

In order to assess the values held by each of our 15 groups, we used the open-access dictionary provided by Ponizovskiy et al. (2020) found at https://osf.io/vt8nf. After installing the dictionary into the LIWC software, we ran the dictionary on the full dataset outlined in Study 3a ($N = 124,065$, $n = 8,271$). For each of the 10 Schwartz values (achievement, power, security, conformity, tradition, benevolence, universalism, self-direction, stimulation, hedonism), we calculated a score for each post which referred to the total percentage of words relating to that value. Using the mean value score for each identity, we calculated 10 value scores for each of the 15 identities. These are displayed in the matrix below (Figure 4.8).

Figure 4.8

Matrix Showing Mean Percentage of Words From Each Social Identity Relating to

Each of the 10 Schwartz Values

	security	conformity	tradition	benevolence	universalism	self-direction	stimulation	hedonism	achievement	power
entrepreneur	0.189	0.349	0.124	0.779	0.465	1.291	0.354	0.172	1.598	0.859
sales	0.18	0.309	0.1	0.751	0.582	1.03	0.324	0.184	1.73	0.779
teachers	0.234	0.443	0.088	0.913	0.355	1.163	0.286	0.236	0.985	0.441
asianamerican	0.184	0.377	0.211	0.841	0.642	1.046	0.324	0.297	0.522	0.478
christianity	0.253	0.545	3.128	0.886	0.652	1.217	0.263	0.299	0.357	0.431
islam	0.215	0.559	2.116	0.818	0.632	1.069	0.232	0.215	0.383	0.422
homeless	0.547	0.353	0.153	1.15	0.378	0.762	0.209	0.153	0.809	0.654
stopdrinking	0.234	0.243	0.13	1.158	0.335	1.007	0.323	0.433	0.676	0.371
asktransgender	0.264	0.366	0.105	1.067	0.422	0.922	0.325	0.413	0.54	0.401
breakingmom	0.29	0.274	0.137	1.424	0.228	0.722	0.201	0.331	0.521	0.328
daddit	0.264	0.246	0.106	1.495	0.243	0.727	0.24	0.39	0.591	0.294
relationships	0.25	0.347	0.151	1.937	0.308	0.944	0.224	0.426	0.534	0.377
conservative	0.336	0.788	0.268	0.587	0.76	1.08	0.257	0.265	0.551	0.673
libertarian	0.407	0.833	0.223	0.534	0.864	1.248	0.26	0.209	0.614	1.033
feminism	0.319	0.568	0.221	0.75	0.783	1.195	0.342	0.471	0.577	0.558

In the next step, we ran 10 regression analyses whereby the coordinates from the MDS were used to predict each of the 10 mean value scores per identity. Thus, we were attempting to understand whether the values that the groups discussed correlated with their location on the MDS plot. The beta-coefficients are listed in Table 4.3 below.

Table 4.3

Multiple Regression of Value Scores on Multidimensional Scaling Coordinates for 15

Identities

Values	Dimension 1			Dimension 2			R^2	p
	B	*SE*	95% CI	B	*SE*	95% CI		
achievement	.87*	.31	0.20, 1.54	-2.01***	.36	-2.80, -1.22	.76	<.001
benevolence	1.38***	.23	0.88, 1.870	1.48***	.27	0.90, 2.06	.85	<.001
conformity	-.88***	.13	-1.16, -0.59	-.05	.15	-0.38, 0.28	.79	<.001
hedonism	.07	.12	-.02, 0.34	.43*	.15	0.11, 0.75	.43	.034
power	-.41	.23	-0.92, 0.10	-.84*	.28	-1.44, -0.24	.51	.014
self-direction	-.63**	.18	-1.01, -0.25	-.55*	.21	-1.00, -0.10	.63	.002
security	-.06	.15	-0.38, 0.26	.03	.17	-0.35, 0.41	.02	.910
stimulation	-.02	.07	-0.16, 0.12	-.18*	.08	-0.35, -0.02	.33	.091
tradition	-2.34	1.19	-4.93, 0.25	.64	1.40	-2.41, 3.67	.25	.173
universalism	-1.01***	.10	-1.23, -0.80	-.33*	.12	-0.58, -0.07	.91	<.001

*p<.05   **p<.01.   ***p<.001.

From Table 4.3, we can see that identities which value achievement tend to be found towards the bottom right corner of the MDS plot, reflecting the positive correlation with the x-axis (Dimension 1) and the negative correlation with the y-axis (Dimension 2). As can be seen on Figure 4.2 (repeated below for convenience), this is where our vocational identities are situated. Conversely, identities high in benevolence tend to be found in the top right corner of the plot where our relational identities are situated. Meanwhile, identities high in conformity and universalism tend to be found on the left side of the plot, where our political and religious identities lie.

Figure 4.2

Repeated: Multidimensional Scaling Plot Illustrating the Relations Between the 15 Different Social Identities (Between-Groups Analysis)
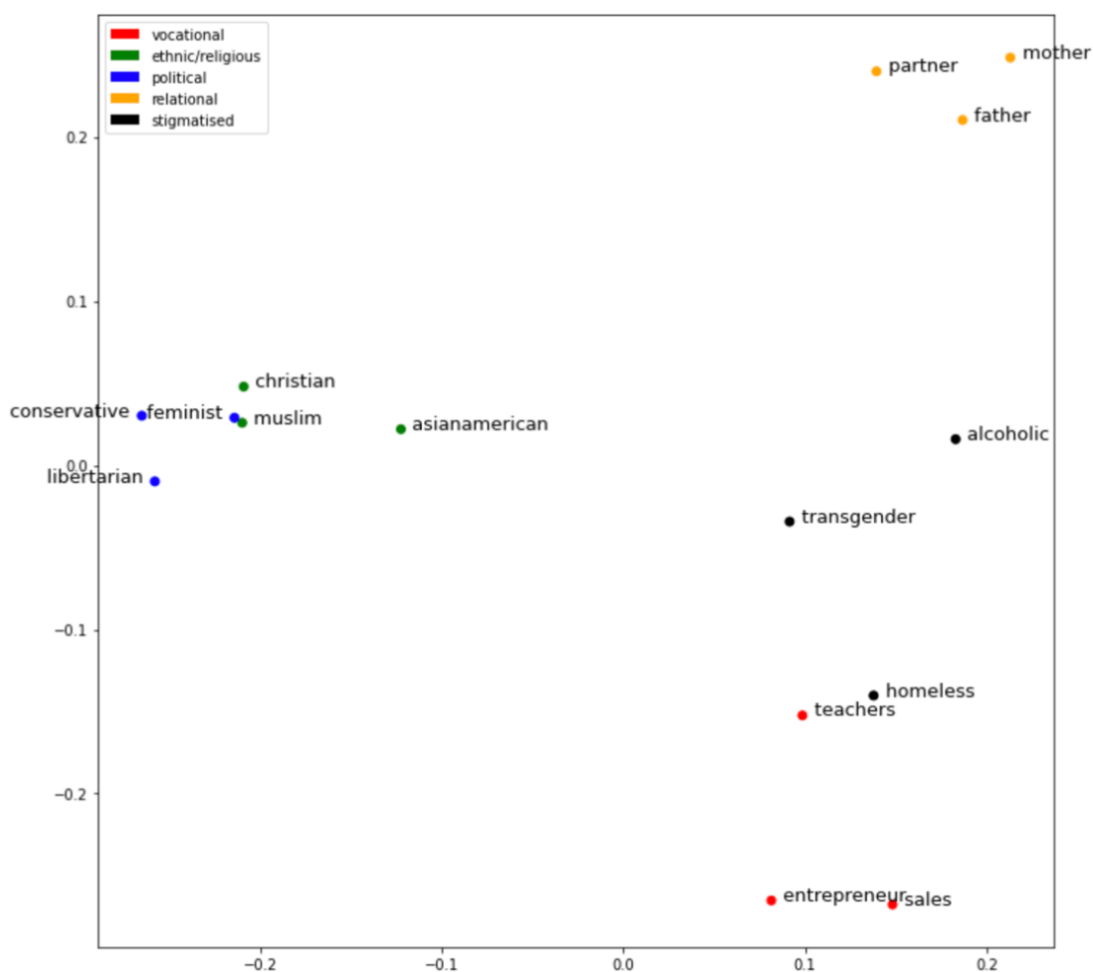
4.3.2.2   Emotions

We then repeated this methodology with an emotion-based dictionary in order to understand how the group's emotional linguistics correspond to the MDS dimensions. We used the open access NRC word-emotion association lexicon (Mohammad & Turney, 2010) found at https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm.

Repeating the method employed above, we calculated the percentage of words that related to each of the eight emotion categories (anger, anticipation, disgust, joy, fear, sadness, surprise and trust) for each post from our 15 social identities. We then took the mean score of each identity. These are displayed in Figure 4.9 below.

Figure 4.9

Matrix Showing Mean Percentage of Words from Each Social Identity Relating to

Each of the Eight Emotions

	anger	anticipation	disgust	fear	joy	sadness	surprise	trust
entrepreneur	0.1	0.17	0.078	0.108	0.145	0.104	0.118	0.177
sales	0.1	0.17	0.078	0.105	0.136	0.103	0.117	0.191
teachers	0.098	0.159	0.083	0.11	0.135	0.114	0.12	0.181
asianamerican	0.113	0.141	0.102	0.117	0.137	0.12	0.103	0.167
christianity	0.109	0.134	0.098	0.151	0.134	0.111	0.103	0.16
islam	0.115	0.142	0.104	0.127	0.132	0.122	0.104	0.154
homeless	0.115	0.154	0.105	0.12	0.127	0.119	0.106	0.154
stopdrinking	0.107	0.151	0.105	0.125	0.134	0.121	0.108	0.15
asktransgender	0.1	0.15	0.1	0.131	0.128	0.128	0.1	0.163
breakingmom	0.115	0.151	0.105	0.123	0.131	0.119	0.107	0.149
daddit	0.098	0.152	0.099	0.121	0.141	0.124	0.108	0.158
relationships	0.116	0.145	0.103	0.123	0.13	0.118	0.106	0.158
conservative	0.124	0.127	0.108	0.13	0.112	0.126	0.108	0.163
libertarian	0.126	0.128	0.109	0.139	0.113	0.124	0.101	0.16
feminism	0.12	0.129	0.109	0.129	0.127	0.126	0.092	0.168

As before, we ran eight regression analyses whereby the coordinates from the

MDS analysis were used to predict each of the eight values per identity. The beta-

coefficients are listed in Table 4.4 below.

Table 4.4

Multiple Regression of Emotion Scores on Multidimensional Scaling Coordinates for 15 Identities

Emotions	Dimension 1			Dimension 2			R^2	p
	B	SE	95% CI	B	SE	95% CI		
anger	-.07**	.02	-0.12, -0.03	.05	.02	-0.00, 0.10	.60	.004
anticipation	.04	.02	-0.01, 0.09	.01	.03	-0.05, 0.07	.20	.267
disgust	-.03	.02	-0.06, 0.00	.05**	.02	0.02, 0.09	.53	.011
fear	-.10**	.03	-0.15, -0.04	.07*	.03	0.00, 0.13	.62	.002
joy	.02	.03	-0.03, 0.08	.06	.03	0.00, 0.13	.30	.122
sadness	-.03	.02	-0.06, 0.01	.07**	.02	0.02, 0.11	.53	.011
surprise	.01*	.01	0.00, 0.03	.01	.01	0.00, 0.03	.41	.044
trust	-.06	.03	-0.14, 0.01	-.01	.04	-0.10, 0.08	.24	.197

*p<.05   **p<.01.   ***p<.001.

From Table 4.4, we note that those with higher anger and fear tend to be found to the left of the plot, where our religious and political identities lie. We also note that identities with greater disgust and sadness tend to be found towards the top of the plot. This is relatively surprising given that this is where our relational identities lie.

4.3.3 Discussion

The results of this semantic analysis demonstrate that features of prototypes such as values and emotions can be used to understand the position of the social identities on the MDS plot. That is, we note that the types of values and emotions reflected in the language of our identities relate to their position on the MDS plot. This provides evidence to support the idea that individuals are expressing their group identities through the way in which they communicate.

4.3.3.1 Values

With regards to the value content of the identities, the results indicate that identities which value achievement tend to be found near where our vocational identities are situated. This shows synthesis with the self-report findings of Deaux et al. (1995, Study 2) who noted that individuals were more likely to perceive vocational identities as achieved, as opposed to ascribed. In Deaux et al.'s (1995) Study 2, the researchers also looked to explain whether certain attributes could be used to predict the MDS coordinates of the 64 identities included in their initial analysis. Deaux and colleagues used indicators such as whether an identity was perceived to be individual or collective, ascribed or achieved and desirable or undesirable. In total, they had 15 different traits that they were using to understand the position of the identities. They did not however consider how specific values related to their five types of social identity. Yet, through using naturally-occurring online data we are able

contribute greater knowledge as to the differences between these five different types of group.

Similarly, the results of the value regression analysis demonstrated that identities high in benevolence tend to be found towards the top right corner of the plot towards our relational identities. This makes intuitive sense given that relational identities could be argued to be more care-focused as compared to other types of identity. However, it is also interesting to note that variations in benevolence may also help to understand within-cluster variations. For example, we note that the 'teacher' identity appears to be higher up the y-axis than the other vocational identities. Here, we can see that whilst teachers are clustered with other vocational identities, their position within this cluster may be explained through understanding their values, in this case the centrality of benevolence to the teaching profession (Neufeld & Hargreaves, 1999). This argument may also help to understand the position of the religious identities on the plot, as compared to the political identities; whilst both types of identity may be associated with collective action, it could be argued that benevolence plays a more fundamental role in understanding religion and religious identities than it does political identities more generally (Lee et al., 2013).

The other strong correlation between the Schwartz values and the position of the social identities on the MDS plot pertains to conformity and universalism. Specifically, we find that identities high in conformity and universalism tend to be located towards the left side of the plot. Again, this aligns with research of Deaux et al. (1995) who found that individuals tended to rate political, ethnic and religious identities as being more collective. Deaux et al. (1995) suggest that these identities

connote some form of common group membership in a way that the other types of identity (stigmatised, vocational and relational) do not.

In sum then, the relationship between the values that are reflected in the language of groups and their position on the MDS plot appears to show relatively high face validity. In order to explore this face validity further, it may be worthwhile for future research to assess whether self-reported values align with the findings outlined here.

4.3.3.2 Emotions

With regards to the emotional content of the groups' discourse, we find weaker correlations between the position of the identities on the MDS plot and their corresponding emotional content. Nonetheless, we note that identities who show greater fear and anger tend to be situated towards the left side of the plot, whereas identities high in disgust and sadness tend to be situated towards the top of the plot. These findings are interesting, as we can see how the 'hot emotions' which are associated with action tendencies predict the position of the collective identities on our plot (Klandermans & Goslinga, 1996; Miller, 2013; van Zomeren et al., 2004). Thus, in line with previous research which has suggested that collective action tendencies are increased through feelings of group-based anger (Thomas et al., 2009; van Zomeren et al., 2004), here, we directly observe that anger is an identity defining attribute of political groups. Similarly, more recent research has explored the direct role of fear in predicting or influencing collective action intentions (Adra et al., 2019), and thus once again we confirm the importance of both of these hot emotions in understanding collective action identities.

On the other hand, the finding that identities towards the top of the plot are associated with disgust and sadness is perhaps more surprising as this is where our

relational identities lie. Whilst previous research has identified the relationship between moral disgust and political action (e.g., Hancock, 2004; Gross, 2004; Seidman, 2013), or even perceptions of disgust towards more stigmatised or invisible identities (e.g., Soldatic & Meekosha, 2012) there appears to be little research that has considered disgust with regards to relational identities. At this stage then, it makes it challenging to understand what is driving the correlation between identities on our y-axis, and their level of disgust-based discourse.

It is, however, interesting to note that only the negative emotions (fear, anger, disgust and sadness) appear to be related with the position of the identities on the MDS plot. As argued by Coleman and Williams (2013) and Frijda (1986) positive emotions tend to be characterised as more diffuse affective states and therefore lack specific action tendencies that may be associated with particular social identities. Coleman and Williams (2013) and Frijda (1986) therefore suggest that due to the breadth and ubiquity of positive emotions, such as joy and surprise, they are less likely to match or be associated with specific social identities or social identity prototypes. In contrast, negative emotions are more readily distinguished and are hence believed to be associated with more specific appraisal tendencies and action readiness states (Frijda et al., 1989). In addition, it has also been suggested that when an individual experiences a positive emotion, it is more challenging to determine whether this is because the emotion fits with the salient social identity prototype, or whether it is because the emotion is hedonistically pleasing (Porat et al., 2016). For these reasons, previous social identity research has tended to concentrate more on establishing links between identities and negative emotions (Coleman & Williams, 2013; cf. Thomas, McGarty et al., 2009b). In relation to this work therefore, our finding that only negative emotions are associated with the

position on the MDS plot corroborates the suggestion that positive emotions are too diffuse to be associated with specific social identity prototypes.

Taken together, the results of this study provide evidence to suggest that the position of the identities on the MDS plot reveals something about the nature of their group prototypes. Whilst here, we have looked at the emotional and value-based content of the identity prototypes, there are other features that may also predict each groups' position on the MDS chart which are more challenging to identify using automated textual analysis. Our focus on emotions and values for this research was based both on the theoretical suggestion of their importance (Bettencourt & Hume, 1999; Tajfel, 1981) as well as the previously validated automated textual analysis tools that exist. However, we contend that there may be other aspects of identity prototypes, such as personality dimensions for example (Reysen et al., 2012), which may also correlate with the dimensions of the MDS plot. This is therefore not an exhaustive analysis to understand the differences in identity prototypes.

Nonetheless, it must be remembered that it was in fact the style of each group's communication that directly contributed to their position on the MDS. Thus, the fact that these prototypical content indicators map to the stylistic MDS dimensions, highlights the value of using style to model identity prototypical behaviour and identity enactment. Even when groups may not be discussing particular identity-consistent values or emotions or other prototypical content therefore, we are still able to identify group prototypical behaviour based on the way in which messages are communicated.

One current limitation to our approach however, concerns the platform from which our data have been collated. At present, all the data used within this analysis has originated from Reddit forums and thus it is not clear whether our findings are

applicable only to Reddit data. For our next study therefore, we decided to include data from a range of other available sources in order to further verify that these findings could be generalised outside of the Reddit platform.

4.4   Study 3d: Testing the Approach on Other Sources of Data

In Study 3d, we aimed to see whether our classifier was still able to detect an identity-prototypical style from posts that were not made on Reddit. By using posts taken from different internet platforms, this enables us to understand how platform-specific norms may influence the results. More specifically, it allows us to understand the robustness of the model and identify weaknesses in applicability outside Reddit.

At a more theoretical level however, we can also discover whether identities which have previously been defined at a local level (e.g., the inductive identities explored in Smith et al., 2015a), may also show behavioural evidence of being conceptualised at the global level. As outlined in the literature review, Smith et al. (2015a) suggested that in a Facebook group dedicated to the Occupy Wall Street movement, individuals induced their local group prototypes based on negotiations with others. Here however, we are interested in ascertaining whether our approach can be used to identify the 'type' of identity to which this group belongs (i.e., vocational, collective action (political, ethnic and religious), relational, stigmatised). We suggest that the individuals within the Occupy Facebook group may have had a generalised 'collective action' identity during communication with others. Thus, they would have already had an approximate idea of their group's identity 'type' before inducing the specifics of this identity through local communication.

Through studying the development of the Occupy identity from the Facebook group used in Smith and colleagues' 2015 analysis, we can begin to understand how

global and local prototypicality may both play a role in identity enactment online. Here then, we explore whether identities that have been induced through regular communication also show evidence of behaving in line with a generalised political identity prototypical style.

4.4.1 Hypotheses

In order to test this idea, this study will include data from the Occupy Wall Street Facebook group which was analysed in Smith et al. (2015a). We predict that this Facebook data will cluster more closely with our collective action groups on the left-hand side of the MDS plot. We will also include other openly accessible data used in previous linguistic research (Koschate et al., 2021). We hypothesize that data taken from the parenting forum 'Mumsnet' (and used within Koschate et al.'s (2021) study) will appear closer to the relational identities on our MDS plot. Thirdly, we suggest that the libertarian forum from the Silk Road cryptomarket (as outlined in Chapter 2 of this thesis) will appear closer to the collective action identities on our plot and that the Silk Road entrepreneur forum data (also used in Chapter 2 of this thesis) will appear closer to the vocational identities on the MDS plot. We decided to use these four data sources to test the robustness of the MDS methodology due to the known social identities of the individuals (parents, political group, libertarians and drug vendors), as well as the accessibility of the data.

4.4.2 Method

4.4.2.1 Additional data sources

For the analysis, we used four secondary data sources; the Occupy Facebook data from Smith et al., (2015a), the parenting Mumsnet data from Koschate et al., (2021) and the Silk Road data, divided into the libertarian forum and the entrepreneur forum, employed earlier in the thesis.

As before, we removed all posts which were under 50 words. The total number of posts included in the analysis is displayed in Table 4.5 below.

Table 4.5

*Number of Posts Included in Analysis from Non-Reddit Based Data Sources*

Data Source	Total data	Data after short posts removed	Predicted Identity Type
Occupy Facebook data from Smith et al. (2015a)	6,481	1,020	Political
Mumsnet parenting forum data from Koschate et al. (2021)	386,864	237,697	Relational
Silk Road 'Vendor Roundtable' forum data	10,056	4,746	Vocational
Silk Road 'Philosophy, Economics and Justice' forum data	10,780	5,748	Political

4.4.3 Analysis

To conduct this analysis, we repeated the methodology of the Study 3a, adding in the four extra identities. We therefore tested and trained 72 new Extra Trees classifiers, comparing each of the four new data sources with the 15 previous Reddit identities and each of the new identities (18 x 4). As per the previous studies, we trained on 50% of the data available, and tested on the remaining 50% being sure to keep both the training and test class sizes balanced. Figure 4.10 below illustrates the total dissimilarity matrix including the data from the four new identities.
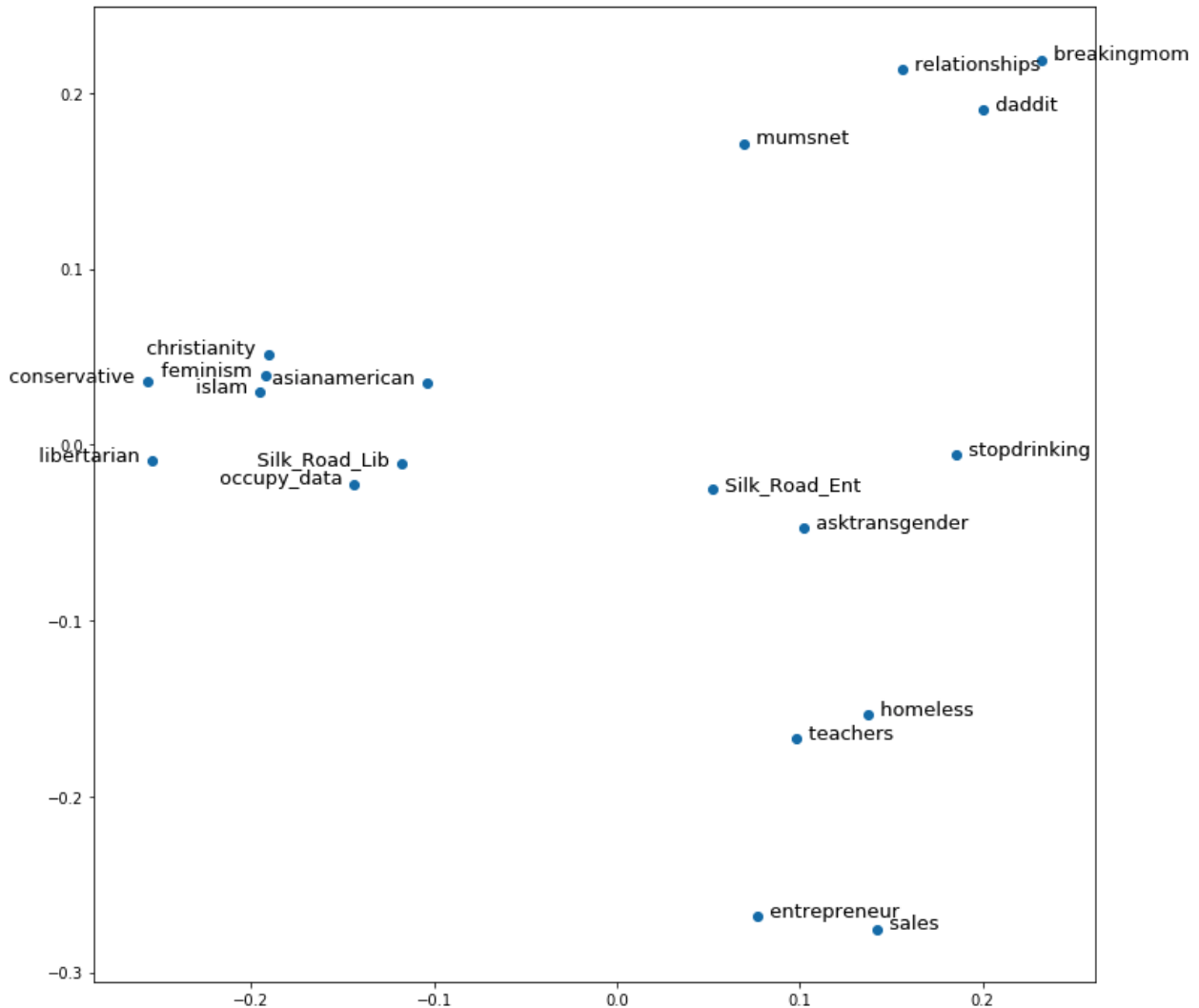
Figure 4.10

Shaded Dissimilarity Matrix of AUC Results Including the Four Additional Data Sources as Compared to 15 Reddit Identities
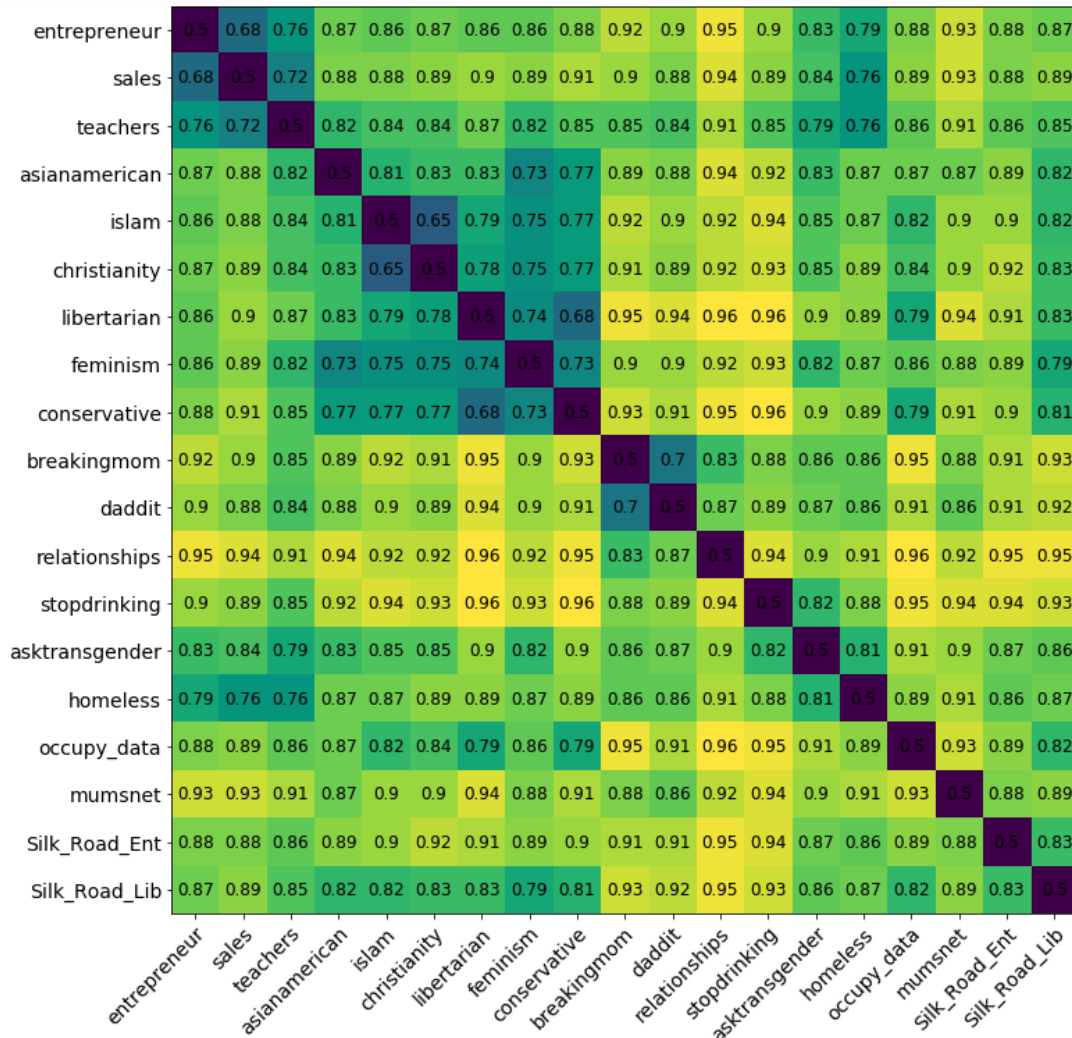
Using this dissimilarity matrix, we then repeated the MDS analysis as before. Once again, the Scree plot demonstrated that two dimensions best fit the data. The resultant MDS visualisation is displayed below in Figure 4.11.

Figure 4.11

Multidimensional Scaling Plot Illustrating the Relations Between the 15 Different

Social Identities Including Data From Four Additional Sources

As we can see from Figure 4.11, both the Occupy data taken from Smith et al. (2015a) as well as the Silk Road libertarian forum data appear closer to the collective action identities, as predicted. We also note that the Mumsnet data is positioned closely to the relational identities in line with our prediction. However, unexpectedly, the Silk Road entrepreneur forum data is further away from the vocational identities than predicted. Possible explanations for this finding are explored in the discussion.

4.4.4 Discussion

The results from this study demonstrated that the methodology was still mostly able to identify the 'type' of identity that a group had regardless of the source of the data. We observed that the Mumsnet forum data, the Silk Road libertarian data and the Occupy Wall Street Facebook data were located close to their predicted group clusters and were thus in line with our hypotheses. Conversely, we also noted that the Silk Road entrepreneur data was located further away from the vocational identities than we hypothesised. Reasons for this disparity are discussed below.

The position of the Mumsnet forum data and the Silk Road libertarian data indicate that this stylistic methodology is robust to platform-specific effects. This finding is particularly interesting as MDS analyses often tend to be quite unstable when relatively small changes are made to the data included in the analysis (Hout et al., 2013). This is because the output of MDS is derived through understanding the relations and distances between *all* the points included in the analysis. Thus, it is not uncommon for researchers to find that when they include more datapoints, the relations between individual points in the MDS analysis (in our case social identities), are significantly shifted. For example, if one were to conduct an MDS analysis that sought to model the differences between types of bird based on their features, an analysis that originally included distinct clusters of sparrows, starlings and pigeons, would completely change when penguins were added to the model. The inclusion of penguins would in turn make the sparrows, starlings and pigeons look much more similar to each other when compared to the very distinctive penguin. Here then, we can see how including new datapoints has the capability to completely change the structure of the MDS results. In light of this, it is therefore very encouraging to note the robustness of our MDS results even after including four new identities in the

analysis. This finding speaks to the validity of the analysis for understanding the position of new identities.

Notably however, the one additional identity which was not clustered in line with our predictions was the Silk Road entrepreneur identity. For the Silk Road entrepreneur identity, we observed more similarity with the stigmatised identities than with the vocational identities (as was predicted in Hypothesis 4). However, this finding is not completely out of place. Previous qualitative research from Masson and Bancroft (2018) has suggested that individuals who were using the Silk Road saw the Silk Road platform as a safe place where those with alternative lifestyles – those interested in the buying and selling of drugs – could discuss drug-related issues with likeminded individuals. As has been noted in their paper, the actors on the Silk Road claimed that they were often stigmatised from society for this socially unacceptable lifestyle choice and thus they relished in the sense of community provided on the Silk Road platform. As the Silk Road was the first darknet cryptomarket to exist, individuals were less aware of the security risks of using the platform and may thus have been more open about the stigma that they faced in choosing to buy and sell drugs. It is possible therefore that this more stigmatised identity is what we are seeing with our MDS analysis here. Again, we observe that by assessing behavioural prototypicality, this can provide information about a group that is not captured by the forum name ("vendor roundtable") or outgroup member's social categorisation ("entrepreneur").

Another interesting point to note is the position of the Occupy Wall Street Facebook forum data on the MDS plot. Specifically, the results indicate that the Facebook data from the Occupy group sits close to our collective action identities. The significance of this finding appears to validate the idea that whilst individuals

may induce specific local group prototypes and norms through communication with other group members (as demonstrated in Smith and colleagues' (2015a) analysis), here we provide evidence to suggest that their behaviour is also impacted by a wider understanding of the purpose, goals and normative way to express one's group identity. That is, whilst the members within the Occupy Facebook group may have been influenced by the individuality and ideas of those with whom they are communicating, we also note that their behaviour is not wholly malleable to influence from local interactions. Instead, we find evidence that their behaviour may also be shaped by a more generalised sense of being a political identity. Here then, we see the value in understanding the 'type' of group to which an identity may belong.

This finding therefore opens up exciting new possibilities to understand which aspects of group identities are more malleable to change, and which are influenced by a more global understanding of who the group is. This notion shows synthesis with the work of Koudenburg et al. (2019), who demonstrated the more complex relationship between global, superordinate identity norms and local group norms. More specifically, Koudenburg and colleagues (2019) demonstrated that when individuals with a shared superordinate identity communicate within local groups, they can both adopt new local norms or continue to conform with the superordinate normative viewpoint, regardless of the local group norm. In relation to our own stylistic research then, we suggest that stylistically prototypical behaviour at the category-level may be more robust to local group influences; it is more context-independent than local group attitudes. This would explain why local groups with similar identities still place close together on our MDS plot. Comparatively, the specific content of the norms, e.g., particular attitudes or opinions, may be more malleable to local group influence. Consequently then, whilst individuals within the

Occupy Facebook group were found to induce the aspects of their prototype which related to the action they were planning to take (Smith et al., 2015a), the behavioural evidence demonstrated in this study suggests that we are still able to detect the 'political' aspect of their identity. This finding thus opens up exciting new avenues for future research to consider the interplay of local and global group prototypes and their combined influences on behaviour. Specifically, we suggest that there may be both higher-order (global) as well as lower-order (local) influences on the way group members behave in intragroup communication.

In sum, this study has demonstrated compelling evidence in support of the idea that prototypical behaviour can be simultaneously locally constructed as well as influenced by a higher-order ingroup representation. By combining the results of this study with those of Smith et al. (2015a), we provide direct evidence to support the idea that real-world identities are impacted by both local and global identity processes. This has important implications for how we study online groups in future research. This idea will be explored further in the discussion of the thesis.

4.5   Study 3e: Identity evolution over long timeframes

In the final study of this chapter, we aimed to understand whether this framework could be used to study the evolution of groups and identities over time. Whilst in the previous discussion, we suggested that global identity prototypes may be more context-independent than local identity prototypes, we want to be clear in stating that this does not mean that the global identity prototype cannot change or evolve over time. Instead, these changes are hypothesised to be slower and thus occur over a period of years or decades as opposed to within immediate interaction. This is in line with the ideas of Pehrson and Reicher (2012) who suggest that 'there

is no reason why a fluid self-categorization process cannot produce stable self-categories to the extent that the context itself is stable and completely different sources of contextual stability are relevant at micro- and macro-contextual levels and timeframes.' (p. 116). In this way then, this study aims to understand whether the stylistic framework outlined in this chapter can be used to monitor the evolution of global group prototypes and prototypical behaviour over time, specifically when the social context is known to change.

In order to understand whether our methodological framework could be used to observe these subtle shifts over time, we chose to focus on the evolution of the transgender identity. As has been argued in Vox, 'few marginalised communities have experienced such a dramatic whiplash of fortunes over the course of the 2010s as trans people' (Burns, December 2019). Or, as stated in The Spectator, 'the 2010s were the decade of trans' (Emmons, December 2019). These mainstream news articles point to the major developments, both legally and culturally, that trans rights movements and transgender individuals have achieved over the past decade. We therefore believe that this makes the transgender identity a suitable case-study for understanding whether this MDS framework can be used to illustrate the evolution of identities over time.

Whilst the transgender rights movement has been in existence since the 1980s/90s (Stryker et al., 2008), the legalisation of same-sex marriages in the Western world during the early 21st century led LGBTQ+ movements to shift from focussing on gay rights, to focusing on transgender rights (Green, 2016; Taylor et al., 2018). For example, in 2016, Obama's Social Security Administration passed a law making it easier for transgender individuals to amend their passports and legal records (White House Press Office, 2016). In the UK, the Equality Act of 2010

declared gender reassignment as a protected characteristic, thus making it illegal to discriminate against transgender individuals in professional environments. Further, the World Health Organisation also withdrew transgender health issues from the ICD-11, proclaiming that transgender issues are no longer considered as mental or behavioural disorders (World Health Organization, 2019). Whilst not all rulings over the past decade have been in favour of transgender rights, for example the Trump administration's ban on transgender individuals serving in the military, it is clear that transgender rights and issues have gone from the political fringe to the political mainstream (Taylor et al., 2018). It is therefore possible that the transgender identity is becoming more political over time.

Similarly, the 2010s decade observed a greater visibility of trans individuals in mainstream media (Green, 2016). In June 2014, Time Magazine dedicated its front page to the 'Transgender Tipping Point', with the accompanying article discussing the growing awareness of trans rights (Steinmetz, 2014). Additionally, Olympic medallist Caitlyn Jenner appeared on the cover of Vanity Fair, after coming out as trans in 2015 (Bissinger, 2015). Further, whilst it is clear that transgender individuals have been more visible in the media since 2010, research by Austin and Goodman (2017) has noted that the increased media attention and political centrality of transgender individuals has led to a positive shift in the social and cultural attitudes towards trans people. Based on this historical context then, it appears that the transgender identity is slowly becoming more politicised whilst simultaneously becoming less stigmatised within mainstream society.

Despite this positive social shift, transgender individuals still report high levels of stigmatisation both during and following their transition (Bry et al., 2018; Miller & Grollman, 2015; Verbeek et al., 2020). As a result of this stigmatisation, the

prevalence of depression and anxiety in trans individuals exceeds that of the general population (Witcomb et al., 2018). Whilst there is a plethora of qualitative research which has focused on the stigmatising experience of transgender individuals within the past decade, it is difficult to assess whether there have been collective-level changes in the way transgender individuals perceive their own identities. Instead, statistical research has demonstrated a steady increase in the proportion of individuals identifying as transgender and non-binary in both Europe and the US (see Nolan, Kuhner & Dy, 2019 for a review). This increase in the number of individuals identifying as transgender may point to the idea that transgender identities are slowly becoming less stigmatised. This makes further sense when considering the greater visibility and positive shift in cultural attitudes towards transgender individuals (Austin & Goodman, 2017; Green, 2017).

4.5.1   Hypotheses

In this study therefore, we seek to explore this idea using our identity typology. Based on the increasing political success of the trans rights movement (Taylor et al., 2018) coupled with the increasing social acceptance of transgender individuals by the general population (Austin & Goodman, 2017), we hypothesize that we will be able to observe a shift in the prototypical behaviour of transgender individuals. Specifically, we suggest that the position of the transgender identity on the MDS plot will shift from being closer to the stigmatised identities at the beginning of 2011 towards the collective action identities by 2019. That is, we hypothesize that using only linguistic style features, we will be able to observe the gradual politicisation and de-stigmatisation of the transgender identity between the years 2011-2019.

4.5.2 Method

4.5.2.1   Data

In order to understand the evolution of the transgender identity over time, we

collected Reddit data dating back to 2011 using Google BigQuery. In line with our

previous analysis, we collected posts and comments from the subreddit

r/asktransgender, a forum consisting of transgender individuals discussing their

personal experiences and identity. We cleaned these data to remove any bots,

deleted authors, deleted posts or posts with less than 50 words. Following data

cleaning, the total number of posts and users included in the analysis is shown in

Table 4.6 below.

Table 4.6

Data Collected from the r/asktransgender Subreddit Over Nine Years

Year	No. of Reddit posts	No. of Users
2011	15,620	1,163
2012	31,823	2,536
2013	51,596	3,985
2014	84,294	6,315
2015	149,808	10,518
2016	183,092	12,426
2017	201,570	15,000
2018	207,589	18,740
2019[a]	177,139	21,035

[a] The 2019 data only includes posts up to October 2019 as this was the most recent

data available on Google BigQuery at the point of collation.

4.5.3   Analysis

4.5.3.1 Multidimensional Scaling Analysis

After running the data through LIWC software in order to quantify the linguistic style of each post (see Appendix A), we repeated the MDS analysis outlined in Study 3a nine times, each with a different year's worth of data from the transgender identity. In this way, all the other identities were held constant so that we could be sure that any movement of the position of the transgender identity on the MDS plot could be attributed to changes in the transgender identity specifically.

To repeat the MDS analysis, we first created a dissimilarity matrix between the yearly transgender sample and the other 14 identities. We took 4,136 training posts and 4,135 testing posts (in line with Study 3a) from the yearly transgender samples and computed the dissimilarity between the transgender identity and the other 14 identities using the Extra Trees classifier as before. Once again, we entered all 41 stylistic variables into the analysis (see Appendix A). We repeated this process twice using a different sample of 4,136 training posts and 4,135 testing posts so that we could calculate variance between the samples.

To complete the MDS analysis, we included only one year's worth of data at a time. The coordinates for each year's worth of data are plotted in Figure 4.12 below. As outlined in the legend, the colour of the transgender points refers to the date of the data. The redder the colour, the earlier the data.

Figure 4.12

Multidimensional Scaling Plot Illustrating the Evolution of the Transgender Identity

Over Time.
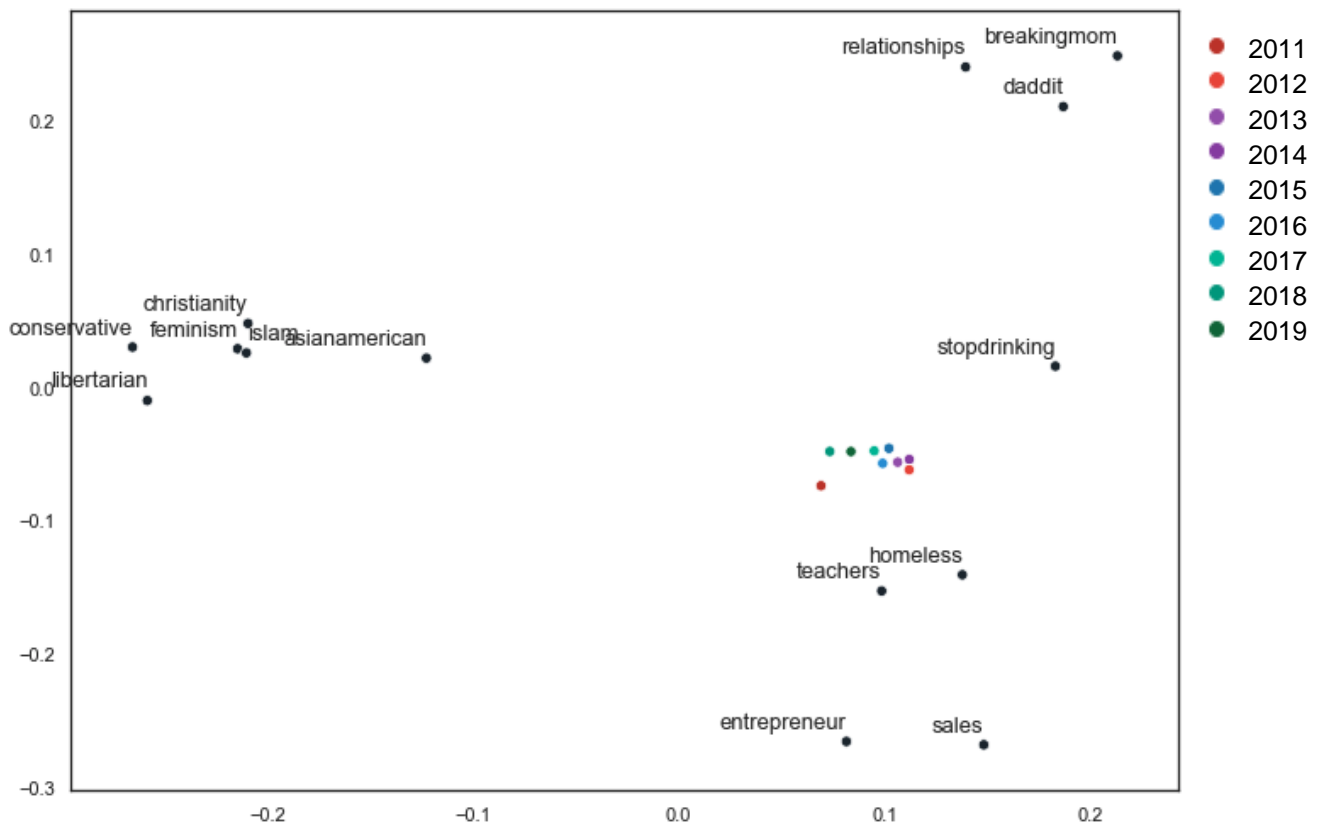
As we can see from Figure 4.12, there is clear evidence to suggest that the

transgender identity is 'moving' away from the more stigmatised identities, and

towards the collective action identities on the left side of the plot. Notably however,

2011 appears to be an anonymous year, and this may be because 2011 was when

the forum was first starting up. To understand the statistical significance of the

relationship between the position on the MDS plot and the year of posting, we

decided to run a correlational analysis.

### 4.5.2.3   Correlational Analysis

We extracted the MDS coordinates of each year's worth of transgender data

and ran a Spearman's correlational analysis in order to understand the correlation

between the year of the Reddit posts and the transgender identity's position on the

MDS plot. For each year, we had three sets of co-ordinates based on the resampling procedure. At first, we noted a significant correlation between the year of posting and the Y coordinates, $r(36) = .67$, $p = <.001$, (see Figure 4.13 below), however no significant correlation between the year of posting and the X coordinates, $r(36) = -.24$, $p = .16$, (see Figure 4.14 below). This finding suggested that the transgender identity was 'moving' towards the top the MDS plot, although was not moving towards the left side of the plot.

After visually inspecting the scatterplot of the X coordinates and the year of posting however, we identified that 2011 appeared to be an anomalous year (Figure 4.14 below). This may be because the year 2011 was when the forum was first established and thus the community was not developed at this point. However, we noted that when the 2011 data were excluded from the analysis, a strongly significant negative correlation was observed between the year of posting and the X coordinate, $r(32)=-.83$, $p = <.001$. This finding aligns with the visualisation demonstrated on the MDS plot in Figure 4.12.

Figure 4.13

Box and Whisker Plot Demonstrating the Positive Relationship Between the MDS Y

Coordinate and Year of Transgender Data
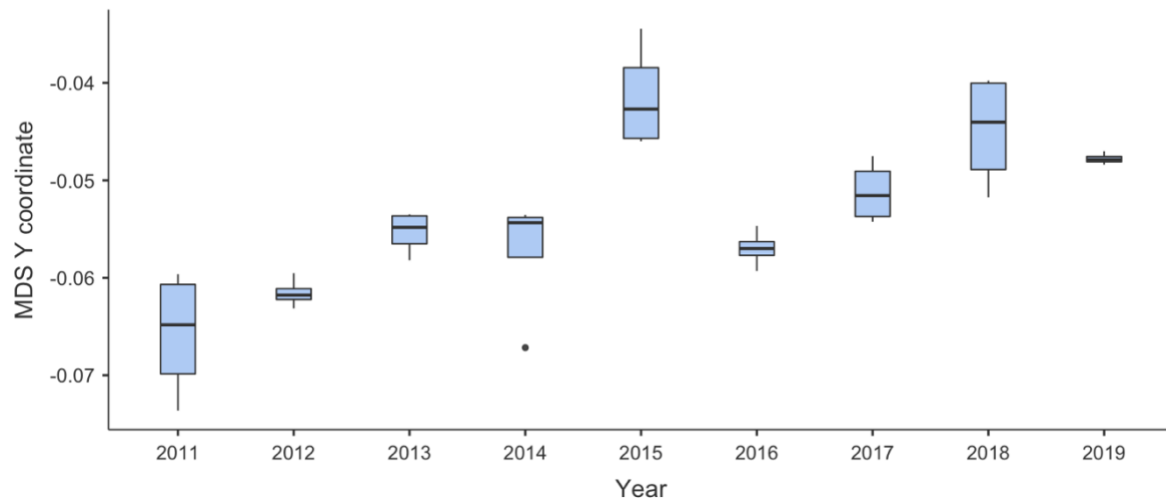
Figure 4.14

Box and Whisker Plot Demonstrating the Positive Relationship Between the MDS X

Coordinate and Year of Transgender Data
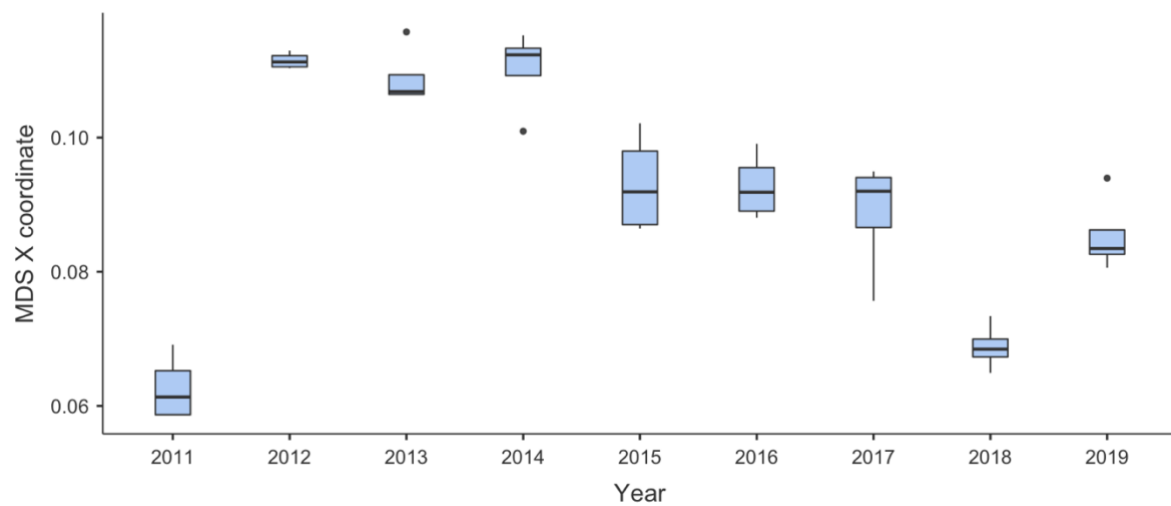
4.5.4   Discussion

The significant correlations between the MDS coordinates and the year of the Reddit data provides compelling evidence to suggest that the transgender identity is indeed becoming less stigmatised and more politicised over time, and moreover that this is detected in the way in which transgender individuals express their identity online. This slow evolution is clearly visualised in Figure 4.12.

This finding corresponds with both the success of the transgender rights movement during the last decade (Taylor et al., 2018), as well as the research demonstrating the positive shift in attitudes towards transgender individuals (Austin & Goodman, 2017). Whilst previous research has focused on outsiders' opinions towards transgender individuals (e.g., Austin & Goodman, 2017), here we can directly observe the shift in how transgender individuals experience and express their own group identity. Of course, this is not to say that transgender individuals do not experience stigma; as we can see, this change is very slow and despite this movement, the transgender identity still appears much closer to the stigmatised identities than the collective action groups. Nonetheless, the results of this study provide direct evidence to support the idea that when the societal context changes, so too does the prototype and thus prototypical behaviour of social identities (Pehrson & Reicher, 2012).

As outlined in the introduction, Pehrson and Reicher (2012) argue that 'there is no reason why a fluid self-categorization process cannot produce stable self-categories to the extent that the context itself is stable and completely different sources of contextual stability are relevant at micro- and macro-contextual levels and timeframes.' (p. 116). Whilst for the most part, the 'contextual stability' of self-categories has been studied at the micro level, for example through changing the

immediate comparative context (e.g., Haslam et al., 1995), here we provide direct evidence demonstrating how self-categories, and by extension prototypical behaviour, may change at a macro-level. Further, by using naturally occurring data, we are able to access longitudinal data with ease in order to study, observe and model how societal level changes lead to changes in individual and group-level behaviour. In turn, this research demonstrates an exciting new approach for studying the development of real-world identities over longer time frames than those which are typically studied within experimental and questionnaire research.

The results of this study thus demonstrate an exciting application for using multidimensional scaling approaches and linguistic style to understand identity evolution over time. We believe that this approach could therefore be used to understand how other identity's prototypes and prototypical behaviour may shift over time. For example, whilst much research has demonstrated the role of leaders in influencing what is seen as group prototypical of a particular identity (Reicher & Hopkins, 2003), this methodology may enable us to assess and study the speed at which social categories develop and change over longer time frames as a result of influential leaders. Moreover, by using naturally occurring online data, these research aims can be achieved in relatively inexpensive ways that are not impacted by the usual barriers to longitudinal research such as high rates of attrition. In addition to these advantages of using naturally occurring data to study longitudinal processes, it must also be noted that we are able to access historical data which can be used to understand present day social identities. In turn, this behavioural approach offers many exciting new research avenues.

In addition, we also suggest that this methodology may have applications for understanding the movement and politicisation of specific online communities over

time. For example, the initial Incel website (involuntary celibate), was originally started in the year 2000 for individuals who struggled to find loving relationships. According to its founder, it was a place where lonely individuals could come to moan about their inability to meet anyone (Jones, 2020). However, over time the Incel movement has become a politicised anti-feminist movement with misogyny at its centre (Jones, 2020). Some Incels have even taken their hatred offline resulting in acts of mass violence (Jones, 2020). We suppose that an automated methodology like the one proposed here may allow tracking of the politicisation of identities over time. Whilst at the moment our methodology is not refined to identify 'radicalisation' as such, this could be an interesting future research avenue to explore. Are radicalised political identities 'extreme' political identities that lie even further left on our MDS plot, or are they indistinguishable from our more mainstream political identities?

In sum, we contend that the exciting findings of this study demonstrate the value of combining social scientific theories with computational, automated textual analysis. Further, we feel that there is much more development to be completed within this interdisciplinary sphere of research.

4.6   Chapter Discussion

In this chapter, we have provided compelling evidence demonstrating that similarities in groups' linguistic style can be used to identify higher-order commonalities and patterns among the communication styles of different identities which may thus be indicative of a higher-order prototype. Specifically, we have shown the synthesis between the typology proposed by Deaux et al. (1995), where identities are grouped based on perceived similarities, and the behavioural typology

that arises through examining linguistic style differences from a diverse range of

social identities. Moreover, we have illustrated the robustness of this typology,

through establishing that it cannot be explained by demographic or personality

factors (Study 3b), and that it can incorporate groups from other online platforms

(Study 3d). Further, we have shown how this linguistic style typology of identities

relates to the values and emotions that group members hold (Study 3c). More

importantly, we have also illustrated that although group behaviour may be relatively

robust to the local context (Study 3d), global-level identity prototypes are not

deterministic or static (Study 3e). Instead, we provide evidence which suggests that

group prototypical behaviour changes over time in line with the wider social structure

(Pehrson & Reicher, 2012). In turn then, the methodology developed and validated in

this chapter provides a novel way of understanding how group level prototypes can

be understood and modelled through comparing similarities in group behaviour.

The findings outlined in this chapter support the idea that social identities are

psychological group memberships, as opposed to group memberships that rely on

the physical congregations of individuals (Turner et al., 1987). For example, here we

can begin to understand how the psychological idea of 'who the group is' is an

internal psychological construct or representation, as opposed to being built solely

upon specific interactions with other group members. Whilst this research has not

sought to address where this psychological representation comes from, it

nonetheless provides evidence to suggest that group members share a global idea

of what the social identity is (and thus the socially appropriate way to behave)

regardless of the immediate local group with which they are interacting.

The complex relationship between local groups and global identities has been

explored in research from Koudenburg et al. (2019). Koudenburg and colleagues

show that when individuals who share a superordinate identity are primed with a positive opinion of an outgroup, this positive opinion is less likely to be changed even when local group norms suggest that the group normative position is in fact negative. Conversely, when group members who share a superordinate identity are primed with a negative opinion of an outgroup, we see that they are more likely to update this opinion dependent on what the local group norm is. Thus, this research suggests that there may be two different levels of prototypicality (local and global) that both interact to influence group attitudes. Here then, we extend this idea by suggesting that global group identities may best be studied using behavioural indicators that are more robust to local group influence. Specifically, we demonstrate the value and robustness of using linguistic style indicators to access the global group prototypicality. This is in contrast to the way in which local group prototypicality has tended to be studied; primarily through understanding how local group communication influences the content of attitudes and opinions (cf. Postmes et al., 2000).

4.6.1   The Behavioural Typology

In addition to this, we also provide a behavioural typology which aligns strongly with the perceptual typology proposed by Deaux et al. (1995) suggesting that there may be different types of identity prototypes. Notably, we find that four groups may be best at explaining the variation in identity-prototypical behaviour, due to the fact that ethnic and religious groups are found to be very similar in their behaviour to our political identities. One explanation for the disparity between Deaux et al.'s (1995) five identity types and our four identity types, pertains to the perspective from which these types were understood. In our own research, we are using the behaviour of individuals to understand how they express and enact their own identities.

Conversely, in Deaux and colleagues' research (1995), the identities are socially categorised based on *perceived* similarities or differences between them. In turn, this perceptual approach may not take into consideration the lived experiences of the individuals who identify with these social identities in the same way that the behavioural approach demonstrated here does.

The difference between the findings of Deaux et al. (1995) and our own results, points to the value of understanding social identities directly in context as they are mobilised (Turner et al., 1990; Durrheim et al., 2016). In this research, we observed how identities are mobilised through intragroup discussion, without focusing on the explicit attitudinal or opinion-based norms of the identities that may be formed within the local context (e.g., Smith et al., 2015a). Instead, it could be argued that by using linguistic style, we are able to focus more on the process of 'being' the identity, rather than the process of 'becoming' the identity (Reicher, 2004). Thus, through using stylistic indicators, we are able to assess the more general sociopsychological perspective through which group members express themselves, as opposed to the conscious process of negotiating an identity through discussions about specific topics or opinions (Reicher, 2004). We therefore propose that there may be a difference between expressing and enacting a salient social identity and negotiating or constructing a social identity (Reicher, 2004; Smith et al., 2015a).

4.6.2   Linguistic Style

Whilst in this analysis we ascertained the relationship between the emotions and values of the social identities and their position on the MDS plot (Study 3c), we contend that there may also be other aspects of the global identity prototype which relate to the way in which our identities communicate. For example, research from Reysen et al. (2015), Reynolds et al. (2012) and Jenkins et al. (2012) suggests that

personality attributes could be used to understand the prototype of different groups. Specifically, Jenkins and colleagues (2012) propose that 'just as social identities have an associated set of attitudes, norms, and behaviours, personality may represent a reflection of the group's normative content' (p.11). This echoes the sentiment made by Turner and Onorato (1999) which suggests that the consistency in personality across a lifespan, may be due to the chronic salience of a particular (personal) identity. Further, research from Jenkins and colleagues (2012) has found that when different social identities are made salient, participants report personality traits that are less consistent with their personal identity, and instead more consistent with the salient social identity.

In addition to this, Reynolds et al. (2012) noted that individuals reported a change in neuroticism when the salience of a social identity was manipulated. In their research, they manipulated the salience of an Australian identity, an Australian National University student identity or a Non-Indigenous Australian identity and then asked individuals to complete the Big Five Inventory (John et al., 1991). They found that individuals with a Non-Indigenous Australian identity reported statistically higher levels of neuroticism than both those with a student identity salient, and those with an Australian identity salient. This was in line with their hypotheses that the Non-Indigenous Australian identity would elicit feelings of guilt, hopelessness and shame due to the negative intergroup history between Indigenous and non-Indigenous Australians. Here, then, it can be argued that the prototype of the non-Indigenous Australian identity impacted the way individuals reported their personality traits. Similar research has replicated the impact of identity salience on reported personality traits in other social identities such as sports fans (Reysen et al., 2015).

In line with this idea, research from psycholinguistics has demonstrated that personality traits can be detected using linguistic style analysis (Boyd, et al., 2017; Mairesse et al., 2007; Pennebaker et al., 2005). It is therefore possible that the differences between different social identities' linguistic style may relate to the 'personality' of their identity prototypes. This is an interesting idea and may be worth exploring in more detail. From a theoretical viewpoint, the link between social identity salience and personality remains relatively understudied, and so this may provide a novel method of furthering this research using naturally occurring data. Similarly, inferring individual-level personality traits using naturally occurring linguistic data is a notably large domain (see Boyd et al., 2020 for a review) and thus the inclusion or appreciation of social identity salience as possible confound for assessing consistency over different social contexts may prove a valuable addition to this research field.

6.4.3 Limitations and Future Research

It must be noted, that at present we do not know whether individuals' prototypical behaviour is in line with a specific social identity (e.g., conservative) or in line only with the type of group to which the identity belongs (e.g., political). Thus, whilst we have demonstrated between group-type differences and shown that identities within each group type communicate in a similar style than those outside of the group-type, the next step is to fully explore whether linguistic style can be used to map similarities in prototypes *within* group clusters. Thus, if we were to isolate the behaviour of, for example, ten political groups, would the patterns in their linguistic style relate to similarities in their group's prototypes? Or, is linguistic style too broad a behavioural indicator to reveal much about the specifics of the social identity content? At present, that answer remains relatively unclear due to the fact that we

cannot be sure exactly what is being expressed using linguistic style. With a greater understanding of which specific aspects of the prototype are being communicated through linguistic style, we can begin to answer this question in order to develop more nuanced applications for this methodology.

4.6.4   Concluding remarks

In sum then, this chapter has demonstrated how we can use similarities in linguistic style behaviour to better understand identity prototypes. Specifically, we noted that the identity typology proposed by Deaux et al. (1995) was remarkably similar to the behavioural typology presented here. We also acknowledged where the typologies differ (i.e., with regards to ethnic and religious identities) and suggested that this can be explained by the difference in stereotypical categorisation and prototypical behaviour. Finally, we suggest that these results also provide further evidence to suggest that certain aspects of the group prototype are more context-independent. Specifically, we noted that whilst globally prototypical behaviour may be more robust to local changes (e.g., domain, Study 3d), it may change over time based on macrolevel societal changes. In this way, changes in global group prototypical behaviour may be said to reflect the structure of social relations (Perhson & Reicher, 2012).

5   General Discussion

In this discussion, I will first provide a brief overview of the aims and purpose of this thesis as outlined in Chapter 1 and will then go on to outline how each empirical chapter directly addresses these aims. I will then highlight the theoretical significance of the results from each chapter when taken together. Specifically, I will argue that the research outlined in this thesis has clear implications for how we study identity prototypicality, specifically with regards to the more context-independent and context-dependent features of prototypes.

In the second half of this discussion, I will explore how the observational behavioural approach to studying social identities in naturally occurring data allows social psychologists a new lens through which to consider the impacts of social psychological phenomena. In turn, I will outline the value of using behavioural approaches, as well as some of the limitations. These limitations arise from both difficulties in operationalising the subject matter, as well as from the methodological constraints of using observations and computational techniques to draw psychological inferences. Finally, I will go on to suggest how the behavioural approach taken within this thesis can be developed in order to further our understanding of how salient social identities impact behaviour in real-world situations.

5.1   Overview of Thesis Aims

In Chapter 1 of this thesis, I outlined that identity prototypes are defined and studied at both the local level, as well as a more global level dependent on the definition of a prototype adhered to. Specifically, I noted that Turner (Turner et al., 1987; Turner, 1991), refers to prototypes and prototypicality in a very localised sense. He suggests that prototypes are synonymous with the prototypical position,

and that prototypes therefore arise directly out of interactions with individuals in very specific contexts (Turner, 1991). Conversely, I noted that much of the social identity theory of leadership literature (e.g., Hogg et al., 2012), tends to define and study prototypes as pre-existing cognitive representations of an identity which comprise the attributes perceived to define the identity and differentiate it from others. Further, I acknowledged that this definition of a prototype taken from Hogg's research (Hogg, 2001; Hogg & Reid, 2006; Hogg & Rinella, 2018), shows great similarity with the idea of 'self-stereotypes' as considered by Turner (1991). Either way, it appears that both researchers agree that a cognitive group representation which comprises group-defining attributes influences behaviour in such a way that shared group identities lead to uniformities in behaviour.

In this thesis then, I aimed to explore these uniformities in group behaviour using naturally-occurring online linguistic data. Specifically, I was interested in exploring the effect of social identity salience on prototypical behaviour which transcended the local group. That is, I wished to test the hypothesis that individuals with shared social identities communicated in similar ways regardless of the platform of their communication. Further, I also aimed to explore how computational methods and behavioural analysis could be used to test traditional self-categorisation hypotheses such as the relationship between prototypicality and influence. Finally, I aimed to explore what we could learn from the similarities between different groups' prototypical behaviour. Below I will present an overview of the empirical findings in relation to the thesis aims.

5.2 Overview of Empirical Findings

In the first empirical chapter of this thesis (Chapter 2), I explored whether we could use linguistic style indicators and machine learning classifiers to detect the

identity-prototypical communication style of two distinct social identities. The first study of this chapter indicated that there were in fact differences in the ways that libertarians and entrepreneurs communicated online. However, at this stage, it was not clear whether these findings were a result of social identity processes, or whether they were instead simply evidence of local group norms. In the following five studies of the chapter, I controlled for various alternative reasons for why the libertarians and entrepreneurs in my first study may have been communicating using different linguistic styles. I therefore ran studies controlling for demographic and individual-level factors (Study 1b, 1e and 1f), local group norms (Study 1c, 1e and 1f), topic of discussion (Studies 1d, 1e and 1f) and purpose of communication (Study 1e and 1f).

The overall results from this chapter demonstrated that when a libertarian or entrepreneur identity is salient, group members behaviourally assimilate to the prototypical way of communicating, regardless of the specific domain of communication. Thus, in line with the first aim of the thesis, this finding demonstrated that individuals with the same identity showed similarities in group behaviour regardless of the local context. Specifically, we observed that we could classify group prototypical behaviour, regardless of whether the behaviour was in a different forum (Study 1c) or in an online experiment (Study 1e). In Study 1f, however, we noted that motivation appears to play a key role in determining whether the salience of a social identity leads to stylistically prototypical behaviour. By forcing individuals to write a certain number of words, we found that the style of communication was less likely to be group prototypical than when the respondents were simply asked to write a few sentences (Study 1e). Here then, we observed the key role of motivation in predicting identity-prototypical behaviour.

Similarly, in Study 1f, I looked to concurrently validate our measure of behavioural prototypicality through comparing behavioural prototypicality to pre-existing measures of self-reported prototypicality. The results of this study indicated that the relationship between how prototypical individuals perceive themselves to be and how prototypical their behaviour actually is, is not straightforward. Interestingly, we observed a positive correlation between self-reported libertarian prototypicality and behavioural libertarian prototypicality, however we also observed a close to significant negative correlation between self-reported entrepreneur prototypicality and behavioural entrepreneur prototypicality. This finding thus suggested that perceptions of prototypicality may be conceptually distinct from behavioural prototypicality. Consequently, this negative finding served to reiterate the importance of studying behavioural prototypicality; with it being hypothesized that prototypes impact both perceptions and behaviours, this finding underlined the importance of not relying purely on perceptions to study group prototypes. It also reiterated the need to be cautious in interpreting what the behavioural data actually indicate.

In Chapter 3, we looked to explore whether there was evidence to suggest that other individuals recognise the prototypical style of ingroup members. In the first study (Study 2a) we applied the Reddit-trained classifier validated in Chapter 2 to understand the relationship between behavioural prototypicality and ability to generate responses from others on Reddit. The results from this study indicated that there was, in fact, a small positive correlation between prototypicality and response generation. However, as the primary aim of this thesis was to understand globally prototypical behaviour as opposed to just locally prototypical behaviour, we decided to control for the influence of local group norms by training a classifier on a different platform. Specifically, we trained a classifier on the Silk Road platform and applied

this classifier to our Reddit data. This allowed us to develop behavioural prototypicality scores which were not confounded by possible local group norms. Using this approach, we noted a very small correlation between behavioural prototypicality and response generation online. Whilst this finding was statistically significant, it is difficult to draw strong conclusions from these results. In turn, the primary takeaway from this chapter was that combining computational methodologies with sociopsychological theory may enable us to advance beyond describing individuals as influential because of their influence alone. Instead, by using sociopsychological theory, this can allow us to understand exactly why some individuals are more likely to become influential than others. Nonetheless, more refined research is needed to explore the link between behavioural prototypicality and influence online.

Finally, the fourth chapter of the thesis explored whether we could use the similarity in different groups' prototypical behaviour to better understand the similarities between group identities. Utilising a diverse array of social identities, we were able to demonstrate remarkable similarity between stereotypical perceptions of groups, and how these groups enact their identities online. Specifically, using multidimensional scaling and cluster analysis, we noted that there were four or five different types of social identity which aligned with the five types proposed by Deaux et al. (1995). These were vocational identities, stigmatised identities, relational identities and collective action identities (political and religious/ethnic identities). This pointed to the idea that there are perhaps higher-order group prototypes that individuals may use to guide their behaviour.

In the following four studies of Chapter 4, we explored the robustness of this behavioural typology. Firstly, we used a within-individual design to understand

whether the typology could be explained by demographic or individual-level differences (Study 3b). We found that even when using within-individual data, our typology was still robust as the data still clustered into four different group types. We then explored how the position of the identities on our MDS plot correlated to differences in each group's values and emotions (Study 3c). The results of this analysis proved to have high face validity and showed synthesis with much existing literature regarding the significance of emotions in collective action groups (Thomas et al., 2009), as well as the value of achievement in vocational groups (Deaux et al., 1995). Next, we tested the robustness of our model through adding four new identities to the model to ascertain whether the addition of new identities would change the structure of the typology (Study 3d). Interestingly, we observed that the typology was robust to the introduction of groups from different online sources and also noted that three out of the four identities clustered with our predicted group types. This served to further demonstrate how this methodological approach could be used to explore identities whose prototypes may not be well known or predictable. It also served to demonstrate that the similarities in group prototypical behaviour that we are observing are not platform-specific. This speaks to the idea that linguistic style enables us to measure something psychological which appears to transcend the local group context.

Finally, in the last study of the thesis (Study 3e), we demonstrated that whilst group prototypical behaviour at the global level is robust to changes in the local context, this does not mean that what is prototypical for an identity cannot or does not evolve over time. Instead, we demonstrated that global group prototypical behaviour can be seen to change predictably in line with changes in the social structure. Specifically, we showed the very strong correlation between the position of

the transgender identity on the MDS plot, and the year of posting. This correlation demonstrated that between the years 2012 and 2019, the transgender identity moved towards the collective action identities and away from the more stigmatised identities. More importantly however, this finding demonstrated the value of this linguistic style approach for studying identity-prototypical behavioural change over longer time frames.

In sum, the empirical results outlined in this thesis provide evidence to suggest that it is possible to model and map uniformities in group behaviour which result from shared social identities. Moreover, the results also suggest that linguistic style is a suitable behavioural indicator that can be used to study global group prototypical behaviour.

5.3   Theoretical Significance of Empirical Findings

The empirical findings outlined above have several theoretical implications for how we understand social identities and their enactment in the real-world. These findings can be summed up in two major theoretical contributions:

i.   Group-based behaviour may be impacted by both local group norms and direct interactions, as well as from an overarching sense of 'who we are'. In this way, it is necessary to distinguish between globally prototypical behaviour and locally prototypical behaviour.

ii.   There is a unique value to understanding and modelling behavioural prototypicality in addition to the research on perceptions of prototypicality as is commonly studied.

I will discuss these two theoretical contributions in more detail below.

5.3.1   Globally Prototypical Behaviour

In the first empirical chapter, we demonstrated that individuals behave in group prototypical ways regardless of the immediate local context. That is, using a machine learning classifier we were able to detect how libertarians and entrepreneurs express their social identities both online and in experimental conditions. This finding supports the idea that group members' behaviour is impacted by the salience of their social identity and moreover that this leads group members to behave in line with the norms of the group (Koschate et al., 2021). Specifically, this behavioural uniformity can be seen regardless of the local group context.

In turn then, it makes sense to propose a distinction between local group prototypicality as studied through polarisation and other interaction-based research (e.g., Postmes et al., 2000; Mackie, 1986; Turner, 1991), and a more context-independent prototypicality which transcends the immediate context and can be observed even when individuals are not directly interacting with others (Study 1e; Koschate et al., 2021). Specifically, I argue that uniformities in group behaviour are impacted by both levels of prototypicality; individuals may accommodate to other group members they see as valid sources of social information (Gallois et al., 2016), but they may also show uniformities in behaviour with other group members even when those group members are not present.

This idea shows synthesis with Postmes et al.'s (2005b) model which suggests that in small groups, individuals may either deduce their identity from a pre-existing understanding of the group or they may induce it locally through interactions. In their model, Postmes et al. (2005b) suggest that deduction is a top-down process whereby identities are formed and influenced through an understanding of the 'supra-ordinate social realities existing in the social structure' (p. 8). Thus, the

content of an identity is established through both implicit and explicit comparison with other groups that exist within the social structure. In this thesis, we have shone a light on the 'implicit social structure' that exists even when groups are communicating in the absence of a direct outgroup or explicit comparative context. In Chapter 4, we compared the similarities of 15 diverse social identities, and found a style of communication that corresponded to each of the five types of identity proposed by the perceptual typology of identities created by Deaux et al. (1995). Whilst Deaux et al. (1995) determined that individuals perceive there to be five different perceptual types of social identity, we have demonstrated that these different types of group appear to behave in distinct ways that separates them from other group types (cf. ethnic/religious and political identities). In turn then, we suggest that the results from Chapter 4 serve to empirically demonstrate the implicit social structure by indicating how the similarities between groups can be mapped out using behavioural data.

Further, through using the behavioural typology constructed through comparing linguistic style similarities between groups, we were also able to understand how groups without a pre-conceived category label identity (e.g., #Occupy Facebook group, Smith et al., 2015a) still behave in line with their wider group type. That is, even though the individuals within the Occupy Facebook group had yet to develop the specific actions and beliefs of their social identity, we were still able to detect the general collective action style of the group. This finding suggests that individuals are likely to have an approximate understanding of their group, possibly a shared purpose, goal or frustration, even when the group is relatively young. In turn, this shared understanding of who the group is determines how they negotiate their identities. For example, whilst the majority of individuals who participated in the

Occupy Wall Street protests were under 35 (Milkman, 2014), we could determine from the way in which they expressed and enacted their shared identity that this group was not merely a shared demographic identity but was instead a group coming together to form a social movement. In this way then, the individuals actively demonstrated the prototypically relevant features of their identities through enactment via their communication style. This shows synthesis with the ideas of self-categorisation theory whereby the features of groups are not enough to define the prototype, instead it is the meaning ascribed to them within context (Turner, 1990, 1991). Here, we observe that individuals are active participants in creating and constructing the meaning of their identities through the way in which they behave. Commonly, this idea is studied in communication content (i.e., using rhetorical or discursive approaches, Billig, 1996), but here we show the additional role that linguistic style plays in identity enactment.

Following on from this argument, we suggest that the style of communication is used by individuals to directly enact and express an identity, and that this identity expression is conceptually distinct from negotiating the content of an identity. This idea shows synthesis with Reicher's distinction between 'being' an identity, and 'becoming' an identity (2001). Reicher suggests that 'being' helps us to define our position within the social structure, whereas 'becoming' refers to the notion that identity can mobilise us and be shaped within interaction. Consequently, it is argued that 'being' may be deduced from 'who we are' whereas 'becoming' is induced from 'who we want to be'. In relation to this idea then, the results presented in this thesis suggest that linguistic style allows us a lens through which to study 'being' whereas content analysis, such as that of Smith et al. (2015a), provides us with a lens to study 'becoming'. Salient social identities may therefore impact what individuals do

and how they do it (linguistic style) but in addition to this, social identities can also be shaped and moulded directly through the content of interactions with others (linguistic content).

In sum, in this section we have argued that the evidence presented in this thesis points to the need to distinguish between global and local behavioural influences. We have indicated that the uniformities in behaviour demonstrated in this thesis appear to transcend local group norms, and thus we suggest that they may be indicative of a more globalised notion of identity prototypicality. In section ___ we explore other possible explanations for our empirical results.

5.3.2 Perceptual versus Behavioural Measures of Identity

The final theoretical contribution of this thesis pertains to the value in studying prototypical behaviour directly. As outlined in the literature review, the vast majority of research into prototypes and prototypicality has focused on how individuals perceive their own group prototypes or the process of depersonalisation whereby self and others are compared to this abstract representation of the group's prototype. However, in this thesis we have demonstrated a methodology that can be used to assess how prototypical behaviour is enacted in the real-world. Whilst there is a plethora of research which has sought to identify how perceptions of prototypicality impact behaviours such as ingroup bias, updating of opinions and responses to group threat (Hogg & Reid, 2006), as far as we are aware, there are very few studies that have aimed to create one behavioural measure that is able to classify a wide variety of different social groups.

5.4 The Behavioural Approach to Studying Social Identities

Within this thesis, we have primarily relied upon observational behavioural analysis in order to explore how individuals with salient social identities behave in

real-world scenarios. In doing so, this has enabled us to explore aspects of social identity theory which have previously only been studied within controlled experimental conditions. However, we contend that by studying social identities 'in the wild', this can provide us with knowledge that is harder to ascertain in rigorously controlled conditions. However, at the same time, there are some limitations to this approach which must be noted. We will address both the value and limitations associated with behavioural approaches below.

5.4.1   The Value of Observational Behavioural Research

In this thesis, the majority of the conclusions that we have drawn have come from naturally occurring behavioural data. That is, we have observed behaviour, used sociopsychological theory to make hypotheses as to why individuals behave in the way that they do, and then designed further behavioural analyses to support or reject our theoretical hypotheses. With the exception of the final two studies in Chapter 2 (Study 1e and Study 1f), the data have originated solely from online forums. This behavioural approach has enabled us to directly ascertain whether the core tenets of social identity theory can be used to empirically understand group behaviour in the real world.

As noted by Doliński (2018) amongst others (e.g., Baumeister et al., 2007), whilst psychology is defined as the science of behaviour, it is increasingly rare for behaviours to be the primary object of study. In fact, Baumeister et al. (2007) suggest that the primary behaviours studied by social psychologists in particular, refer to those which are "almost always performed in a seated position, usually seated in front of a computer. Finger movements, as in keystrokes and pencil marks, constitute the vast majority of human actions" (p. 397). The authors note that the exception to this statement is in areas of psychology such as animal behaviour or

development psychology, whereby subjects' behaviour is directly observed, likely because the subjects involved in these studies are simply less capable of completing surveys or using computers. Whilst Baumeister and colleagues' paper is now over a decade old, the more recent article from Doliński (2018) indicates that in the then-most recent volume of the Journal of Personality and Social Psychology (volume 113, comprising articles from the latter half of 2017) there were only 18 studies (out of a possible 290) whereby the dependent variable constituted real behaviour. A possible reason that Doliński provides to explain social psychologists' aversion to studying behaviour directly, is the fact that behaviours often constitute binary outcomes; a person does something, or they do not. In turn then, Doliński argues that 'This dichotomous character of the dependent variable excludes the application of many refined statistical analysis techniques' (p.10).

However, as demonstrated in our own research, there are ways that one can study real-world behaviours without being limited to dichotomous variables. Moreover, with the increase in time spent using digital technologies, it seems that this is the ideal time for social psychologists to branch away from a reliance on self-report survey methodologies and contrived experimental behaviours, and instead to develop new behavioural techniques that aim to understand not just the cognitive components that underpin group related outcomes, but also the specific conditions under which certain behaviours are enacted.

The value of observing behaviour directly can be seen in our own research. It has long been hypothesized that individuals become depersonalised when they self-categorise with a salient social identity (Hogg & Turner, 1992; Turner et al., 1987), however the direct impact that this depersonalisation has on behaviour has remained understudied. Whereas stereotypes have been treated as perceptual tools used to

make sense of social reality (McGarty, 1999), by observing real-world behaviour we are able to access directly when and why individuals categorise with particular labels. For example, in Study 3e, we showed that whilst the stigmatising component of the transgender identity may have been the primary defining attribute, and thus that individuals may have congregated in online forums to discuss the shared grievances and struggles of their common social identity, by observing their behaviour directly we are able to detect a shift in the reason for their continued congregation. Specifically, the results of Study 3e demonstrated that transgender individuals may now be congregating in online forums with a more pronounced aim of mobilising their stigmatised identity into action.

Further, the broader findings of Chapter 4 point to the value in comparing layperson perceptions and actual group behaviours. By studying the behaviour of particular groups directly and comparing this behavioural prototypicality to the layperson perspective of group categories as developed in Deaux et al.'s (1995) research, we were able to identify both similarities and differences between outgroup perceptions and ingroup behaviour. Specifically, we noted that the ethnic and religious identities expressed themselves in ways that were very similar to the political identities within our analysis. This acknowledgement of the similarities between these identities brings into focus the distinction between the outgroup perception of a group and the ingroup lived experience of a category label. More specifically, it suggests that there is a discrepancy between stereotypes as tools of categorisation (McGarty et al., 2002), and prototypes as guides of identity-relevant behaviour (Hogg, 2001). For example, ethnic labels are often used by outgroup members in order to differentiate a minority from the majority class. In this way, these category labels serve an explanatory purpose, and are used by perceivers to

make sense of the differences between groups (McGarty et al., 2002). However, the simple act of providing a label to a minority (ethnic) group may in fact serve to reinforce the political tensions that exist. Thus, whilst a white student partaking in Deaux et al.'s (1995) original study may have used the term Asian American to draw distinctions between their categories of White Americans and Asian Americans, the existence of this category label is in itself an act of political power. Consequently, it followed that the prototypical style of the Asian American identity was notably close to the political identities. Here, we see the particular importance of whose point of view we are assessing a 'group representation' from. The prototype of an identity and the lived experience of the individuals under a category label may be notably distinct from the category label *given* to a group of people by virtue of their perceived differences to others within the social structure.

This argument brings us round to the importance and value of studying behaviour when assessing prototypes and prototypicality. Through studying intragroup behaviour in real-world environments, we are able to ascertain not only how a group perceives its own identity, but the direct purpose of that identity. Whilst in experiments we can ask participants to report details about ingroup biases or stereotypical traits, this approach to the study of social identities struggles to directly appreciate the *point* of categorising oneself with a category label. As argued by McGarty (1999; McGarty et al., 2002), the process of social categorisation serves as a response to help explain the differences between sets of stimuli (people or otherwise). Any collective of instances can be categorised in a variety of different ways owing to the perceived explanatory ability of the categorisation to understand the situation (Oakes & Turner, 1990). At the same time, social categorisation gives stimuli meaning, and thus produces knowledge about the differences between the

stimuli (McGarty et al., 2002; Oakes & Turner, 1990). However, the apparent explanatory value of social categories concentrates primarily on the value and purpose of categorisations from the perspective of an observer (e.g., McGarty et al., 2002). Conversely, when the object being perceived relates to self, the qualitative nature and purpose of the categorisation is unique (Simon, 1993). That is, whilst explaining differences between groups may be the primary motivation for developing consensual stereotypes in social categorisation theory, the development of consensual ingroup prototypes likely serves a different purpose altogether.

In the studies presented within this thesis, we rely on the behavioural data of individuals who have actively come together under a chosen category label. That is, we are using data from individuals who have explicitly searched for a particular identity-labelled forum (e.g., entrepreneur), and then participated in its discussion. This is a notably different perspective of identity categorisation than an observer using social categories to simplify and make sense of their reality (McGarty et al., 2002). Similarly, it is a notably different perspective from which to understand the prototype than the perspective employed in controlled experimental conditions (e.g., Haslam et al., 1999; Hogg & Hardie, 1996). Here, we are seeing directly how social groups express, enact and 'do' their identity. Consequently, by studying identity-prototypical behaviour in naturally occurring contexts, we can access a better idea of what an identity is actually constructed for and why it exists (Tajfel, 1991). Whilst psychology's 'ecological validity crisis' has received much attention (e.g., Stokols et al., 2009), there is nonetheless a dearth of research which allows individuals to explicitly label their own social identities (Durrheim et al., 2016). In this way then, whilst the fit hypothesis may explain *when* an identity becomes salient in a relatively uncontrolled manner (e.g., when the identity fits the situation and can be used to

make sense of differences between groups of people), this hypothesis does not take into account how identities are experienced when individuals *make* the situation fit the identity, i.e., when an individual actively seeks out those with which they share a sense of reality with. This conscious desire to seek out one's ingroup members is lost within experimental research, and therefore the meaning and purpose imbued to social categories is fundamentally different (Deaux, 1992).

The need to pay direct attention to the motivation behind self-ascribed category labels was pointed out twice within this thesis. The first time the motivation of the identity on prototypical behaviour was seen to matter was in the first chapter where we attempted to externally motivate participants to express their identities. The results of Study 1f suggested that even though individuals identified with our groups of interest, they did not behave how we had seen other group members behave due to the fact that we had unintentionally removed their motivation to express their identities. That is, participants were not writing to enact their identities, they were instead writing to fill a minimum character limit. In much the same way, the ethnic and religious groups included in the MDS analysis throughout Chapter 4 were communicating purposefully in a way that may not have been predicted by the individuals distinguishing between groups in Deaux et al.'s (1995) study. Instead, they were writing for a purpose. This finding elucidates the distinction between a category label as an explanatory tool to make sense of the similarities and differences between groups (McGarty et al., 2002), and a category label as a functional tool for achieving an identity goal. Herein lies the value of using naturally occurring data to directly understand how social identities impact behaviour. Here we capture not only the abstract 'traits' of an identity, but also the reason for its

existence in the first place; we are explicitly able to capture the group members 'being' their identity.

5.4.2 Limitations of Behavioural Research

Whilst I have argued that the observational behavioural research outlined in this thesis allows us to directly understand how individuals come to ascribe meaning to their category labels, there are nonetheless certain limitations with relying predominantly on observed behaviour as a data source. Firstly, we are unable to control for a variety of variables that are hypothesized to have significant roles in understanding group behaviour. For example, we do not know whether all the individuals who contribute to these online forums identify with the category label, and if they do, we do not know the strength of that identification. Both of these factors have been hypothesized to have an important role in predicting identity-prototypical behaviour (Steffens et al., 2015) and thus the lack of knowledge surrounding these issues may constitute a significant limitation of the research. However, these limitations are mitigated in two ways; i) the scale of the data, and ii) the combination of offline studies in order to support the online conclusions.

Firstly, by using a large sample size and a large volume of data, we are able to understand the group or forum on aggregate. In this way, whilst we cannot be sure that all members who are participating in the forum identify with the category label, instead we can observe group behaviour at an aggregate level. This, of course, is not a perfect solution to the problem and thus another way to mitigate the possibility that we are drawing erroneous conclusions is to use experimental studies in order to confirm the causal role of identity in leading to stylistically prototypical behaviour (Study 1e and Study 1f). In this way, the combination of offline and online research allows us to identify both the causal pathway to identity-prototypical behaviour, whilst

also understanding how behaviour in the real-world can be understood and studied. By relying on both naturally occurring data and experimental data, we are therefore able to contribute not only to theory, but also to the creation of tools and methods that can be used to study social identities outside of the laboratory.

Consequently, throughout this thesis, I have aimed to design studies to prove myself wrong. I have repeatedly tested the idea that the reason that groups communicate in different styles is due to something unrelated to their social identity. However, when the findings of this thesis are taken together, we can see that this group-level behaviour cannot be explained by individual level factors such as personality or demographics (Study 1b, Study 3b), or post-level factors such as platform or topic (Study 1c, Study 1e, Study 1f and Study 3c) and not only does the group behaviour transcend local group norms (Study 1c, Study 1e & Study 1f), but it does so in a predictable way (Study 3a). In addition to this, we have demonstrated that the wider social context plays a key role in understanding this group behaviour (Study 3e), and have also demonstrated that there is a predictable, albeit small, link between stylistic behaviour and whether posts are responded to (Study 2a). In turn, all of these findings appear to confirm the core tenet of self-categorisation theory, namely that psychological group membership predictably impacts the way that individuals behave when a particular social identity is salient. However, as this section is titled limitations, I will now point out some methodological improvements that could have been made to this thesis.

Firstly, throughout this thesis I have suggested that the findings relate to the salience of a particular identity. Thus, I make the assumption that an individual commenting in an online forum has the respective identity salient. This idea stems from Koschate and colleagues' 2021 paper, which demonstrates that the linguistic

style of individuals commenting in forums is comparable to the linguistic style of individuals writing texts in isolation when the salience of a social identity has been experimentally manipulated. The results from Study 1e in this thesis show synthesis with Koschate et al.'s results, however the results from our Study 1f appear less convincing. In turn then, it is possible that the findings of this thesis are not related to social identity salience as such but may instead be related to other factors such as impression management (Klein et al., 2007), or perhaps simply linked to the purpose of communication within online forums. If this is the case, it would suggest that the behavioural prototypicality empirically observed within this thesis is not a result of shared identity salience as such, but perhaps simply a shared understanding of what the purpose of the group or forum is. That is, the uniformities in behaviour are linked more closely to the performance of an identity as opposed to the more subtle changes in self-perception.

However, at the same time, the use of forum as a proxy for identity salience is not unjustified. Previous studies have used changes in the topic of a conversation or interaction in order to manipulate the salience of an identity. For example, Lea et al. (2007) used discussion topics to manipulate the salience of a national identity as compared to a personal identity. Further, if it were the case that forum label was not a strong proxy for identity salience, we would not find evidence for individuals with the same salient identities communicating in similar patterns. That is, we would not find similarities in behaviour across platforms as shown in Studies 1c and 1e. Nonetheless, in order to further validate the idea that the communication style differences that we are finding link directly to the salience of an identity, it may be worthwhile to test our methodology on more data from different sources. This could include other written communication generated in real-world situations outside of

online forums. By testing this approach on data wherein the participants are communicating with a different purpose (i.e., not a forum dedicated to the group identity), this may allow us to more concretely conclude whether the uniformities in behaviour that we are observing are directly related to the salience of an identity, or whether they are instead indicative only of the purpose of the communication.

This limitation pertains directly to what we can actually conclude through observing the behaviour of individuals instead of through collecting self-reported data. Through self-reported data, we can assess exactly what an individual thinks (or claims to think) about a given situation, group, or person. Conversely, through using behavioural evidence alone, we can only infer the role of the psychological constructs at a given point in time. Empirically then, we can see that individuals in groups with similar labels communicate in similar linguistic styles. An alternative to our prototype-based explanation would perhaps suggest that these uniformities in behaviour can be understood simply as norms. Turner defines social norms as 'social uniformities among the members of a social group that arise more or less directly from their social interaction and relationships' (1991, p.2). Here, we have specifically observed social uniformities in behaviour, and thus the empirical evidence demonstrated in this thesis appears to show synthesis with Turner's idea of a norm. However, in the second part of his definition, Turner states that these uniformities arise directly from social interactions. Conversely, in this research we have demonstrated that uniformities in behaviour appear even when individual members may not have had direct interactions with each other (e.g., in the across platform analysis and in the experimental manipulations). It therefore appears that perhaps Turner's definition of a norm does not wholly fit the data presented in this thesis. Whilst norms appear to be the correct terminology when applied to local

group processes (e.g., Postmes et al., 2000), it seems that something more central to the identity is the explanatory factor needed within this research. Therefore, we have chosen to interpret our findings in light of prototypes and prototypicality.

Turner (1991) also states that ingroup socially normative responses are a 'function of a 'us x situation' interaction' (p. 162). In this research, we have changed the features of the 'situation' aspect of this equation by using data from different platforms and also from experimental conditions. Through changing the situation and yet keeping the 'us' (social identity label) consistent, we still note uniformities in group behaviour. Thus again, this points to the idea that Turner considers norms to be more localised than the uniformities that we demonstrate in this research. Once again however, and as noted previously, in order to further validate the idea that it is the salience of an identity that is impacting communication, it would be beneficial to change more aspects of the 'situation' in order to fully appreciate the limitations of this approach. I noted earlier that changing the purpose of the communication may be one such aspect that could be manipulated to assess whether this leads to a change in communication style, however other factors that we could also change include the salience of an outgroup, the format of the communication (spoken or written) and the explicit audience of the communication. Whilst in our survey methodology in Studies 1e and 1f the audience of communication changed from other ingroup members to researchers, in future we may benefit from making this more explicit to our participants by having them communicate directly to outgroup members. In turn, this would enable us to comprehend when individuals communicate in prototypical ways, and when they may communicate more strategically (Reicher et al., 1998).

Additionally, in this thesis whilst we have suggested that it is the cognitive notion of a prototype which guides behaviour, we have not attempted to directly assess what individuals believe their identity prototype to be. Instead, in Study 3c, we chose to use the content of their communication to indirectly assess aspects of the group prototype, such as emotions and values, and observed how these content-based indicators related to the prototypical linguistic style behaviour of each group. One could therefore argue that this content-based analysis has demonstrated that the values and emotions that each group has (the content of their prototypes, Hogg & Reid, 2006), relates strongly to how they behave. Whilst this analysis did not involve directly asking individuals to report their values or the traits of their identities, this more subtle approach enables us to transcend the demand characteristics of such methodologies. Through this synthesis of linguistic content and linguistic style therefore, we suggest a clear relationship between how a group behaves and what the group values.

Once again, however, this opens up the criticism of whether we can rely on the groups' actual behaviour as a good indicator of the values that the group holds. I would argue that this behavioural approach allows us to gain a more objective measure of what the group actually values, as compared to what they may say they value if asked. Here, we circumvent issues of social desirability in favour of seeing how the group actually behaves in practice. Again then, we return to the argument that behaviour allows us to understand how identities are constructed, enacted and mobilised directly *in situ*. One limitation though, would be whether emotions and values alone are 'good enough' indicators of what prototypes actually are. In future, it may be worthwhile to assess other group-defining attributes which may impact the way in which groups communicate.

Through using behavioural measures instead of self-report measures, this also opens up a debate about whether explicit awareness of the group prototype is necessary to conclude that we are observing group prototypical behaviour. That is, are individuals consciously aware of their prototype and relying on this idea to inform their behaviour or are the regularities in behaviour that we have observed more implicit than this? Perhaps the question of importance here is whether explicit awareness of a prototype is necessary to its influence on behaviour. Due to the lack of theorising around how prototypes come to impact behaviour, this question seems unanswered. However, as many have demonstrated that awareness is not necessary for salience to impact behaviour (e.g., Coppin et al., 2016), I would thus contend that explicit awareness is perhaps therefore not a prerequisite for prototypical behaviour. Nonetheless, now we have developed a behavioural measure of prototypicality, perhaps this is a question or avenue of research which can now be explored in more detail.

Further, another key limitation pertains to the difficulty in understanding the variation of each groups' behaviour in Chapter 4. Thus, in the research outlined in Chapter 4, we analysed each subreddits' data as one group identity. However, it is entirely possible that there are subgroups within these subreddits which we fail to assess through grouping all the behaviour together. Therefore, one future aim of this work may be to develop a way to understand the level of homogeneity in behaviour within each subreddit. In turn, this would enable us to understand whether certain subreddits are actually made up of a variety of different distinctive subgroups, or instead whether their behaviour is more homogenous across the entire subreddit. This approach of modelling heterogeneity and homogeneity within subreddits may also show synthesis with the cultural work of Gelfand et al. (2011). In their influential

paper, Gelfand and colleagues demonstrate the distinction between tight and loose cultures and how these different types of culture impact future behaviours such as responses to the Covid19 pandemic (Gelfand et al., 2021). In a similar way, it may be possible to model how the homogeneity of a subreddit (or other online forum) can be used to predict or influence social dynamics within the forum such as information dissemination.

In addition to these limitations, the work outlined in this thesis adopts a broader understanding of group prototypes than has previously been suggested. The expression of identity in online forums links to both the perception that one has of the identity, as well as the purpose of the communication. In this discussion, I have argued that the purpose that one has for assigning a category label to oneself may be just as important, if not more important to understanding a particular social prototype, as opposed to only referring to the prototype as a more abstract set of traits. However, it must be acknowledged that others may view this conflation between purpose and abstract traits as a limitation.

In addition to this, throughout this thesis I have tended to avoid discussing exactly what is being communicated through linguistic style. At the beginning of the literature review, I outlined that scholars from a variety of different domains (e.g., sociolinguistics, psycholinguistics, computer science) have suggested that style may be linked to personality (Boyd et al., 2021), demographics (Löckenhoff et al., 2008), psychopathologies (Junghaenel et al., 2008) and so on. However, what precisely is being communicated via linguistic style remains relatively elusive. It is possible, for example, that social identity impacts communication in that individuals seek to consciously display aspects of their identity that they see as integral to their social group, for example intelligence or passion. This would be in line with the more

strategic aspects of the SIDE model (Klein et al., 2007). Conversely, as has been argued by some personality scholars (Boyd et al. 2020; Tausczik & Pennebaker, 2010), linguistic style, and moreover word frequencies, provide a lens through which to understand attentional habits. For example, people who use articles and prepositions more frequently in their texts tend to be more concerned about formal concepts and their inter-relations (Pennebaker et al., 2014) and those who hold higher social statuses tend to use more other-oriented words such as 'you' and 'we', in order to demonstrate their focus on the external social environment (Kacewicz et al., 2014). In this argument, style is a relatively non-conscious behaviour that provides a window to an individual's psychological reality and to the aspects of the situation to which they pay attention.

To be clear then, and in line with the previous argument about the explicit awareness needed to define prototypical behaviour, it is not directly evident whether style is consciously constructed to portray an identity, or whether it is the accidental by-product of the individual's motivations and attention. However, I would also suggest that is perhaps naïve to assume a dichotomy here; behaviour can be impacted both consciously and non-consciously when a social identity is made salient. Whilst it serves as an interesting idea to understand the level of control that individuals have over their stylistic behaviour, this aim is out of the scope of this thesis. Regardless of whether prototypically stylistic behaviour is conscious or non-conscious, the evidence in this thesis still points to the finding that individuals change their linguistic style to match that of the salient social identity when a particular social identity is salient.

Finally, one of the broader methodological constraints of this research pertains to using black box algorithms. As noted in the discussion of Chapter 2, whilst black

box algorithms can enable us to improve our predictive classification abilities, this comes at a cost to explainability. Thus, due to the non-linear relationships between our linguistic indicators and the class (social identity) that we are trying to predict, it is challenging to understand precisely how each indicator contributes to the classification result. In other words, we cannot understand why one particular post is classified as 'entrepreneur' where another is classified as 'libertarian'. Instead, all we can access is the importance of each feature in determining that classification. Moreover, we cannot understand how the classifier calculates a specific 'score' (between 0 and 1) to illustrate the certainty that the classifier has that a post has originated in one of two forums. In this way, it makes it challenging to draw any robust conclusions regarding what makes a particular post seem libertarian, and what makes another post seem entrepreneurial.

Due to the black box nature of this research therefore, perhaps it would be beneficial to complement the current body of work with self-reported research. For example, in Chapter 3 whilst we used response generation to assess whether individuals who behaved more prototypically were also more influential, a more transparent analysis may involve asking individuals to judge the prototypicality of different posts in order to see whether human-judgments align with the 'judgments' of the machine learning algorithm. In doing so, it may in fact also be possible to update the parameters of the algorithm to align with the human judgments more precisely. This of course relies on there being consensus in the judgments of the human coders. As noted in the literature review, whilst human judgments are apt at deciphering a prototypical leader from a non-prototypical leader when given carefully constructed vignettes (e.g., Giessner et al., 2009), it is not entirely clear whether this group-level consensus will translate when assessing messier real-world behaviours.

Nonetheless, this analysis would definitely allow for more robust conclusions regarding the use of the algorithm for measuring the behavioural prototypicality of online forum posts.

5.5 Concluding Remarks

In sum, this thesis has demonstrated a new behavioural perspective through which we can study social identities in online environments. We have shown empirical evidence which suggests that individuals behave prototypically when communicating in online intragroup forums. We have also argued for the distinction between locally-derived prototypicality, and more context-independent measures of prototypicality. Moreover, we have also demonstrated the value in using naturally occurring real-world data in order to extend and test sociopsychological principles. However, we note that behavioural approaches come at the detriment of understanding precisely how different psychological constructs interact to create behaviour. Nonetheless, we argue that by studying social identities in real world contexts, we can gain greater understanding as to how individuals actively express their identities and ascribe meaning to their category-labels in real time. Thus, we prioritise the notion that identities do not exist in a social vacuum, instead they are mobilised in order to achieve specific goals which should not be considered independently to the identity itself.

References

Abrams, D. (1992). Processes of social identification. In G. Breakwell (Ed.), *Social psychology of identity and the self-concept* (pp. 57-99). Academic Press.

Abrams, D. (1994). Social self-regulation. *Personality and Social Psychology Bulletin, 20*(5), 473-483. https://doi.org/10.1177/0146167294205004

Abrams, D., Wetherell, M., Cochrane, S., Hogg, M. A., & Turner, J. C. (1990). Knowing what to think by knowing who you are: Self-categorization and the nature of norm formation, conformity and group polarization. *British Journal of Social Psychology, 29*(2), 97-119. https://doi.org/10.1111/j.2044-8309.1990.tb00892.x

Adra, A., Harb, C., Li, M., & Baumert, A. (2020). Predicting collective action tendencies among Filipina domestic workers in Lebanon: Integrating the Social Identity Model of Collective Action and the role of fear. *Group Processes & Intergroup Relations, 23*(7), 967–978. https://doi.org/10.1177/1368430219885180

Ahmed, R., & Bates, B. R. (2016). To accommodate, or not to accommodate: Exploring patient satisfaction with doctors' accommodative behavior during the clinical encounter. *Journal of Communication in Healthcare, 9*(1), 22-32. https://doi.org/10.1080/17538068.2015.1126936

Aldridge, J., & Décary-Hétu, D. (2016). Hidden wholesale: The drug diffusing

  capacity of online drug cryptomarkets. *International Journal of Drug Policy,*

  35, 7-15. https://doi.org/10.1016/j.drugpo.2016.04.020

Austin, A., & Goodman, R. (2017). The impact of social connectedness and

  internalized transphobic stigma on self-esteem among transgender and

  gender non-conforming adults. *Journal of homosexuality, 64*(6), 825-841.

  https://doi.org/10.1080/00918369.2016.1236587

Baddeley, J. L., & Singer, J. A. (2008). Telling losses: Personality correlates and

  functions of bereavement narratives. *Journal of Research in Personality,*

  42(2), 421-438. https://doi.org/10.1016/j.jrp.2007.07.006

Bagozzi, R. P., & Dholakia, U. M. (2002). Intentional social action in virtual

  communities. *Journal of Interactive Marketing, 16*(2), 2-21.

  https://doi.org/10.1002/dir.10006

Bargh, J. A., Lombardi, W. J., & Higgins, E. T. (1988). Automaticity of chronically

  accessible constructs in person× situation effects on person perception: It's

  just a matter of time. *Journal of Personality and Social Psychology, 55*(4),

  599-625. https://doi.org/10.1037/0022-3514.55.4.599

Barratt, M. J., & Aldridge, J. (2016). Everything you always wanted to know about

  drug cryptomarkets* (*but were afraid to ask*). International Journal of Drug*

  Policy, 35, 1-6. https://doi.org/10.1016/j.drugpo.2016.07.005

Barreto, M., Ellemers, N., & Banal, S. (2006). Working under cover: Performance-related self-confidence among members of contextually devalued groups who try to pass. *European Journal of Social Psychology, 36*(3), 337-352. https://doi.org/10.1002/ejsp.314

Barreto, M., Spears, R., Ellemers, N., & Shahinper, K. (2003). Who wants to know? The effect of audience on identity expression among minority group members*. British Journal of Social Psychology, 42*(2), 299-318.https://doi.org/10.1348/014466603322127265

Barreto, N. B., & Hogg, M. A. (2017). Evaluation of and support for group prototypical leaders: A meta-analysis of twenty years of empirical research. *Social Influence, 12*(1), 41-55.

Bartel, C. A., & Wiesenfeld, B. M. (2013). The social negotiation of group prototype ambiguity in dynamic organizational contexts. *Academy of Management Review, 38*(4), 503-524. https://doi.org/10.5465/amr.2011.0300

Baumeister, R. F. (1986). *Identity: Cultural change and the struggle for self*. Oxford University Press.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?.

Perspectives on psychological science, 2(4), 396-403.

https://doi.org/10.1111/j.1745-6916.2007.00051.x

Bettencourt, B. A., & Hume, D. (1999). The cognitive contents of social-group

identity: Values, emotions, and relationships. *European Journal of Social*

Psychology, 29(1), 113–121. https://doi.org/10.1002/(SICI)1099-

0992(199902)29:1<113::AID-EJSP911>3.0.CO;2-G

Billig, M. (1996). Arguing and thinking: A rhetorical approach to social psychology.

Cambridge University Press.

Bissinger, B. (2015, July). Caitlyn Jenner: The Full Story. *Vanity Fair*.

https://www.vanityfair.com/hollywood/2015/06/caitlyn-jenner-bruce-cover-

annie-leibovitz

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET for Windows:*

Software for social network analysis. Analytic Technologies

Boyd, R. L. (2017). Psychological Text Analysis in the Digital Humanities. In S. Hai-

Jew (Ed*.), Data Analytics in Digital Humanities* (pp. 161-189). Springer

International Publishing. https://doi.org/10.1007/978-3-319-54499-1_7

Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: a new

approach to personality in a digital world. *Current Opinion in Behavioral*

Sciences, 18, 63-68. https://doi.org/10.1016/j.cobeha.2017.07.017

Boyd, R. L., & Schwartz, H. A. (2021). Natural Language Analysis and the

  Psychology of Verbal Behavior: The Past, Present, and Future States of the

  Field. *Journal of Language and Social Psychology, 40*(1), 21-41.

  https://doi.org/10.1177/0261927X20967028

Boyd, R. L., Pasca, P., & Lanning, K. (2020). The personality panorama:

  Conceptualizing personality through big behavioural data. *European Journal*

  of Personality, 34(5), 599-612. https://doi.org/10.1002/per.2254

Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., &

  Mihalcea, R. (2015). Values in Words: Using Language to Evaluate and

  Understand Personal Values. In D. Quercia & B. Hogan (Eds.), *Proceedings*

  of the Ninth AAAI International Conference on Web and Social Media, (pp.

  31–40). The AAAI Press.

Branscombe, N. R., & Wann, D. L. (1994). Collective self-esteem consequences of

  outgroup derogation when a valued social identity is on trial. *European*

  Journal of Social Psychology, 24(6), 641-657.

  https://doi.org/10.1002/ejsp.2420240603

Bravo-Marquez, F., Mendoza, M., & Poblete, B. (2014). Meta-level sentiment models

  for big social data analysis. *Knowledge-based systems, 69*, 86-99.

  https://doi.org/10.1016/j.knosys.2014.05.016

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin, 17*(5), 475-482. https://doi.org/10.1177/0146167291175001

Brown, R. (2000), Social identity theory: past achievements, current problems and future challenges. *European Journal of Social Psychology, 30*(1), 745-778. https://doi.org/10.1002/1099-0992(200011/12)30:6<745::AID-EJSP24>3.0.CO;2-O

Brown, R., & Williams, J. (1984). Group Identification: The Same Thing to All People? *Human Relations, 37*(7), 547–564. https://doi.org/10.1177/001872678403700704

Brown, T. E., Davidsson, P., & Wiklund, J. (2001). An operationalization of Stevenson's conceptualization of entrepreneurship as opportunity-based firm behavior. *Strategic Management Journal, 22*(10), 953-968. https://doi.org/10.1002/smj.190

Bry, L. J., Mustanski, B., Garofalo, R., & Burns, M. N. (2018). Resilience to discrimination and rejection among young sexual minority males and transgender females: A qualitative study on coping with minority stress. *Journal of Homosexuality, 65*(11), 1435-1456. https://doi.org/10.1080/00918369.2017.1375367

Burns, K. (2019, December 27). The internet made trans people more visible. It also left them more vulnerable. *Vox.com.* https://www.vox.com/identities/2019/12/27/21028342/trans-visibility-backlash-internet-2010

Caporael, L. R., Lukaszewski, M. P., & Culbertson, G. H. (1983). Secondary baby talk: Judgments by institutionalized elderly and their caregivers. *Journal of Personality and Social Psychology, 44*(4), 746-754. https://doi.org/10.1037/0022-3514.44.4.746

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. P. (2010). Measuring user influence in Twitter: The million follower fallacy. In Fourth International AAAI *Conference on Weblogs and Social Media. AAAI*. https://www.aaai.org/ ocs/index.php/ICWSM/ICWSM10/paper/ view/1538/1826

Chen, V. H. H., & Wu, Y. (2015). Group identification as a mediator of the effect of players' anonymity on cheating in online games. *Behaviour & Information Technology, 34*(7), 658-667. https://doi.org/10.1080/0144929X.2013.843721

Chung, C. K., & Pennebaker, J. W. (2019). Textual analysis. In H. Blanton, J. M. LaCroix, & G. D. Webster (Eds.), *Frontiers of social psychology. Measurement in social psychology* (pp.153-173). Routledge/Taylor & Francis Group. https://doi.org/10.4324/9780429452925-7

Chung, C., & Pennebaker, J. (2007). The Psychological Functions of Function

Words. In K. Fiedler (Ed.), *Frontiers of social psychology. Social

communication* (p. 343-359). Psychology Press.

Coleman, N. V., & Williams, P. (2013). Feeling like my self: Emotion profiles and

social identity. *Journal of Consumer Research, 40*(2), 203-222.

https://doi.org/10.1086/669483

Coppin, G., Pool, E., Delplanque, S., Oud, B., Margot, C., Sander, D., & Van Bavel,

J. J. (2016). Swiss identity smells like chocolate: Social identity shapes

olfactory judgments. *Scientific Reports*, 6(1), 1-10.

https://doi.org/10.1038/srep34979

Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. In C. Zhang &

Y.Q. Ma (Eds.), *Ensemble machine learning* (pp. 157-175). Springer.

http://dx.doi.org/10.1007/978-1-4419-9326-7_5

Danbold, F., & Huo, Y. J. (2015). No Longer "All-American"? Whites' Defensive

Reactions to Their Numerical Decline. *Social Psychological and Personality

Science, 6*(2), 210–218. https://doi.org/10.1177/1948550614546355

Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C.

(2013). No Country for Old Members: User Lifecycle and Linguistic Change

in Online Communities. In D. Schwabe (Ed.), *WWW '13: Proceedings of the

22nd international conference on World Wide Web* (pp. 307-318).

Association for Computing Machinery.
https://doi.org/10.1145/2488388.2488416

De Cremer, D., van Dijke, M., & Mayer, D. M. (2010). Cooperating When "You" and
"I" Are Treated Fairly: The Moderating Role of Leader Prototypicality. *Journal
of Applied Psychology, 95*(6), 1121-1133. https://doi.org/10.1037/a0020419

Deaux, K. (1991). Social identities: Thoughts on structure and change. In R. C.
Curtis (Ed.), *The relational self: Theoretical convergences in psychoanalysis
and social psychology* (pp. 77-93). The Guilford Press.

Deaux, K. (1993). Reconstructing social identity. *Personality and Social Psychology
Bulletin, 19*(1), 4–12. https://doi.org/10.1177/0146167293191001

Deaux, K. (2000). Models, meanings and motivations. In D. Capozza & R. Brown
(Eds.), *Social identity processes: Trends in theory and research* (p. 1–14).
Sage Publications Ltd. https://doi.org/10.4135/9781446218617.n1

Deaux, K., & Major, B. (1987). Putting gender into context: An interactive model of
gender-related behavior. *Psychological Review, 94*(3), 369–389.
https://doi.org/10.1037/0033-295X.94.3.369

Deaux, K., Reid, A., Mizrahi, K., & Cotting, D. (1999). Connecting the person to the
social: The functions of social identification. In T. R. Tyler, R. M. Kramer, &

O. P. John (Eds.), *Applied social research. The psychology of the social self* (p. 91–113). Lawrence Erlbaum Associates Publishers.

Deaux, K., Reid, A., Mizrahi, K., & Ethier, K. A. (1995). Parameters of social identity. *Journal of personality and social psychology, 68*(2), 280-291. https://doi.org/10.1037/0022-3514.68.2.280

Devos, T., & Banaji, M. R. (2005). American= white?. *Journal of personality and social psychology, 88*(3), 447-466. https://doi.org/10.1037/0022-3514.88.3.447

Doise, W., Sinclair, A., & Bourhis, R. Y. (1976). Evaluation of Accent Convergence and Divergence in Cooperative and Competitive Intergroup Situations*. British Journal of Social and Clinical Psychology, 15*(3), 247-252. https://doi.org/10.1111/j.2044-8260.1976.tb00031.x

Doliński, D. (2018). Is psychology still a science of behaviour?. *Social Psychological Bulletin, 13*(2), 1-14. https://doi.org/10.5964/spb.v13i2.25025

Doosje, B., Ellemers, N., & Spears, R. (1995). Perceived intragroup variability as a function of group status and identification. *Journal of experimental social psychology, 31*(5), 410-436. https://doi.org/10.1006/jesp.1995.1018

Doosje, B., Haslam, S. A., Spears, R., Oakes, P. J., & Koomen, W. (1998). The effect of comparative context on central tendency and variability judgements

and the evaluation of group characteristics. European *Journal of Social Psychology, 28*(2), 173-184. https://doi.org/10.1002/(SICI)1099-0992(199803/04)28:2<173::AID-EJSP862>3.0.CO;2-Z

Dovidio, J. F., Gaertner, S. L., Hodson, G., Houlette, M. A., & Johnson, K. M. (2005). Social Inclusion and Exclusion: Recategorization and the Perception of Intergroup Boundaries. In D. Abrams, M. A. Hogg, & J. M. Marques (Eds.), *The social psychology of inclusion and exclusion* (p. 245–264). Psychology Press.

Duckitt, J. H. (1992). *The social psychology of prejudice*. Praeger Publishers/Greenwood Publishing Group

Durrheim, K., Quayle, M., & Dixon, J. (2016). The Struggle for the Nature of "Prejudice": "Prejudice" Expression as Identity Performance. *Political Psychology, 37*(1), 17-35. https://doi.org/10.1111/pops.12310

Eisenberger, R., Karagonlar, G., Stinglhamber, F., Neves, P., Becker, T. E., Gonzalez-Morales, M. G., & Steiger-Mueller, M. (2010). Leader-Member Exchange and Affective Organizational Commitment: The Contribution of Supervisor's Organizational Embodiment. *Journal of Applied Psychology, 95*(6), 1085-1103. https://doi.org/10.1037/a0020858

Ellemers, N., & Jetten, J. (2013). The many ways to be marginal in a group. *Personality and Social Psychology Review, 17*(1), 3-21. https://doi.org/10.1177/1088868312453086

Ellemers, N., Spears, R., & Doosje, B. (1997). Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of Personality and Social Psychology, 72*(3), 617–626. https://doi.org/10.1037/0022-3514.72.3.617

Emmons, L. (2019, December 18). The 2010s were the decade of trans. *US Spectator.* https://spectator.us/topic/2010s-decade-trans/

Eronen, M. I., & Bringmann, L. F. (2021). The Theory Crisis in Psychology: How to Move Forward. *Perspectives on Psychological Science.* https://doi.org/10.1177/1745691620970586

Ethier, K., & Deaux, K. (1990). Hispanics in ivy: Assessing identity and perceived threat. *Sex roles, 22*(7-8), 427-440. https://doi.org/10.1007/BF00288162

Fast, L. A., & Funder, D. C. (2008). Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology, 94*(2), 334–346. https://doi.org/10.1037/0022-3514.94.2.334

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2), 175-191. https://doi.org/10.3758/BF03193146

Fielding, K. S., & Hogg, M. A. (1997). Social identity, self-categorization, and leadership: A field study of small interactive groups. *Group Dynamics, 1(*1), 39-51. https://doi.org/10.1037/1089-2699.1.1.39

Firth, J.R. (1957), Applications of General Linguistics. *Transactions of the Philological Society, 56*, 1-14. https://doi.org/10.1111/j.1467-968X.1957.tb00568.x

Frijda, N. H. (1986). *The emotions.* Cambridge University Press.

Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology, 57*(2), 212–228. https://doi.org/10.1037/0022-3514.57.2.212

Gallois, C., Gasiorek, J., Giles, H., & Soliz, J. (2016). Communication Accommodation Theory: Integrations and New Framework Developments. In H. Giles (Ed.), *Communication Accommodation Theory: Negotiating Personal Relationships and Social Identities across Contexts* (pp. 192-210). Cambridge University Press. doi:10.1017/CBO9781316226537.010

Gallois, C., Ogay, T., & Giles, H. (2005). Communication Accommodation Theory: A Look Back and a Look Ahead. In W.B. Gudykunst (Ed.), *Theorizing About Communication and Culture* (pp. 121-148). Sage

Ganegedara, T. (2018). *Natural Language Processing with TensorFlow*. Packt Publishing

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Ferrer, M., Fischlmayr, I.R., …Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *science, 332*(6033), 1100-1104. https://doi.org/10.1126/science.1197754

Gelfand, M. J., Jackson, J. C., Pan, X., Nau, D., Pieper, D., Denison, E., Dagher, M., van Lange, P., Chiu, C. Y., & Wang, M. (2021). The relationship between cultural tightness-looseness and COVID-19 cases and deaths: a global analysis. *The Lancet. Planetary health, 5*(3), e135–e144. https://doi.org/10.1016/S2542-5196(20)30301-6

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning, 63*(1), 3-42. https://doi.org/10.1007/s10994-006-6226-1

Gibson, R. (2002). *Intercultural Business Communication: An Introduction to the Theory and Practice of Intercultural Business Communication for Teachers, Language Trainers, and Business People*. Oxford University Press.

Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age.* Stanford university press.

Giessner, S. R., & van Knippenberg, D. (2008). "License to Fail": Goal definition, leader group prototypicality, and perceptions of leadership effectiveness after leader failure. *Organizational Behavior and Human Decision Processes, 105*(1), 14-35. https://doi.org/10.1016/j.obhdp.2007.04.002

Giessner, S. R., van Knippenberg, D., & Sleebos, E. (2009). License to fail? How leader group prototypicality moderates the effects of leader performance on perceptions of leadership effectiveness*. Leadership Quarterly, 20*(3), 434-451. https://doi.org/10.1016/j.leaqua.2009.03.012

Gilbert, G. M. (1951). Stereotype persistance and change among college students. *Journal of Abnormal and Social Psychology, 46*(2), 245-254. https://doi.org/10.1037/h0053696

Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of Accommodation: Developments in Applied*

Sociolinguistics (Studies in Emotion and Social Interaction) (pp. 1-68). Cambridge University Press. doi:10.1017/CBO9780511663673.001

Goksuluk, D., Korkmaz, S., Zararsiz, G., & Karaagaoglu, A. E. (2016). easyROC: an interactive web-tool for ROC curve analysis using R language environment. *The R Journal, 8*(2), 213-230.

Goldman, L., & Hogg, M. A. (2016). Going to extremes for one's group: the role of prototypicality and group acceptance. *Journal of Applied Social Psychology, 46(*9), 544-553. https://doi.org/10.1111/jasp.12382

Gonzales, A. L., Hancock, J. T., & Pennebaker, J. W. (2010). Language Style Matching as a Predictor of Social Dynamics in Small Groups. *Communication Research, 37*(1), 3-19. https://doi.org/10.1177/0093650209351468

Google Developers (n.d.). *"Embeddings can produce remarkable analogies."[*Image]. Google Developers https://developers.google.com/machine-learning/crash-course/embeddings/translating-to-a-lower-dimensional-space

Graen, G. B., & Uhl-Bien, M. (1995). Relationshipbased approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective. *The Leadership Quarterly, 6*(2), 219–247. https://doi.org/10.1016/1048- 9843(95)90036-5

Green, K. M. (2016). Troubling the Waters. In E. P. Johnson (Ed.), *No Tea, No Shade* (pp. 65-82). Duke University Press. https://doi.org/10.1215/9780822373711-004

Greenberg, J. (1983). Self-image versus impression management in adherence to distributive justice standards: The influence of self-awareness and self-consciousness. *Journal of Personality and Social Psychology, 44*, 5-19. https://doi.org/10.1037/0022-3514.44.1.5

Grieve, J., Nini, A., & Guo, D. (2017). Analyzing lexical emergence in Modern American English online. *English Language and Linguistics, 21*(1), 99-127. https://doi.org/10.1017/S1360674316000113

Gross, K., & D'Ambrosio, L. (2004). Framing emotional response. *Political Psychology, 25*(1), 1–29. https://doi.org/10.1111/j.1467-9221.2004.00354.x

Gurin, P., & Markus, H. (1988). Group identity: The psychological mechanisms of durable salience. *Revue Internationale de Psychologie Sociale, 1*(2), 257-274

Gómez, Á., Jetten, J., & Swann, W. B. (2014). The more prototypical the better? The allure of being seen as one sees oneself. *Group Processes & Intergroup Relations, 17*(2), 221-239. https://doi.org/10.1177/1368430213510570

Halevy, N., Berson, Y., & Galinsky, A. D. (2011). The mainstream is not electable: When vision triumphs over representativeness in leader emergence and effectiveness. *Personality and Social Psychology Bulletin, 37*(7), 893-904. https://doi.org/10.1177/0146167211402836

Hallier-Willi, C., Nguyen, B., Melewar, T. C., & Dennis, C. (2014). Corporate impression formation in online communities: a qualitative study. *Qualitative Market Research: An International Journal, 17*(4), 410-440. http://dx.doi.org/10.1108/QMR-07-2013-0049

Hancock, A. M. (2004). T*he politics of disgust: The public identity of the welfare queen*. NYU Press.

Haslam, S. A. (2004). *Social identity in organizations: The social identity approach.* Sage.

Haslam, S. A., & Turner, J. C. (1992). Context-dependent variation in social stereotyping 2: The relationship between frame of reference, self-categorization and accentuation. *European Journal of Social Psychology, 22*(3), 251-277. https://doi.org/10.1002/ejsp.2420220305

Haslam, S. A., Oakes, P. J., & Turner, J. C. (1996a). Social identity, self-categorization, and the perceived homogeneity of ingroups and outgroups: The interaction between social motivation and cognition. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition. Handbook of*

motivation and cognition (Volume 3). The interpersonal context (pp. 182-222). The Guilford Press.

Haslam, S. A., Oakes, P. J., McGarty, C., Turner, J. C., & Onorato, R. S. (1995). Contextual changes in the prototypicality of extreme and moderate outgroup members. *European Journal of Social Psychology, 25*(5), 509-530. https://doi.org/10.1002/ejsp.2420250504

Haslam, S. A., Oakes, P. J., McGarty, C., Turner, J. C., Reynolds, K. J., & Eggins, R. A. (1996b). Stereotyping and social influence: The mediation of stereotype applicability and sharedness by the views of in-group and out-group members. *British Journal of Social Psychology, 35*(3), 369-397. https://doi.org/10.1111/j.2044-8309.1996.tb01103.x

Haslam, S. A., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social identity salience and the emergence of stereotype consensus. *Personality and Social Psychology Bulletin, 25*(7), 809-818. https://doi.org/10.1177/0146167299025007004

Haslam, S. A., Turner, J. C., Oakes, P. J., McGarty, C., & Hayes, B. K. (1992). Context-dependent variation in social stereotyping 1: The effects of intergroup relations as mediated by social change and frame of reference. *European Journal of Social Psychology, 22*(1), 3-20. https://doi.org/10.1002/ejsp.2420220104

Haslam, S. A., Turner, J. C., Oakes, P. J., McGarty, C., & Reynolds, K. J. (1997).

    The Group as a Basis for Emergent Stereotype Consensus. *European*

    Review of Social Psychology, 8(1), 203-239.

    https://doi.org/10.1080/14792779643000128

Hayward, P., & Bright, J. A. (1997). Stigma and mental illness: A review and critique.

    Journal of Mental Health, 6(4), 345-354.

    https://doi.org/10.1080/09638239718671

Hirst, G., van Dick, R., & van Knippenberg, D. (2009). A social identity perspective

    on leadership and employee creativity. *Journal of Organizational Behavior:*

    The International Journal of Industrial, Occupational and Organizational

    Psychology and Behavior, 30(7), 963-982. https://doi.org/10.1002/job.600

Hogg, M. A. (2001). A social identity theory of leadership. *Personality and Social*

    *Psychology Review, 5(*3), 184-200.

    https://doi.org/10.1207/S15327957PSPR0503_1

Hogg, M. A. (2006). Social Identity Theory. In P. J. Burke (Ed.), *Contemporary social*

    psychological theories (pp. 111-136). Stanford University Press.

Hogg, M. A. (2007). Uncertainty-Identity Theory. *Advances in Experimental Social*

    Psychology, 39, 69-126. https://doi.org/10.1016/S0065-2601(06)39002-8

Hogg, M. A., & Abrams, D. (1993). Towards a single-process uncertainty-reduction model of social motivation in groups. In M. A. Hogg & D. Abrams (Eds.), *Group motivation: Social psychological perspectives* (pp. 173-190). Harvester Wheatsheaf.

Hogg, M. A., & Hains, S. C. (1996). Intergroup Relations and Group Solidarity: Effects of Group Identification and Social Beliefs on Depersonalized Attraction. *Journal of Personality and Social Psychology, 70*(2), 295-309. https://doi.org/10.1037/0022-3514.70.2.295

Hogg, M. A., & Hardie, E. A. (1991). Social Attraction, Personal Attraction, and Self-Categorization-, A Field Study. *Personality and Social Psychology Bulletin, 17*(2), 175–180. https://doi.org/10.1177/014616729101700209

Hogg, M. A., & Knippenberg, D. van. (2003). Social Identity and Leadership Processes in Groups. *Advances in Experimental Social Psychology, 35*, 1-52. https://doi.org/10.1016/S0065-2601(03)01001-3

Hogg, M. A., & Reid, S. A. (2006). Social identity, self-categorization, and the communication of group norms. *Communication Theory, 16*(1), 7-30. https://doi.org/10.1111/j.1468-2885.2006.00003.x

Hogg, M. A., & Smith, J. R. (2007). Attitudes in social context: A social identity perspective. *European Review of Social Psychology, 18*(1), 89-131. https://doi.org/10.1080/10463280701592070

Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology, 26*(4), 325-340. https://doi.org/10.1111/j.2044-8309.1987.tb00795.x

Hogg, M. A., Giles, H. (2012). Norm talk and identity in intergroup communication. In H. Giles. (Ed.), *The handbook of intergroup communication* (pp. 373-388). Routledge

Hogg, M. A., Rast, D. E., & van Knippenberg, D. (2012). The social identity theory of leadership: Theoretical origins research findings and conceptual developments. *European Review of Social Psychology, 23*(1), 258-304. https://doi.org/10.1080/10463283.2012.741134

Hogg, M. A., Turner, J. C., & Davidson, B. (1990). Polarized Norms and Social Frames of Reference: A Test of the Self-Categorization Theory of Group Polarization. *Basic and Applied Social Psychology, 11*(1), 77-100. https://doi.org/10.1207/s15324834basp1101_6

Hogg, M., & Terry, D. (2000). Social Identity and Self-Categorization Processes in Organizational Contexts. *The Academy of Management Review, 25*(1), 121-140. https://doi.org/10.5465/amr.2000.2791606

Hollander, E. P. (1980). Leadership and social exchange processes. In K.J. Gergen, M.S. Greenberg & R.H. Willis (Eds.), *Social exchange* (pp. 103– 118). Springer. https://doi.org/10.1007/978-1- 4613-3087-5_5

Hollander, E. P. (1992). Leadership, followership, self, and others. The Leadership

    Quarterly, 3(1), 43–54. https://doi.org/10.1016/1048-9843(92)90005-Z

Hopkins, N., Regan, M., & Abell, J. (1997). On the context dependence of national

    stereotypes: Some Scottish data. British Journal of Social Psychology, 36(4),

    553-563. https://doi.org/10.1111/j.2044-8309.1997.tb01149.x

Hout, M. C., Papesh, M. H., & Goldinger, S. D. (2013). Multidimensional scaling.

    Wiley Interdisciplinary Reviews: Cognitive Science, 4(1), 93-103.

    https://doi.org/10.1002/wcs.1203

Huddy, L. (2001). From social to political identity: A critical examination of social

    identity theory. *Political psychology, 22*(1), 127-156.

    https://doi.org/10.1111/0162-895X.00230

Huddy, L. (2004). Contrasting theoretical approaches to intergroup relations. *Political*

    Psychology, 25(6), 947-967. https://doi.org/10.1111/j.1467-

    9221.2004.00404.x

Huffaker, D. (2010). Dimensions of leadership and social influence in online

    communities. *Human Communication Research, 36*(4), 593–617. https://

    doi.org/10.1111/j.1468-2958.2010.01390.x

Isenberg, D. J. (1986). Group polarization: A critical review and meta-analysis. *Journal of Personality and Social Psychology, 50*(6), 1141-1151. https://doi.org/10.1037/0022-3514.50.6.1141

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding libertarian morality: The psychological dispositions of self-identified libertarians. *PloS one, 7*(8), e42366.

Jackson, J. M., & Saltzstein, H. D. (1958). The effect of person-group relationships on conformity processes. *The Journal of Abnormal and Social Psychology, 57*(1), 17-24. https://doi.org/10.1037/h0048403

Jackson, L. A., Sullivan, L. A., Harnish, R., & Hodge, C. N. (1996). Achieving positive social identity: Social mobility, social creativity, and permeability of group boundaries. *Journal of Personality and Social Psychology, 70*(2), 241–254. https://doi.org/10.1037/0022-3514.70.2.241

Jans, L., Leach, C. W., Garcia, R. L., & Postmes, T. (2015). The development of group influence on in-group identification: A multilevel approach. *Group Processes and Intergroup Relations, 18*(2), 190-209. https://doi.org/10.1177/1368430214540757

Jans, L., Postmes, T., & Van der Zee, K. I. (2012). Sharing differences: The inductive route to social identity formation. *Journal of Experimental Social Psychology, 48*(5), 1145-1149. https://doi.org/10.1016/j.jesp.2012.04.013

Janson, A., Levy, L., Sitkin, S. B., & Lind, E. A. (2008). Fairness and other leadership heuristics: A four-nation study. *European Journal of Work and Organizational Psychology, 17*(2), 251-272. https://doi.org/10.1080/13594320701746510

Jenkins, S. T., Reysen, S., & Katzarska-Miller, I. (2012). Ingroup identification and personality. *Journal of Interpersonal Relations, Intergroup Relations and Identity, 5*(1), 9-16.

Jetten, J., & Hornsey, M. J. (2011). *Rebels in Groups*. Wiley-Blackwell.

Jetten, J., Branscombe, N. R., & Spears, R. (2002). On being peripheral: Effects of identity insecurity on personal and collective self-esteem. *European Journal of Social Psychology, 32*(1), 105-123. https://doi.org/10.1002/ejsp.64

Jetten, J., Branscombe, N. R., Schmitt, M. T., & Spears, R. (2001). Rebels with a cause: Group identification as a response to perceived discrimination from the mainstream. *Personality and Social Psychology Bulletin, 27*(9), 1204-1213. https://doi.org/10.1177/0146167201279012

Jetten, J., Branscombe, N. R., Spears, R., & McKimmie, B. M. (2003). Predicting the Paths of Peripherals: The Interaction of Identification and Future Possibilities. *Personality and Social Psychology Bulletin, 29*(1), 130-140. https://doi.org/10.1177/0146167202238378

Jetten, J., Hornsey, M. J., & Adarves-Yorno, I. (2006). When Group Members Admit to Being Conformist: The Role of Relative Intragroup Status in Conformity Self-Reports. *Personality and Social Psychology Bulletin, 32*(2), 162-173. https://doi.org/10.1177/0146167205279904

Jetten, J., Postmes, T., & McAuliffe, B. J. (2002). "We're all individuals": Group norms of individualism and collectivism, levels of identification and identity threat. *European Journal of Social Psychology, 32*(2), 189-207. https://doi.org/10.1002/ejsp.65

Jetten, J., Spears, R., & Manstead, A. S. R. (1996). Intergroup norms and intergroup discrimination: Distinctive self-categorization and social identity effects. *Journal of Personality and Social Psychology, 71*(6), 1222–1233. https://doi.org/10.1037/0022-3514.71.6.1222

Jetten, J., Spears, R., & Manstead, A. S. R. (1997). Strength of identification and intergroup differentiation: The influence of group norms. *European Journal of Social Psychology, 27*(5), 603-609. https://doi.org/10.1002/(SICI)1099-0992(199709/10)27:5<603::AID-EJSP816>3.0.CO;2-B

Johar, G. V., Moreau, P., & Schwarz, N. (2003). Gender typed advertisements and impression formation: The role of chronic and temporary accessibility. Journal of *Consumer Psychology, 13(*3), 220-229. https://doi.org/10.1207/S15327663JCP1303_04

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54.* Institute of Personality and Social Research. https://doi.org/10.1037/t07550-000

Johnson, S. L., Safadi, H., & Faraj, S. (2015). The emergence of online community leadership. *Information Systems Research, 26*(1), 165-187. https://doi.org/10.1287/isre.2014.0562

Jones, A. (2020*). Incels and the Manosphere: Tracking Men's Movements Online.* [Master's Thesis, University of Central Florida]. Showcase of Text, Archives, Research and Scholarship.

Jones, S., Cotterill, R., Dewdney, N., Muir, K., & Joinson, A. (August, 2014). Finding Zelig in text: A measure for normalising linguistic accommodation. Paper presented at *25th International Conference on Computational Linguistics, Dublin, UK* (pp. 455–465). https:// www.aclweb.org/anthology/C14-1044

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology, 87*(4), 765. https://doi.org/10.1037/0021-9010.87.4.765

Junghaenel, D. U., Smyth, J. M., & Santner, L. (2008). Linguistic dimensions of psychopathology: A quantitative analysis. *Journal of Social and Clinical Psychology, 27*(1), 36-55. https://doi.org/10.1521/jscp.2008.27.1.36

Jussim, L., Crawford, J. T., & Rubinstein, R. S. (2015). Stereotype (In)Accuracy in Perceptions of Groups and Individuals. *Current Directions in Psychological Science, 24(*6), 490–497. https://doi.org/10.1177/0963721415605257

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology, 33*(2), 125-143. https://doi.org/10.1177/0261927X13502654

Kameda, T., Ohtsubo, Y., & Takezawa, M. (1997). Centrality in Sociocognitive Networks and Social Influence: An Illustration in a Group Decision-Making Context. *Journal of Personality and Social Psychology, 73*(2), 296-309. https://doi.org/10.1037/0022-3514.73.2.296

Karlins, M., Coffman, T. L., & Walters, G. (1969). On the fading of social stereotypes: Studies in three generations of college students. *Journal of Personality and Social Psychology, 13*(1), 1-16. https://doi.org/10.1037/h0027994

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *Journal of Abnormal and Social Psychology, 28*(3), 280-290. https://doi.org/10.1037/h0074049

Katzarska-Miller, I., Barnsley, C. A., & Reysen, S. (2014). Global Citizenship Identification and Religiosity. *Archive for the Psychology of Religion, 36*(3), 344–367. https://doi.org/10.1163/15736121-12341291

Klandermans, B., & Goslinga, S. (1996). Media discourse, movement publicity, and

    the generation of collective action frames: Theoretical and empirical

    exercises in meaning construction. In D. McAdam, J. McCarthy, & M. Zald

    (Eds.), *Comparative Perspectives on Social Movements: Political*

    Opportunities, Mobilizing Structures, and Cultural Framings (pp. 312-337).

    Cambridge University Press. doi:10.1017/CBO9780511803987.016

Klein, O., Spears, R., & Reicher, S. (2007). Social identity performance: Extending

    the strategic side of SIDE. *Personality and Social Psychology Review, 11*(1),

    28-45. https://doi.org/10.1177/1088868306294588

Klemz, A. (1977). *Blindness and Partial Sight: A Guide for Social Workers and*

    Others Concerned with the Care and Rehabilitation of the Visually

    Handicapped. Woodhead-Faulkner.

Klenowski, P. M., Copes, H., & Mullins, C. W. (2011). Gender, identity, and

    accounts: How white collar offenders do gender when making sense of their

    crimes. *Justice Quarterly, 28*(1), 46-69.

    https://doi.org/10.1080/07418825.2010.482536

Knippenberg, D. Van, & Wilke, H. (1992). Prototypicality of arguments and

    conformity to ingroup norms. *European Journal of Social Psychology, 22*(2),

    141-155. https://doi.org/10.1002/ejsp.2420220204

Komisin, M., Guinn, C, Reinicke, B., Simmonds, D. (2011) *Identifying Personality Types Using Document Classification Methods*. [Master's Thesis, UNC Wilmington]. Annals of the Master of Science in Computer Science and Information Systems at UNC Wilmington. http://csbapp.uncw.edu/data/mscsis/full.aspx.

Koschate, M., Naserian, E., Dickens, L., Stuart, A., Russo, A., & Levine, M. (2021). ASIA: Automated Social Identity Assessment using linguistic style. *Behavior Research Methods, 53*(1) 1-20. https://doi.org/10.3758/s13428-020-01511-3

Koudenburg, N., Greijdanus, H., & Scheepers, D. (2019). The polarizing effects of group discussion in a negative normative context: Integrating societal-, group-, and individual-level factors. *British Journal of Social Psychology, 58*(1), 150-174. https://doi.org/10.1111/bjso.12282

Koudenburg, N., Postmes, T., & Gordijn, E. H. (2017). Beyond Content of Conversation: The Role of Conversational Form in the Emergence and Regulation of Social Structure. *Personality and Social Psychology Review, 21*(1), 50-71. https://doi.org/10.1177/1088868315626022

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. SAGE Publications, Inc. https://www.doi.org/10.4135/9781412985130

Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modelling (Vol. 26).* Springer. https://doi. org/10.1007/978-1-4614-6849-3

Larrimore, L., Jiang, L., Larrimore, J., Markowitz, D., & Gorski, S. (2011). Peer to peer lending: The relationship between language features, trustworthiness, and persuasion success. *Journal of Applied Communication Research, 39*(1), 19-37. https://doi.org/10.1080/00909882.2010.536844

Latrofa, M., Vaes, J., Cadinu, M., & Carnaghi, A. (2010). The cognitive representation of self-stereotyping. *Personality and Social Psychology Bulletin, 36(*7), 911-922. https://doi.org/10.1177/0146167210373907

Lau, J. H., & Baldwin, T. (2016). An Empirical Evaluation of doc2vec with Practical Insights into Document Embedding Generation. In P. Blunsom, K. Cho, E. Grefenstette, K. M. Hermann, L. Rimell, J. Weston & S. W. Yih (Eds.), *Proceedings of the 1st Workshop on Representation Learning for NLP* (pp. 78-86). Association for Computational Linguistics https://doi.org/10.18653/v1/W16-1609

Le Hénaff, B., Michinov, N., Le Bohec, O., & Delaval, M. (2015). Social gaming is inSIDE: Impact of anonymity and group identity on performance in a team game-based learning environment. *Computers & Education, 82*, 84-95. https://doi.org/10.1016/j.compedu.2014.11.002

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. In E.P. Xing & T. Jebara (Eds.), *Proceedings of the 31st*

International Conference on Machine Learning (pp. 1188-1196).

Proceedings of Machine Learning Research.

Lea, M., & Spears, R. (1991). Computer-mediated communication, de-individuation and group decision-making. *International Journal of Man-Machine Studies, 34*(2), 283-301. https://doi.org/10.1016/0020-7373(91)90045-9

Lea, M., Spears, R., & de Groot, D. (2001). Knowing Me, Knowing You: Anonymity Effects on Social Identity Processes within Groups. *Personality and Social Psychology Bulletin, 27*(5), 526-537. https://doi.org/10.1177/0146167201275002

Lea, M., Spears, R., & Watt, S. E. (2007). Visibility and anonymity effects on attraction and group cohesiveness*. European Journal of Social Psychology, 37*(4), 761-773. https://doi.org/10.1002/ejsp.398

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., Ouwerkerk, J. W., & Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology, 95*(1), 144-165. https://doi.org/10.1037/0022-3514.95.1.14

Lee, E.J. (2006). When and How Does Depersonalization Increase Conformity to Group Norms in Computer-Mediated Communication? *Communication Research, 33*(6), 423-447. https://doi.org/10.1177/0093650206293248

Lee, E.J. (2007). Deindividuation Effects on Group Polarization in Computer-Mediated Communication: The Role of Group Identification, Public-Self-Awareness, and Perceived Argument Quality. *Journal of Communication, 57*(2), 385-403. https://doi.org/10.1111/j.1460-2466.2007.00348.x

Lee, M. T., Poloma, M. M., & Post, S. G. (2013). *The heart of religion: Spiritual empowerment, benevolence, and the experience of God's love.* Oxford University Press.

Lenard, D. B. (2016). Gender differences in the personal pronouns usage on the corpus of congressional speeches. Journal of Research Design and Statistics in *Linguistics and Communication, 3*(2), 161-188. https://doi.org/10.1558/jrds.30111

León, J. A. M., Cantisano, G. T., Mangin, J. P. L. (2009). Leadership in nonprofit organizations of Nicaragua and El Salvador: A study from the social identity theory. *The Spanish Journal of Psychology, 12*(1), 667-676. https://doi.org/10.1017/S1138741600002031

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology, 78*(2), 223-246. https://doi.org/10.1037/0022-3514.78.2.223

Livingstone, A., & Haslam, S. A. (2008). The importance of social identity content in a setting of chronic social conflict: Understanding intergroup relations in Northern Ireland. *British Journal of Social Psychology, 47*(1), 1-21. https://doi.org/10.1348/014466607X200419

Long, R. T. (1998). Toward a Libertarian theory of class. *Social Philosophy and Policy, 15*(2), 303-347. https://doi.org/10.1017/s0265052500002028

Luhtanen, R. K. (2003). Identity, stigma management, and well-being: A comparison of lesbians/bisexual women and gay/bisexual men. *Journal of Lesbian Studies, 7*(1), 85-100. https://doi.org/10.1300/J155v07n01_06

Luhtanen, R., & Crocker, J. (1992). A Collective Self-Esteem Scale: Self-Evaluation of One's Social Identity. *Personality and Social Psychology Bulletin, 18*(3), 302–318. https://doi.org/10.1177/0146167292183006

Löckenhoff, C. E., Costa, P. T., & Lane, R. D. (2008). Age differences in descriptions of emotional experiences in oneself and others. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 63*(2). https://doi.org/10.1093/geronb/63.2.P92

Mackenzie, J. (2017). Identifying informational norms in Mumsnet Talk: A reflexive-linguistic approach to internet research ethics. *Applied Linguistics Review, 8*(3), 293-314. https://doi.org/10.1515/applirev-2016-1042

Mackie, D. M. (1986). Social Identification Effects in Group Polarization. Journal of

Personality and Social Psychology, 50(4), 720-728.

https://doi.org/10.1037/0022-3514.50.4.720

Mackie, D., & Cooper, J. (1984). Attitude polarization: Effects of group membership.

Journal of Personality and Social Psychology, 46(3), 575-585.

https://doi.org/10.1037/0022-3514.46.3.575

Maddox, A., Barratt, M. J., Allen, M., & Lenton, S. (2016). Constructive activism in

the dark web: cryptomarkets and illicit drugs in the digital 'demimonde'.

Information, Communication & Society, 19(1), 111-126.

https://doi.org/10.1080/1369118X.2015.1093531

Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic

cues for the automatic recognition of personality in conversation and text.

Journal of Artificial Intelligence Research, 30(1), 457-500.

https://doi.org/10.1613/jair.2349

Majumder, N., Poria, S., Gelbukh, A., & Cambria, E. (2017). Deep learning-based

document modeling for personality detection from text. IEEE Intelligent

Systems, 32(2), 74-79. https://doi.org/10.1109/MIS.2017.23

Massanari, A. L. (2016). Socialbots and Their Friends. Routledge.

https://doi.org/10.4324/9781315637228-13

Masson, K., & Bancroft, A. (2018). 'Nice people doing shady things': Drugs and the morality of exchange in the darknet cryptomarkets. *International Journal of Drug Policy, 58*(1), 78-84. https://doi.org/10.1016/j.drugpo.2018.05.008

McGarty, C. (1999) *Social Categorization in Social Psychology*. Sage

McGarty, C., Turner, J. C., Hogg, M. A., David, B., & Wetherell, M. S. (1992). Group polarization as conformity to the prototypical group member. *British Journal of Social Psychology, 31*(1), 1-19. https://doi.org/10.1111/j.2044-8309.1992.tb00952.x

McGarty, C., Yzerbyt, V., & Spears, R. (Eds.). (2002). *Stereotypes as Explanations: The Formation of Meaningful Beliefs about Social Groups*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511489877

Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology, 90*(5), 862-877. https://doi.org/10.1037/0022-3514.90.5.862

Michel, J. B., Kui Shen, Y., Presser Aiden, A., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science, 331*(6014), 176-182. https://doi.org/10.1126/science.1199644

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. In Y. Bengio & Y. LeCun (Eds.), *Proceeding of the International Conference on Learning Representations Workshop Track* (pp. 1301–3781).

Milkman, R. (2014). Millennial movements: Occupy wall street and the dreamers. *Dissent, 61*(3), 55-59. https://doi.org/10.1353/dss.2014.0053

Miller, D. L. (2013). *Introduction to collective behavior and collective action.* Waveland Press.

Miller, L. R., & Grollman, E. A. (2015). The social costs of gender nonconformity for transgender adults: Implications for discrimination and health. *Sociological Forum, 30*(3), 809-831. https://doi.org/10.1111/socf.12193

Mislove, A. E. (2009). *Online social networks: Measurement, analysis, and applications to distributed information systems* [Doctoral Dissertation, Rice University]. Rice Scholarship. https://scholarship.rice.edu/handle/1911/61861.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational intelligence, 29*(3), 436-465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Mohammad, S., & Turney, P. (2010, June). Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In D. Inkpen & C. Strapparava (Eds.), *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text* (pp. 26-34). Association for Computational Linguistics.

Moscovici, S., & Zavalloni, M. (1969). The group as a polarizer of attitudes. *Journal of Personality and Social Psychology, 12*(2), 125-135. https://doi.org/10.1037/h0027568

Mulac, A., Blau, S. and Bauquier, L. (1983) Gender-linked language differences and their effects in male and female students' impromptu essays. [Paper presentation]. *Meeting of the Speech Communication Association*, Washington, D.C.

Mulac, A., Bradac, J., & Gibbons, P. (2006). Empirical support for the gender-as-culture hypothesis. *Human Communication Research, 27*(1), 121-152. https://doi.org/10.1111/j.1468-2958.2001.tb00778.x

Mulac, A., Seibold, D. R., & Farris, J. L. (2000). Female and male managers' and professionals' criticism giving: Differences in language use and effects. *Journal of Language and Social Psychology, 19*(4), 389-415.

Mulac, A., Studley, L. B., & Blau, S. (1990). The gender-linked language effect in

    primary and secondary students' impromptu essays. *Sex Roles: A Journal of*

    Research, 23(9-10), 439-469. https://doi.org/10.1007/BF00289762

Munger, T., & Zhao, J. (2015). Identifying influential users in on-line support forums

    using topical expertise and social network analysis. In J. Pei, F. Silvestri, & J.

    Tang (Eds.), *2015 IEEE/ACM International Conference on Advances in*

    Social Networks Analysis and Mining (ASONAM) (pp. 721–728). Association

    for Computing Machinery. https:// doi.org/10.1145/2808797.2810059.

Munksgaard, R., & Demant, J. (2016). Title: Mixing politics and crime-the prevalence

    and decline of political discourse on the cryptomarket. *International Journal*

    of Drug Policy, 25(1), 77-83. https://doi.org/10.1016/j.drugpo.2016.04.021

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon.

    Psychological Bulletin, 83(4), 602-627. https://doi.org/10.1037/0033-

    2909.83.4.602

Nagel, T. (1995). *Equality and partiality.* Oxford University Press.

Nelson, S. K., Robbins, M. L., Andrews, S. E., & Sweeny, K. (2017). Disrupted

    transition to parenthood: Gender moderates the association between

    miscarriage and uncertainty about conception. *Sex Roles: A Journal of*

    Research, 76(5-6), 380-392. https://doi.org/10.1007/s11199-015-0564-z

Neufeld, J., & Hargreaves, A. (1999). What's Worth Loving in Your School: Teacher Development and Universal Benevolence Changing Teachers, Changing Times: Teachers' Work and Culture in the Postmodern Age. *Canadian Journal of Education / Revue Canadienne de l'éducation, 24*(1), 87. https://doi.org/10.2307/1585775

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes, 45*(3), 211-236. https://doi.org/10.1080/01638530802073712

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665-675. https://doi.org/10.1177/0146167203029005010

Nguyen, D., & Rosé, C. (2011). Language use as a reflection of socialization in online communities. In M. Nagarajan & M. Gamon (Eds.), *Proceedings of the Workshop on Languages in Social Media* (pp 76-85). Association for Computational Linguistics.

Nguyen, D., Doğruöz, A. S., Rosé, C. P., & de Jong, F. (2016). Computational sociolinguistics: A survey. *Computational linguistics, 42*(3), 537-593. https://doi.org/10.1162/COLI_a_00258

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology, 21*(4), 337–360. https://doi.org/10.1177/026192702237953

Nielsen, F. (2016). Introduction to HPC with MPI for Data Science. Springer.

Noel, J. G., Wann, D. L., & Branscombe, N. R. (1995). Peripheral Ingroup Membership Status and Public Negativity Toward Outgroups. *Journal of Personality and Social Psychology, 68*(1), 127-137. https://doi.org/10.1037/0022-3514.68.1.127

Nolan, I. T., Kuhner, C. J., & Dy, G. W. (2019). Demographic and temporal trends in transgender identities and gender confirming surgery. *Translational andrology and urology, 8*(3), 184-190. https://doi.org/10.21037/tau.2019.04.09

Oakes, P. J. (1987). The salience of social categories. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.

Oakes, P. J., & Turner, J. C. (1990). Is limited information processing capacity the cause of social stereotyping? *European Review of Social Psychology, 1*(1), 111-135. https://doi.org/10.1080/14792779108401859

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality*. Basil Blackwell

Oakes, P. J., Turner, J. C., & Haslam, S. A. (1991). Perceiving people as group members: The role of fit in the salience of social categorizations. *British Journal of Social Psychology, 30*(2), 125-144. https://doi.org/10.1111/j.2044-8309.1991.tb00930.x

Pehrson, S., & Reicher, S. (2014). On the meaning, validity and importance of the distinction between personal and social identity: A social identity perspective on Identity Process Theory. In R. Jaspal & G. Breakwell (Eds.), *Identity Process Theory: Identity, Social Action and Social Change* (pp. 97-117). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139136983.009

Pehrson, S., Brown, R., & Zagefka, H. (2009). When does national identification lead to the rejection of immigrants? Cross-sectional and longitudinal evidence for the role of essentialist in-group definitions. *British Journal of Social Psychology, 48*(1), 61-76. https://doi.org/10.1348/014466608X288827

Pendry, L. F., & Salvatore, J. (2015). Individual and social benefits of online discussion forums. *Computers in Human Behavior, 50*, 211-220.

Peng, S., Zhou, Y., Cao, L., Yu, S., Niu, J., & Jia, W. (2018). Influence analysis in social networks: A survey. *Journal of Network and Computer Applications, 106*(1), 17–32. https://doi.org/10.1016/j.jnca.2018.01.005

Pennebaker Conglomerates Inc. (2017) *LIWC 2015: How it works.*
LIWC.wpengine.com. https://liwc. wpengine.com/how-it-works/

Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us.*
Bloomsbury Press. https://doi.org/10.1093/llc/fqt006

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. University of Texas at Austin. https://doi.org/10.15781/T29G6Z

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLoS ONE, 9*(12). https://doi.org/10.1371/journal.pone.0115844

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology, 54*(1), 547-577. https://doi.org/10.1146/annurev.psych.54.101601.145041

Pennebaker. (2011). Using computer analyses to identify language style and aggressive intent: the secret life of function words. *Dynamics of Asymmetric Conflict, 4*(2), 92-102. https://doi.org/10.1080/17467586.2011.627932

Petrie, K. J., Pennebaker, J. W., & Sivertsen, B. (2008). Things we said today: A linguistic analysis of the Beatles. *Psychology of Aesthetics, Creativity, and the Arts, 2*(4), 197-202. https://doi.org/10.1037/a0013117

Pierro, A., Cicero, L., & Tory Higgins, E. (2009). Followers' satisfaction from working with group-prototypic leaders: Promotion focus as moderator. Journal of *Experimental Social Psychology, 45*(5), 1105-1110. https://doi.org/10.1016/j.jesp.2009.06.020

Platow, M. J., & Van Knippenberg, D. (2001). A social identity analysis of leadership endorsement: The effects of leader ingroup prototypicality and distributive intergroup fairness. *Personality and Social Psychology Bulletin, 27*(11), 1508-1519. https://doi.org/10.1177/01461672012711011

Platow, M. J., van Knippenberg, D. (2001). A social identity analysis of leadership endorsement: The effects of leader ingroup prototypicality and distributive intergroup fairness. *Personality and Social Psychology Bulletin, 27*(1), 1508-1519.

Platow, M. J., Van Knippenberg, D., Haslam, S. A., Van Knippenberg, B., & Spears, R. (2006). A special gift we bestow on you for being representative of us: Considering leader charisma from a self-categorization perspective. *British Journal of Social Psychology, 45*(2), 303-320. https://doi.org/10.1348/014466605X41986

Plutchik, R. (1984). Emotions: A general psychoevolutionary theory. In K. R. Scherer

    & P. Ekman (Eds.), *Approaches to emotion* (pp. 197-219). Lawrence

    Erlbaum Associates

Ponizovskiy, V., Ardag, M., Grigoryan, L., Boyd, R., Dobewall, H., & Holtz, P. (2020).

    Development and Validation of the Personal Values Dictionary: A Theory–

    Driven Tool for Investigating References to Basic Human Values in Text.

    European Journal of Personality, 34(5), 885–902.

    https://doi.org/10.1002/per.2294

Porat, R., Halperin, E., & Tamir, M. (2016). What we want is what we get: Group-

    based emotional preferences and conflict resolution. *Journal of Personality*

    and Social Psychology, 110(2), 167–190.

    https://doi.org/10.1037/pspa0000043

Postmes, T, Spears, R., & Lea, M. (2000). The formation of group norms in

    computer-mediated communication. *Human Communication Research,*

    26(3), 341-371. https://doi.org/10.1111/j.1468-2958.2000.tb00761.x

Postmes, T., & Jetten, J. (2006). *Individuality and the group: Advances in social*

    identity. Sage.

Postmes, T., Haslam, S. A., & Swaab, R. I. (2005). Social influence in small groups:

    An interactive model of social identity formation. *European Review of Social*

    Psychology, 16(1), 1-42. https://doi.org/10.1080/10463280440000062

Postmes, T., Spears, R., & Cihangir, S. (2001b). Quality of decision making and

   group norms. *Journal of Personality and Social Psychology, 80*(6), 918–930.

   https://doi.org/10.1037/0022-3514.80.6.918

Postmes, T., Spears, R., & Lea, M. (2000). The formation of group norms in

   computer-mediated communication. *Human communication research, 26*(3),

   341-371. https://doi.org/10.1111/j.1468-2958.2000.tb00761.x

Postmes, T., Spears, R., & Lea, M. (2002). Intergroup differentiation in computer-

   mediated communication: Effects of depersonalization. *Group Dynamics,

   6*(1), 3-16. https://doi.org/10.1037/1089-2699.6.1.3

Postmes, T., Spears, R., Lee, A. T., & Novak, R. J. (2005b). Individuality and social

   influence in groups: Inductive and deductive routes to group identity. *Journal

   of Personality and Social Psychology, 89*(5), 747-763.

   https://doi.org/10.1037/0022-3514.89.5.747

Postmes, T., Spears, R., Sakhel, K., & De Groot, D. (2001). Social influence in

   computer-mediated communication: The effects of anonymity on group

   behavior. *Personality and Social Psychology Bulletin, 27*(10), 1243-1254.

   https://doi.org/10.1177/01461672012710001

Prentice, D. A., Miller, D. T., & Lightdale, J. R. (1994). Asymmetries in attachments

   to groups and to their members: Distinguishing between common-identity

and common-bond groups. *Personality and Social Psychology Bulletin, 20*(5), 484-493. https://doi.org/10.1177/0146167294205005

Razaque, A., Rizvi, S., Almiani, M., & Al Rahayfeh, A. (2019). State-of-art review of information diffusion models and their impact on social network vulnerabilities. *Journal of King Saud University-Computer and Information Sciences.* Advance online publication. https://doi.org/10.1016/j. jksuci.2019.08.008

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science, 5*(1), 31. https://doi.org/10.1140/epjds/s13688-016-0093-1

Reicher, S. (1984a). Social influence in the crowd: Attitudinal and behavioural effects of de-individuation in conditions of high and low group salience. *British Journal of Social Psychology, 23*(4), 341-350. https://doi.org/10.1111/j.2044-8309.1984.tb00650.x

Reicher, S. (1984b). The St. Pauls' riot: An explanation of the limits of crowd action in terms of a social identity model. *European Journal of Social Psychology, 14*(1), 1–21. https://doi.org/10.1002/ejsp.2420140102

Reicher, S. (2004). The Context of Social Identity: Domination, Resistance, and Change. *Political Psychology, 25*(6), 921-945. https://doi.org/10.1111/j.1467-9221.2004.00403.x

Reicher, S. (2008). The Psychology of Crowd Dynamics. In M.A. Hogg and R.S. Tindale (Eds.), *Blackwell Handbook of Social Psychology: Group Processes* (pp. 182-208). Blackwell Publishers Ltd. https://doi.org/10.1002/9780470998458.ch8

Reicher, S. D., Spears, R., & Postmes, T. (1995). A Social Identity Model of Deindividuation Phenomena. *European Review of Social Psychology, 6*(1), 161-198. https://doi.org/10.1080/14792779443000049

Reicher, S., & Hopkins, N. (1996). Seeking influence through characterizing self-categories: An analysis of anti-abortionist rhetoric. *British Journal of Social Psychology, 35*(2), 297-311. https://doi.org/10.1111/j.2044-8309.1996.tb01099.x

Reicher, S., & Hopkins, N. (2003). On the science of the art of leadership. In D. van Knippenberg, & M. A. Hogg (Eds.), *Leadership and power: Identity processes in groups and organizations* (pp. 197-209). SAGE Publications Ltd, https://www.doi.org/10.4135/9781446216170.n15

Reicher, S., Levine, R. M., & Gordijn, E. (1998). More on deindividuation, power relations between groups and the expression of social identity: Three studies on the effects of visibility to the in-group. *British Journal of Social Psychology, 37*(1), 15-40. https://doi.org/10.1111/j.2044-8309.1998.tb01155.x

Reicher, S., Haslam, S. A., & Hopkins, N. (2005). Social identity and the dynamics of leadership: Leaders and followers as collaborative agents in the transformation of social reality. *Leadership Quarterly, 16*(4), 547-568. https://doi.org/10.1016/j.leaqua.2005.06.007

Reynolds, K. J., Bizumic, B., Subasic, E., Turner, J. C., Branscombe, N., Mavor, K. I., & Batalha, L. (2012). Social identity and personality processes: Non-aboriginal Australian identity and neuroticism. *European Journal of Social Psychology, 42*(2), 252-262. https://doi.org/10.1002/ejsp.1841

Reysen, S., & Branscombe, N. R. (2008). Belief in collective emotions as conforming to the group. *Social Influence, 3*(3), 171-188. https://doi.org/10.1080/15534510802247438

Reysen, S., & Katzarska-Miller, I. (2013). A model of global citizenship: Antecedents and outcomes. *International Journal of Psychology, 48*(5), 858-870. https://doi.org/10.1080/00207594.2012.701749

Reysen, S., Plante, C. N., Roberts, S. E., & Gerbasi, K. C. (2015). A social identity perspective of personality differences between fan and non-fan identities. *World Journal of Social Science Research, 2*(1). https://doi.org/10.22158/wjssr.v2n1p91

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior, 29*(5), 615–620. https://doi.org/10.1007/s10979-005-6832-7

Rickford, J. R. (2011). Le Page's Theoretical and Applied Legacy in Sociolinguistics. In L. Hinrichs & J. T. Farquharson (Eds.), *Variation in the Caribbean: From Creole Continua to Individual Agency* (pp. 251-272). John Benjamins Publishing Company

Rosch, E. (1978). Principles of Categorization. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27-48), Lawrence Erlbaum Associates Publishers.

Rude, S. S., Gortner, E. M., & Pennebaker, J. W. (2004). Language use of depressed and depression-vulnerable college students. *Cognition and Emotion, 18*(8), 1121-1133. https://doi.org/10.1080/02699930441000030

Ryan, C. S., & Bogart, L. M. (2001). Longitudinal changes in the accuracy of new group members' in-group and out-group stereotypes. *Journal of Experimental Social Psychology, 37*(2), 118-133. https://doi.org/10.1006/jesp.2000.1439

Ryan, E. B., Giles, H., Bartolucci, G., & Henwood, K. (1986). Psycholinguistic and social psychological components of communication by and with the elderly.

Language and Communication, 6(2), 1-24. https://doi.org/10.1016/0271-5309(86)90002-9

Sagiv, L., Roccas, S., & Hazan, O. (2012). Identification with groups: The role of personality and context. *Journal of Personality, 80*(2), 345-374. https://doi.org/10.1111/j.1467-6494.2011.00733.x

Saroglou, V., Delpierre, V., & Dernelle, R. (2004). Values and religiosity: A meta-analysis of studies using Schwartz's model. *Personality and individual differences, 37*(4), 721-734. https://doi.org/10.1016/j.paid.2003.10.005

Sassenberg, K., & Postmes, T. (2002). Cognitive and strategic processes in small groups: Effects of anonymity of the self and anonymity of the group on social influence. *British Journal of Social Psychology, 41*(3), 463-480. https://doi.org/10.1348/014466602760344313

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E. P., Ungar, L. H., & Preis, T. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLoS One, 8*(9), https://doi.org/10.1371/journal.pone.0073791

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–65), Academic Press.

Schwartz, S. H., & Boehnke, K. (2004). Evaluating the structure of human values with confirmatory factor analysis. *Journal of Research in Personality, 38*(3), 230-255. https://doi.org/10.1016/S0092-6566(03)00069-2

Seidman, S. (2013). Defilement and disgust: Theorizing the other. *American Journal of Cultural Sociology, 1*(1), 3-25. https://doi.org/10.1057/ajcs.2012.3

Simon, B. (1993). On the asymmetry in the cognitive construal of ingroup and outgroup: A model of egocentric social categorization. *European Journal of Social Psychology, 23*(2), 131-147. https://doi.org/10.1037/a0014717

Simon, B., & Stürmer, S. (2003). Respect for group members: Intragroup determinants of collective identification and group-serving behavior. *Personality and Social Psychology Bulletin, 29*(2), 183-193. https://doi.org/10.1177/0146167202239043

Sinclair, S., Hardin, C. D., & Lowery, B. S. (2006). Self-stereotyping in the context of multiple social identities. *Journal of Personality and Social Psychology, 90*(4), 529–542. https://doi.org/10.1037/0022-3514.90.4.529

Slatcher, R. B., Vazire, S., & Pennebaker, J. W. (2008). Am "I" more important than "we"? Couples' word use in instant messages. *Personal Relationships, 15(*4), 407-424. https://doi.org/10.1111/j.1475-6811.2008.00207.x

Sluss, D. M. (2006). *Generalizing relational identification to and from organizational identification* [Doctoral dissertation, Arizona State University]. ProQuest Dissertations Publishing.

Smith, E. R., Seger, C. R., & Mackie, D. M. (2007). Can emotions be truly group level? Evidence regarding four conceptual criteria. *Journal of Personality and Social Psychology, 93*(3), 431–446. https://doi.org/10.1037/0022-3514.93.3.431

Smith, L. G. E., & Postmes, T. (2009). Intra-group interaction and the development of norms which promote inter-group hostility. *European Journal of Social Psychology, 39*(1), 130-144. https://doi.org/10.1002/ejsp.464

Smith, L. G. E., & Postmes, T. (2011). The power of talk: Developing discriminatory group norms through discussion. *British Journal of Social Psychology, 50*(2), 193-215. https://doi.org/10.1348/014466610X504805

Smith, L. G. E., Gavin, J., & Sharp, E. (2015a). Social identity formation during the emergence of the occupy movement. *European Journal of Social Psychology, 45*(7), 818-832. https://doi.org/10.1002/ejsp.2150

Smith, L. G. E., Thomas, E. F., & McGarty, C. (2015b). "We Must Be the Change We Want to See in the World": Integrating Norms and Identities through Social Interaction. *Political Psychology, 36*(5), 543-557. https://doi.org/10.1111/pops.12180

Smith, R. M. (2004). Identities, interests, and the future of political science. *Perspectives on Politics, 2*(2), 301-312. https://doi.org/10.1017/S1537592704040174

Snider, J. S., Reysen, S., & Katzarska-Miller, I. (2013). How we frame the message of globalization matters. *Journal of Applied Social Psychology, 43*(8), 1599-1607. https://doi.org/10.1111/jasp.12111

Soldatic, K., & Meekosha, H. (2012). The place of disgust: Disability, class and gender in spaces of workfare. *Societies, 2*(3), 139-156. https://doi.org/10.3390/soc2030139

Song, X., Chi, Y., Hino, K., & Tseng, B. (2007). Identifying opinion leaders in the blogosphere. In *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management* (pp. 971–974). Association for Computing Machinery. https://doi. org/10.1145/1321440.1321588

Spears, R., & Lea, M. (1994). Panacea or Panopticon? *Communication Research, 21*(4), 427-459. https://doi.org/10.1177/009365094021004001

Spears, R., Doosje, B., & Ellemers, N. (1997). Self-stereotyping in the face of threats to group status and distinctiveness: The role of group identification. *Personality and Social Psychology Bulletin, 23*(5), 538-553. https://doi.org/10.1177/0146167297235009

Spears, R., Doosje, B., & Ellemers, N. (1999). Commitment and the context of social perception. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content* (pp. 59-83). Blackwell Science.

Spears, R., Lea, M., Corneliussen, R. A., Postmes, T., & Haar, W. Ter. (2002). Computer-Mediated Communication as a Channel for Social Resistance. *Small Group Research, 33*(5), 555-574. https://doi.org/10.1177/104649602237170

Steffens, N. K., Haslam, S. A., & Reicher, S. D. (2014). Up close and personal: Evidence that shared social identity is a basis for the "special" relationship that binds followers to leaders. *Leadership Quarterly, 25(2*), 296-313. https://doi.org/10.1016/j.leaqua.2013.08.008

Steffens, N. K., Haslam, S. A., Reicher, S. D., Platow, M. J., Fransen, K., Yang, J., Ryan, M. K., Jetten, J., Peters, K., & Boen, F. (2014). Leadership as social identity management: Introducing the Identity Leadership Inventory (ILI) to assess and validate a four-dimensional model. *Leadership Quarterly, 25*(5), 1001-1024. https://doi.org/10.1016/j.leaqua.2014.05.002

Steffens, N. K., Munt, K. A., van Knippenberg, D., Platow, M. J., & Haslam, S. A. (2021). Advancing the social identity theory of leadership: A meta-analytic review of leader group prototypicality. *Organizational Psychology Review, 11*(1), 35-72. https://doi.org/10.1177/2041386620962569

Steinmetz, K. (2014, May). The Transgender Tipping Point. *Time Magazine.*

   https://time.com/135480/transgender-tipping-point/

Stokols, D., Misra, S., Runnerstrom, M. G., & Hipp, J. A. (2009). Psychology in an

   age of ecological crisis: From personal angst to collective action. *American*

   *Psychologist, 64(*3), 181-193

Stryker, S., Currah, P., & Moore, L. J. (2008). Introduction: Trans-, trans, or

   transgender?. *Women's Studies Quarterly, 36*(3), 11-22.

   https://doi.org/10.1353/wsq.0.0112

Tajfel, H. (1981). *Human Groups and Social Categories: Studies in Social*

   Psychology. Cambridge University Press.

Tanis, M., & Postmes, T. (2005). A social identity approach to trust: interpersonal

   perception, group membership and trusting behaviour. *European Journal of*

   Social Psychology, 35(3), 413-424. https://doi.org/10.1002/ejsp.256

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words:

   LIWC and Computerized Text Analysis Methods. *Journal of Language and*

   Social Psychology, 29(1), 24-54. https://doi.org/10.1177/0261927X09351676

Taylor, J. K., Haider-Markel, D. P., & Lewis, D. C. (2018). *The remarkable rise of transgender rights.* University of Michigan Press. https://doi.org/10.3998/mpub.9448956

Thakerar, J. N., Giles, H., & Cheshire, J. (1982). Psychological and linguistic parameters of speech accommodation theory. In C. Fraser & K. R. Scherer (Eds.), *Advances in the social psychology of language* (pp. 205-255). Cambridge University Press.

Thoits, P. A. (1992). Identity structures and psychological well-being: Gender and marital status comparisons. *Social Psychology Quarterly, 55*(3), 236–256. https://doi.org/10.2307/2786794

Thomas, E. F., McGarty, C., & Mavor, K. I. (2009a). Aligning identities, emotions, and beliefs to create commitment to sustainable social and political action. *Personality and Social Psychology Review, 13*(3), 194-218. https://doi.org/10.1177/1088868309341563

Thomas, E. F., McGarty, C., & Mavor, K. I. (2009). Transforming "apathy into movement": The role of prosocial emotions in motivating action for social change. *Personality and Social Psychology Review, 13*(4), 310-333. https://doi.org/10.1177/1088868309343290

Toma, C. L., & D'Angelo, J. D. (2015). Tell-tale words: Linguistic cues used to infer the expertise of online medical advice. *Journal of Language and Social Psychology, 34*(1), 25-45. https://doi.org/10.1177/0261927X14554484

Tong, S. T., Corriero, E. F., Wibowo, K. A., Makki, T. W., & Slatcher, R. B. (2020). Self-presentation and impressions of personality through text-based online dating profiles: A lens model analysis. *New Media & Society, 22*(5), 875-895. https://doi.org/10.1177/1461444819872678

Turner, J. (1982). Toward a cognitive definition of the group. In H. Tajfel (Ed.), *Social Identity and Intergroup Relations*. Cambridge University Press

Turner, J. C. (1985). Social categorization and the self-concept: A social-cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research* (pp. 77-122). JAI Press.

Turner, J. C. (1991). *Mapping social psychology series. Social influence.* Thomson Brooks/Cole Publishing Co.

Turner, J. C., & Haslam, S. A. (2001). Social identity, organizations, and leadership. In M. E. Turner (Ed.), *Applied social research. Groups at work: Theory and research* (pp. 25-65). Lawrence Erlbaum Associates Publishers.

Turner, J. C., & Onorato, R. S. (1999). Social identity, personality, and the self-concept: A self-categorizing perspective. In T. R. Tyler, R. M. Kramer, & O.

P. John (Eds.), *Applied social research. The psychology of the social self* (p. 11–46). Lawrence Erlbaum Associates Publishers.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and Collective: Cognition and social context. *Personality & Social Psychology Bulletin, 20*(5), 454-463. https://doi.org/10.1177/0146167294205002

Turner, J. C., Wetherell, M. S., & Hogg, M. A. (1989). Referent informational influence and group polarization. *British Journal of Social Psychology, 28*(2), 135-147. https://doi.org/10.1111/j.2044-8309.1989.tb00855.x

Ullrich, J., Christ, O., & van Dick, R. (2009). Substitutes for Procedural Fairness: Prototypical Leaders Are Endorsed Whether They Are Fair or Not. *Journal of Applied Psychology, 94*(1), 235-244. https://doi.org/10.1037/a0012936

United Nations Office on Drugs and Crime, World Drug Report (2018) *World Drug Report: Executive summary, conclusions and policy implications (Report no. 1)* Retrieved from https://www.unodc.org/wdr2018/prelaunch/WDR18_Booklet_1_EXSUM.pdf

Van Kleef, G. A., Steinel, W., Van Knippenberg, D., Hogg, M. A., & Svensson, A. (2007). Group member prototypicality and intergroup negotiation: How one's

standing in the group affects negotiation behaviour. *British Journal of Social Psychology, 46*(1), 129-152. https://doi.org/10.1348/014466605X89353

Van Knippenberg, B., & Van Knippenberg, D. (2005). Leader self-sacrifice and leadership effectiveness: The moderating role of leader prototypicality. *Journal of Applied Psychology, 90*(1), 25-37. https://doi.org/10.1037/0021-9010.90.1.25

Van Knippenberg, D. (2011). Embodying who we are: Leader group prototypicality and leadership effectiveness. *Leadership Quarterly, 22*(6), 1078-1091. https://doi.org/10.1016/j.leaqua.2011.09.004

Van Knippenberg, D., & Hogg, M. A. (2003). A Social Identity Model of Leadership Effectiveness in Organisations. *Research in Organizational Behavior, 25*(1), 243-295). https://doi.org/10.1016/S0191-3085(03)25006-1

Van Knippenberg, D., Lossie, N., & Wilke, H. (1994). In-group prototypicality and persuasion: Determinants of heuristic and systematic message processing. British *Journal of Social Psychology, 33*(3), 289-300. https://doi.org/10.1111/j.2044-8309.1994.tb01026.x

Van Veelen, R., Otten, S., Cadinu, M., & Hansen, N. (2016). An Integrative Model of Social Identification: Self-Stereotyping and Self-Anchoring as Two Cognitive Pathways. *Personality and Social Psychology Review, 20*(1), 3-26. https://doi.org/10.1177/1088868315576642

van Zomeren, M., Spears, R., Fischer, A. H., & Leach, C. W. (2004). Put Your Money Where Your Mouth Is! Explaining Collective Action Tendencies Through Group-Based Anger and Group Efficacy. *Journal of Personality and Social Psychology, 87*(5), 649–664. https://doi.org/10.1037/0022-3514.87.5.649

Verbeek, M. J., Hommes, M. A., Stutterheim, S. E., van Lankveld, J. J., & Bos, A. E. (2020). Experiences with stigmatization among transgender individuals after transition: A qualitative study in the Netherlands. *International Journal of Transgender Health, 21*(2), 220-233. https://doi.org/10.1080/26895269.2020.1750529

Verkuyten, M., & Hagendoorn, L. (1998). Prejudice and Self-Categorization: The Variable Role of Authoritarianism and In-Group Stereotypes. *Personality and Social Psychology Bulletin, 24*(1), 99-110. https://doi.org/10.1177/0146167298241008

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

Wang, G., Jiang, W., Wu, J., & Xiong, Z. (2013). Finegrained feature-based social influence evaluation in online social networks. *IEEE Transactions on Parallel and Distributed Systems, 25*(9), 2286–2296. https://doi.org/10.1109/TPDS.2013.135

Welbers, K., & de Nooy, W. (2014). Stylistic Accommodation on an Internet Forum as Bonding. *American Behavioral Scientist, 58*(10), 1361-1375. https://doi.org/10.1177/0002764214527086

White House Press Office (2016). *FACT SHEET: Obama Administration's Record and the LGBT Community* [Press release]. https://obamawhitehouse.archives.gov/the-press-office/2016/06/09/fact-sheet-obama-administrations-record-and-lgbt-community

Wilson, S. (2019). *Natural language processing for personal values and human activities* [Doctoral dissertation, University of Michigan]. Deep Blue. https://deepblue.lib.umich.edu/handle/2027.42/150025

Witcomb, G. L., Bouman, W. P., Claes, L., Brewin, N., Crawford, J. R., & Arcelus, J. (2018). Levels of depression in transgender people and its predictors: Results of a large matched control study with transgender people accessing clinical services. *Journal of Affective Disorders, 235*, 308-315. https://doi.org/10.1016/j.jad.2018.02.051

Wood, W. (2000). Attitude Change: Persuasion and Social Influence. *Annual Review of Psychology, 51*(1), 539-570. https://doi.org/10.1146/annurev.psych.51.1.539

World Health Organization. (2019). *International statistical classification of diseases and related health problems (11th ed.)*. https://icd.who.int/

Yang, C. C., Tang, X., & Thuraisingham, B. M. (2010). An analysis of user influence ranking algorithms on dark web forums. In *Association for Computing Machinery (ACM) Special Interest Group on Knowledge Discovery and Data Mining Workshop on Intelligence and Security Informatics* (pp. 10–16). Association for Computing Machinery. https://doi.org/10.1145/1938606.1938616

Yarkoni, T. (2010). Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers. *Journal of Research in Personality, 44(*3), 363-373. https://doi.org/10.1016/j.jrp.2010.04.001

Young, I. M. (2011). *Justice and the Politics of Difference.* Princeton University Press.

Zhou, Y., & Liu, L. (2015). Social influence based clustering and optimization over heterogeneous information networks. *ACM Transactions on Knowledge Discovery from Data (TKDD), 10*(1), 1–53. https://doi.org/10.1145/2717314

Appendices

Appendix A: LIWC Linguistic Style Categories

Outlined in Figure A1 are the LIWC stylistic categories used in this research. These categories are part-of-speech categories that are used widely regardless of topic.

Figure A1

41 LIWC stylistic variables used initial analysis

Pronouns	Function words	Other grammar	Psychological	Time	Perception/Relativity	Structural/ Informal
1st person singular	Articles	Adjectives	Affect	Past focus	See	Words per sentence
1st person plural	Prepositions	Verbs	Insight	Future focus	Hear	6 letter words
2nd person	Auxiliary verbs	Interrogatives	Causation	Present focus	Feel	Swear
3rd person singular	Adverbs	Numbers	Discrepancy		Motion	Netspeak
3rd person plural	Conjunctions	Quantifiers	Tentative		Space	Assent
Impersonal pronouns	Negations	Comparisons	Certainty		Time	Non-fluencies
			Differentiation			Fillers

Appendix B: Diagram of K-fold Cross Validation

Figure A2
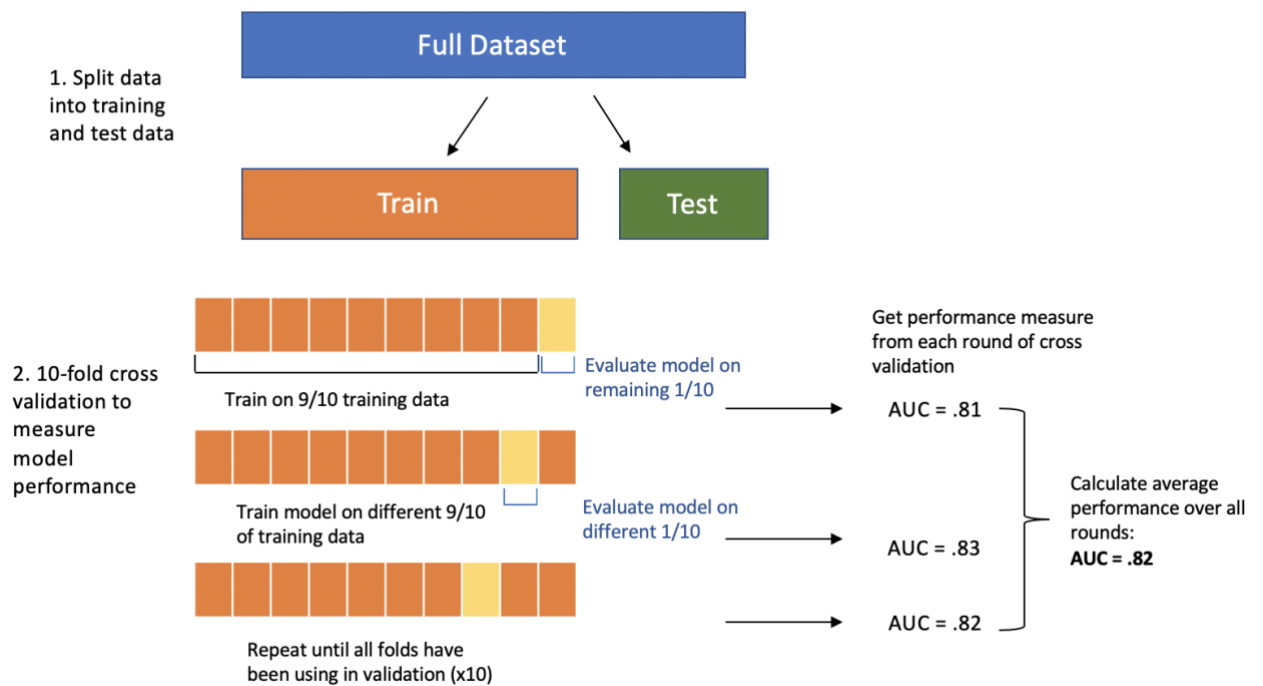
Diagram to illustrate cross validation procedure

Figure A2 demonstrates the cross-validation procedure. This process allows us to understand the performance of the training model before it is tested on the test data. We can use cross validation to prevent the model overfitting the training data.

Appendix C: Multidimensional Scaling Using Function Words and Pronouns

Figure A3 below depicts the outcome of completing the multidimensional scaling analysis from study 4.1 when using only the 12 function words and pronoun categories outlined in Figure A1 above. Here we can see that even when we use relatively few linguistic categories, the general pattern still stands thus speaking to the robustness of this linguistic style approach.

Figure A3

Multidimensional Scaling Using Function Word and Pronoun Categorie