universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

"Vermin, Parasites, Puppeteers - The Contribution of Antisemitic Metaphors to Social Categorization"

verfasst von / submitted by

Elena Annemarie Marx, B.A.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 066 867
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Allgemeine Linguistik: Grammatiktheorie und kognitive Sprachwissenschaft
Betreut von / Supervisor:	Univ.-Prof. Dr. Daniel Büring

Table of Content

Table of Figures

1. Introduction

Antisemitic metaphors are diverse in many respects: While some involve zoomorphisms such as the characterization of Jews as pigs or parasites, others pick up the subject of systemic conspiracy theories. Such metaphors refer to Jews as puppeteers or – by using both metaphorical attributions – as a web spinning spider, a giant kraken or a snake that entwines the entire globe. Irrespective of the individual content of antisemitic metaphors, their different historical genesis and their prevalence in public discourse, one commonality strikes the eye: Jews are identified across the board as a social outgroup whose members share negative properties and behaviours and, furthermore, can be sharply delineated from other members of society. In this respect, antisemitic metaphors do not only convey particular antisemitic beliefs (i.e. their explicit content) but additionally mark out general aspects of the structure and identity of a Jewish social category. Arguably, antisemitic metaphors may have a deep and fundamental influence on social cognition as they modify people's mental representation of Jews as a social category and induce further reasoning processes about our social world.

The present master's thesis picks up on this particular thought: At its core, the overarching objective of the thesis is to establish the ways in which antisemitic metaphors unfold their impact on the representation of a Jewish social category from a cognitive point of view. In this regard, I suggest that antisemitic metaphors represent powerful devices to shape social knowledge representations because, aside from their negative content, they bring together two linguistic elements which have extensive cognitive effects – that is to say, a generic and a metaphorical component. In my thesis, I outline how each of these two components promotes certain ways of representing a Jewish social group in the conceptual system and how they invoke further reasoning processes about the alleged properties and behaviours of Jewish people. Importantly, I do not assume that antisemitic metaphors inevitably result in explicit antisemitic beliefs. For the most part, people are driven by their egalitarian beliefs and by pro-social norms so that the explicit content of antisemitic metaphors can easily be countered by use of self-regulation and cognitive control (Amodio, 2014). Instead, I argue that antisemitic metaphors take their full effect in more subtle and covert ways, arising from conceptual structure rather than content: They render the boundaries of our social representations clear and sharp, increase the inductive potential of a Jewish social group as regards their characteristic properties and, furthermore, reinforce the perception of homogeneity between category members. In that sense, antisemitic metaphors exert their main influence on an implicit and unconscious level of cognition so that their effects are all the more pervasive and may have an impact on the intuitive judgements and behaviours of each and every one of us. As such,

antisemitic metaphors provide a strong foundation for the growing trend of endorsing prejudices and stereotypes against Jewish people and, in the extreme, for hostility, antisemitic discrimination, or even war and genocide.

This master's thesis divides into two theoretical parts, one focusing on social cognition and the other on language: In the first part, some general aspects of social categorization will be reviewed in order to understand how social concepts are generally represented in the conceptual system. While sorting individuals into different social categories is a useful tool to "organize our vast knowledge about human attributes and about the complex relationship networks that comprise human social life" (Lieberman, Woodward, & Kinzler, 2017, p. 556), it also comes with serious consequences: People tend to represent social categories other than their own as immutable, homogenous entities, with sharp category boundaries and a fixed set of properties and behaviours that holds true for all category members. Furthermore, such properties are often foregrounded in their intuitive judgements and in further reasoning processes. Thus, even though social categorization provides the basis to successfully navigate our social world, it is also the starting point for a range of negative psychological phenomena including stereotyping, prejudices as well as dehumanization and hostile behaviour towards particular social groups (Fiske & Russell, 2010; Lieberman et al., 2017). To understand such downstream negative attitudes is therefore to understand how a social category's immutability and homogeneity arise in the first place. Within the categorization research literature, concepts are generally conceived as the "mechanisms or organs of the mind with specific structures that afford specific ways of thinking" (Prasada, 2010, p. 55). In that sense, concepts are not only part and parcel of broader mental structures, they are, moreover, associated with particular reasoning styles and mechanisms, providing the routes along which people think about their social others. Against this background, the first section of the present thesis will give a detailed overview over different cognitive structures and mechanisms that are assumed to be implemented in the representation of social categories. It will be described in further detail how they constrain the structure of a social concept so that representations become increasingly immutable, homogeneous and conceptually rich.

The second section of this thesis will turn towards the question how language contributes to social categorization. More precisely, it will be examined in which ways the different constraints on social category representation are driven by linguistic input. In general, there is broad consensus that language plays a crucial role in social knowledge and the transmission of cultural conceptions (e.g. Gelman & Roberts, 2017; Rhodes, Leslie, & Tworek, 2012): For one thing, language may supply us with the particular contents of our social concepts. However, the

impact of language reaches far beyond merely establishing associations between a social group in their alleged properties and behaviours: Language can furthermore operate as a cognitive tool that shapes the way we think and reason about a social category. In other words, there is a variety of linguistic devices which boost the aforementioned constraints on social categorization and rework the structure of our social category representations. Recalling that the central aim of this master's thesis is to establish how antisemitic metaphors exert a significant impact on the representation of a Jewish social category, the second part of the thesis maps out two linguistic devices that are essential components of antisemitic metaphors: genericity and metaphoricity.

Firstly, it has been observed that a large part of our generic knowledge is expressed in the form of generic statements. Generics thus offer a glimpse into the representation of general knowledge in the conceptual system (Carlson, 2010; Gelman, 2010). What is more, they represent a crucial means to impart and acquire generic knowledge about the world. Generics are produced and understood effortlessly at an early age (e.g. Gelman, Goetz, Sarnecka, & Flukes, 2008), and children's inferences are guided according to the difference between generic and non-generic input (e.g. Graham, Nayer, & Gelman, 2011). In light of such evidence, generics have been associated with some of the constraints on social concepts as outlined in the first section of the thesis. Antisemitic metaphors display many of the characteristics of generic statements so that the first chapters of the second section will explain their influence on people's representation of a Jewish social concept by virtue of their generic nature.

Secondly, metaphorical language represents a central phenomenon in general cognition: Metaphors have been identified as a fundamental part of our everyday discourses (Lakoff & Johnson, 1980a; Ortony & Fainsilber, 1987) and, above all, they seem to be particularly important when people talk and reason about rather abstract and complex topics (Gentner & Grudin, 1985; Kuhn, 1993). It is therefore unsurprising that metaphors are in use when we communicate about the complex networks that constitute our social world. However, while there is general agreement on the significance of metaphors for the organization of abstract knowledge, their actual role in language and cognition is still a matter of debate: Unresolved issues relate to what the true nature of metaphors is, how we think of their meaning and effects, and which computational mechanisms underlie their comprehension, thus enabling metaphorical reasoning from the outset. The later chapters of the second section zoom in on this last question and discuss different theories of metaphor comprehension more thoroughly. In doing so, special attention will be payed to the question how each of the proposed mechanisms relates to the aforesaid constraints on social categorization. By and large, I argue

that metaphoricity warrants a profound impact of antisemitic metaphors on social category representation, but beyond this, it bears on different constraints than the generic component. In this way, antisemitic metaphors provide two separate roots along which people may further think and reason about Jews as a social group so that their effects are reinforced by their twofold nature.

Before cutting to the chase, it should be noted that different strategies are pursued as regards the two components of antisemitic metaphors: While the chapters on generic language make strong reference to various empirical findings of the field, the chapters on metaphors will be more concerned with the respective theoretical literature. The reason for this difference in approaches lies in different objectives that are guiding the two fields of research: As for generics, it is widely acknowledged that they play a crucial role in categorization. Much of the experimental work is therefore set out to establish in which ways generics affect category representations and subsequent reasoning processes. Put differently, experimental findings are generally linked to the sub-question regarding how social categories are represented in the conceptual system and how a category's immutability and homogeneity arise in the first place. Metaphor theories, on the other hand, are extremely vast and diverse in their basic premises so that empirical research is in large part designed to determine the most adequate theoretical claims. Given that it goes beyond the goal of this master's thesis to decide for a correct account of metaphor comprehension, experimental evidence that speaks in favour of or against each theory will be left aside and their central claims will be disentangled primarily on a theoretical level. At bottom, the guiding question of the chapters on metaphors is how each of the theories of metaphor comprehension can relate to the constraints on social categorization and may, thus, elucidate the twofold influence of antisemitic metaphors on the structure of a Jewish social category.

2. Constraints on Social Categorization

When thinking about social groups, we often have certain characteristics in mind which are shared by all or at least most of their members. We may expect them to display some physical qualities (e.g. Scandinavians are light-haired), particular behaviours (e.g. homeless people beg), psychological traits (e.g. women are empathic) or deontic properties (e.g. Jews do not eat pork). Arguably, we assume that category membership is generally determined by commonly shared properties. In consequence, all members of a social group are perceived as homogenous along various lines and their boundaries with other social groups appear to be sharp and impermeable. However, not every property of a social category seems to be of the same importance: While we readily assume that police officers wear uniforms and badges, their habit of driving around in cars is not as pivotal for identifying a person as a policeman. The observation that some features are more relevant to a concept than others is not only true for the social domain, but rather applies to conceptual knowledge in general: For instance, the property "curved" seems more relevant to a boomerang than to bananas, though valid for both concepts (Gelman & Diesendruck, 1999). Likewise, we infer that all carnivores demonstrate a common feeding behaviour, but we would not generalize their prey or any physical trait to all category members. At bottom, many categories seem to be represented with rich structure in that some of their properties are more permanent, stable and true to all of their members. All those considerations touch upon fundamental questions in the psychology of concepts: How is conceptual knowledge represented in the human conceptual system? Which are the central properties of a concept and how are they associated with it? Moreover, what are the implications for cognitive processes such as categorization and inductive reasoning?

One of the most influential theories on this matter is the so-called theory-based view according to which concepts are embedded in our general understanding of the world (e.g. Carey, 2004; Gopnik, 1998; Gopnik & Meltzoff, 1997; Murphy, 2004). Accordingly, concepts are the "constituents of **larger mental structures** – of beliefs that are formed out of them and of systems of representation such as intuitive theories" (Carey, 2004, p. 60). Emphasizing on the relation between concepts and mental structures comes with important implications (Gelman & Diesendruck, 1999): Firstly, people identify the relevant properties of a category by virtue of their intuitions and beliefs about the world. Instead of depending on statistical information that is obtained from everyday perception (e.g. by encountering a great number of non-pork eating Jews), we build upon the things we already know (e.g the existence of religious dietary rules, normativity principles, etc.). Since beliefs vary across concepts and domains, different

properties may be attributed to different concepts. In this way, naïve theories and beliefs constrain the content of conceptual representations. Secondly, broader mental structures also constrain how similarity between category members is computed (Gelman & Diesendruck, 1999). People have different cognitive mechanisms at hand that allow them to make quick and efficient category judgements, generate explanations of observed properties and draw inferences about further category members (e.g. representational mechanisms, heuristics and reasoning styles). By this means, mechanisms constrain the lines along which the relevant properties of a category are generalized to all category members.

Drawing on this theory-based perspective, the following chapter will give a detailed outline of approaches which have identified a number of naïve belief systems and cognitive mechanisms as constraints on social concepts. Importantly, while each approach is fundamentally different in its assumptions about the nature and scope of a given constraint, a broad division can be drawn between so-called internalist and externalist theories (Vasilyeva, Gopnik, & Lombrozo, 2018): Internalist constraints generally focus on the restrictive nature of a category's internal features. By contrast, externalist theories identify the important constraints outside of a social category.

2.1. Internalist Theories

A variety of internalist constraints have been proposed for the representation of conceptual knowledge: For one thing, **Psychological Essentialism** describes a conceptual bias which attributes an underlying essence to each natural and social category. People, furthermore, represent essences as the triggering causes for the observable characteristics and behaviours of a social group or a natural kind (Gelman, 2003, 2004; Gelman & Ware, 2012). The **Inherence Heuristic Hypothesis** proposes a more general cognitive mechanism which explains all sorts of observed patterns by means of the inherent features of its components (Cimpian, 2015; Cimpian & Salomon, 2014). Generally speaking, the Inherence Heuristic is a broad intuitive reasoning process which highlights the inherent properties of a concept and appeals to their inductive potential. As a third approach, the **Aspect Hypothesis** assumes that connections between a category and its internal properties are represented differently in the conceptual system (Prasada & Dillingham, 2006, 2009). Connections between properties and individual instances of a category are referred to as statistical connections, whereas principled connections relate to the properties of an entire category. Accordingly, only the latter are represented as the aspects of a category which can be generalized to all category members.

2.1.1. The Essentialist Framework

2.1.1.1. Metaphysical vs. Psychological Essentialism

The notion of essences goes back to Aristotle who defined "the essence of a thing [as] what it is said to be *per se*" (350 BC/1924, 1029b13). Rather than material substances, he considered essences to be inherent principles which – by determining characteristic properties and development – have causal force and constitute the identity of a kind (Ellis, 2014). In this way, they exist independently of the subjective mind. The idea that things, or rather natural kinds, are defined and predetermined by intrinsic essences has been discussed intensely throughout the history of philosophy. For instance, empiricists like Locke and Hume remained sceptical about the knowability and, consequently, the existence of such essences (Oderberg, 2007). According to the empiricist view, knowledge about the world only derives from direct experience (Ellis, 2014). Since essences are indirectly observable by means of universal properties of kinds, they are unsuited for scientific reasoning. During the 1970s, essentialism was revived in modal logic and theory of word meaning: Saul Kripke defined the essential properties of an object on the basis of possible worlds. More precisely, properties are essential if they are "true of that object in any case where it would have existed" (Kripke, 1972/1994, p. 200). Meanwhile, Hilary Putnam (1975) illustrated a thought experiment of a planet called Twin Earth, a spitting image of the Earth as we know it. The main difference, however, is the chemical structure of water, that is XYZ. Apart from its microstructure, XYZ is indistinguishable from H_2O – it has the same taste, behaves the same with regard to its physical qualities and is likewise called "water" by the Twin Earthlings. Putnam (1975) draws an essentialist conclusion: Since the essential properties differ, the term "water" will never have the same extension on both planets. Interestingly enough, microstructural essentialism constituted the last stronghold of essentialist theorizing in science for a long time. Yet lately, it has been equally called into question (see Bursten, 2014).

Medin and Ortony (1989) introduced essentialism to the field of psychology. Their notion of psychological essentialism, however, entails a very different concept than its philosophical counterpart (Gelman, 2003; Gelman & Ware, 2012; Haslam & Abou-Abdallah, 2015; Medin & Ortony, 1989): While essentialism in philosophy makes ontological claims about the structure of the world by assuming the existence of intrinsic essences, the psychological account does not engage in metaphysical considerations. Psychological essentialism describes people's implicit ontological intuitions and how they mentally represent the world, rather than making statements about objective reality. It thus refers to a belief system – a conceptual bias – which contains two main components (Gelman & Ware, 2012; Gelman, Ware, & Kleinberg, 2010):

(1) The natural kind assumption: People believe "that certain categories are natural kinds (i.e. real, discovered, and rooted in nature)" (Gelman & Ware, 2012, p. 3). As opposed to artificial categories, they are not manufactured or invented by humans, but occur independently in nature.
(2) The essence assumption: People assume for such categories "that there is some unobservable property (part, substance, or quality) — the essence — that causes things to be the way they are and gives rise to the observable similarities shared by members of a category" (Gelman & Ware, 2012, p. 3).

Both components come with various implications: First of all, as unobservable properties themselves, essences serve as "placeholder" notions (Medin & Ortony, 1989, p. 184). That is to say, people may believe in the existence of essences and have theories about what they are without knowing their actual identity. Essences are not definitions, they neither contribute to the meaning of a concept, nor do they directly identify members of a category (Gelman, 2003, 2004; Gelman & Ware, 2012). Nonetheless, essences fill in for yet unknown properties and, thus, bear great inductive potential: People readily generalize properties from one category member to other members of the same category even if their perceptual dissimilarity indicates otherwise. For instance, Yamauchi and Markman (2000) found in four experiments that people were more likely to predict missing features of an imaginary bug based on category information, even in cases where the test item displayed stronger similarity with an opposing kind of bug. The same effect has been found in children as early as the age of one and two: Graham, Kilbreath, and Welder (2004) showed that 13-month-old infants expected non-obvious properties (e.g. a rattle sound) in objects of the same category. Accordingly, they equally imitated eliciting actions (e.g. shaking) on similar objects and objects of different shape when they were introduced by the same category label. The imitation occurred significantly less often for dissimilar objects when no labels were used or when the objects were presented with different labels (for similar findings with 24-month-olds see Jaswal & Markman, 2007). These results generally suggest that categories are represented twofold in the conceptual system: once on a "level of observable reality" and once on a "level of explanation and cause" (Gelman, 2004, p. 407). They also indicate that the levels are organized hierarchically: The first is suspended by the later, when some underlying causal part or substance is suspected. In other words, people attribute variation between category members only to a superficial level, whereas underlying causes pertain to the entire category.

The importance of cause in essentialist reasoning also becomes evident in a different way: Firstly, non-obvious properties are more likely to be generalized to the entire category when they are linked to another property by causal explanation (Rehder, 2006). Secondly, category membership ratings significally decline when a causal feature is missing, but not in case of a missing effect feature (Ahn, Kim, Lassaline, & Dennis, 2000). Overall, this suggests that causes play a greater role in conceptual representation than effects or associations (Gelman, 2003). Thirdly, people are more inclined to infer stable and internal properties as compared to transient or accidental ones (Gelman, 2003, 2004; Heit, 2000). Since essences are considered to be stable, people assume that their observable effects are just as perpetual.

The last assumption points at another aspect of the essentialist belief system, namely the "nativist bias" (Gelman, 2004, p. 406). Following from the natural kind assumption, people believe that biological essences are inheritable and passed on from one generation to the other. Accordingly, category membership as well as characteristic properties are determined by birth and mediated by essences (Gelman, 2003, 2004; Gelman & Ware, 2012). This belief is frequently found in people's naïve theories of biology, even though contradicted by the Darwinist theory of evolution (e.g. Shtulman, 2006), genetics (Dar-Nimrod & Heine, 2011) or medical and psychiatric research (Ahn, Flanagan, Marsh, & Sanislow, 2006). The nativist assumption emerges relatively early in childhood by the age of three (Gelman & Wellman, 1991). In so-called switched-at-birth-tasks, a story of a newborn animal which is raised by parents of a different species is used to assess children's conception of nature and nurture. When reasoning about the animal's properties or traits, children tend to ignore environmental influences through adoption and focus on heredity (Astuti, Solomon, Carey, Ingold, & Miller, 2004; Sousa, Atran, & Medin, 2002; Taylor, Rhodes, & Gelman, 2009; Waxman et al., 2007), whereby physical and known characteristics are affected to a greater extent than behavioural and unknown properties, respectively. Similar results have been reported for adults (Astuti et al., 2004; Sousa et al., 2002; Taylor et al., 2009; Waxman et al., 2007; for gender see Eidson & Coley, 2014). Interestingly, these findings were replicated with children and adults from various cultural backgrounds, confirming the basic premise that essentialism constitutes a general constraint on conceptual representations.

Essences are furthermore considered to be immutable (Gelman, 2003, 2004; Gelman & Ware, 2012). For instance, category identity does not seem to be influenced by severe changes in a category member's appearance. Rosengren, Gelman, Kalish, and Mccormick (1991) found that, at the age of three, children display a basic understanding of growth. At the age of five, they furthermore accept that biological kinds can undergo drastic transformations such as

metamorphosis, while still maintaining their underlying identity. More importantly, children's category judgements are not affected by the removal of superficial properties, whereas they are affected when internal ones are removed (Gelman & Wellman, 1991; for cross-cultural differences see Sousa et al., 2002; Waxman et al., 2007). Again, the experimental evidence indicates two levels of category representation: Variation within as well as between category members is ascribed to a superficial level. Underlying properties, on the other hand, are represented on a deeper level which is assumed to be immune to change or transformation.

Psychological essentialism also gives rise to sharper category boundaries (Gelman, 2003; Gelman & Ware, 2012; Newman & Knobe, 2019). Various studies investigated this effect in adults (Diesendruck & Gelman, 1999; Estes, 2003; Kalish, 2002): By comparing category membership judgements and typicality ratings, results showed that people were generally more inclined to pass absolute category judgements on atypical members of natural kinds. Graded membership, in turn, was more often ascribed to artificial concepts, or rather, atypical artefacts were less likely to be judged as full members of a corresponding category. As for children, Rhodes and Gelman (2009) observed that five-year-olds identified atypical category members in equal measure for both animal and artefact categories such that typicality ratings did not vary across domains. However, when asked about category membership, children did not view peacocks as "sort of" birds but as absolute category members, while they were more prone to agree with belts and headbands as only partially belonging to the category "clothing". Overall, these results suggest that natural kind categories tend to be conceived as dichotomous instead of as a continuum with grey areas at both edges.

2.1.1.2. Privileged Levels and Domain Specificity

It should be pointed out that not all category representations seem to be affected by psychological essentialism (Gelman, 2003). That is to say, people do not infer properties from dolphins to armadillos based on the fact that both are mammals. At the same time, it is rather difficult to imagine what the underlying property of brown spotted tabbies would be and, over and above, how it would differ from the essential properties of other types of brown cats. In accordance with artificial objects, natural kinds appear to have a "basic level of abstraction at which the concrete objects of the world are most naturally divided into categories" (Rosch & Mervis, 1975, p. 586). Yet, what exactly corresponds to this basic level in living kinds? Coley, Medin, and Atran (1997) addressed this question more thoroughly by testing for upward entailment between biological category levels. Participants significantly declined to infer

properties to levels above the generic species level (e.g. from generic species such as vultures to life forms such as birds). The results led to the conclusion that only generic species closely correspond to people's essentialist conception of natural kinds and, in consequence, allow for category-based inferences (see also Atran, Estin, Coley, & Medin, 1997).

Apart from taxonomy, essentialism is subject to additional limitations. Scott Atran argued that essentialist beliefs "issue from a specific cognitive structure, which may be a faculty of the human mind that is innately and uniquely attuned to perceiving and conceptually organizing living kinds" (Atran, 1998, p. 548). In this sense, psychological essentialism is the product of a domain-specific information processing module which, due to evolutionary pressure, is biological at core. Essentialist reasoning outside the realm of biology either signifies an expansion of the folk-biological module into similarly structured domains (Gil-White, 2001), or employs a discrete modular device in the cognitive architecture (e.g. for folk sociology see Hirschfeld, 1996, 2007; Sperber & Hirschfeld, 2004). By contrast, Susan Carey claimed that essentialism derives from linguistic input, paired with children's domain-general logical capacities. In concrete terms, children are equipped with the default assumption that count nouns refer to substance sortals (Carey, 1996; Kelemen & Carey, 2007). Since substance sortals imply unchanging identity, children's attention is directed towards the origins of a kind. This focus, in turn, acts as a starting point to the conception of inherent essences. According to Carey (1996), all sortal categories are essentialized in terms of their origins at an early stage of development. It is only after conceptual change that origins receive different emphasis in different domains. In the case of biological kinds, origins become associated with essential properties. Artefacts, on the other hand, are combined with a causal explanatory structure, the so-called "design stance", which implies that objects originate from an intended function so that creator's intent causes their superficial properties (Kelemen & Carey, 2007). From Carey's perspective, essentialism represents a domain-general bias which is simply instantiated in different ways in different domains (Gelman, 2003).

Empirical evidence on this matter is indirect and rather inconclusive: Indeed, a large amount of the empirical work attests essentialist beliefs early in childhood with reference to biological categories (Gelman & Wellman, 1991; Rhodes & Gelman, 2009; Waxman et al., 2007; for extensive review see Gelmann, 2003). Biological essentialism, moreover, seems to be cross-culturally coherent (Astuti et al., 2004; Atran, Medin, & Ross, 2002; Sousa et al., 2002). At the same time, effects usually attributed to psychological essentialism have not been replicated for artefacts: Heyman and Gelman (2000) found that three- to four-year-old children drew inferences between sketched humans, but not when the drawings were introduced as dolls. The

same difference was reported regarding children's expectations about category identity and superficial change and regarding sharp category boundaries: In their study, Rosengren et al. (1991) showed that children as early as three readily apply the concept of growth to natural kinds by choosing a bigger over a similar size picture as the matured version of an animal. For older versions of artefacts, they merely performed at chance level. At the same time, Gelman and colleagues (Diesendruck & Gelman, 1999; Rhodes & Gelman, 2009) found that both adults and children rated the majority of atypical animal items (e.g. penguin, starfish) as absolute members of their corresponding category (e.g. bird, fish), while granting only partial membership to atypical artefacts (e.g. earmuff as clothing). However, a clear-cut distinction between domains – or rather, the presumption that essentialism only applies to biological kinds, not artefacts (see e.g. Gelman, 2003) – has lately come under scrutiny: Essentialist beliefs have been identified in adults who reason about non-obvious properties in objects of art (Rabb, Brownell, & Winner, 2018) or celebrity possessions (Newman, Diesendruck, & Bloom, 2011) as well as in children when artificial objects display distinctive history (Pesowski & Friedman, 2019). It should be noted though that essentialism only extends to *individual* objects, rather than to *entire* artificial categories such that only particular artefacts are considered to entail essence-like qualities (Gelman, 2013).

Finally, essentialism is implicated in the social domain (Haslam, Rothschild, & Ernst, 2000; Prentice & Miller, 2007). As first suggested by Gordon Allport (1954, p. 174), "[t]here is an inherent 'Jewishness' in every Jew. The 'soul of the Oriental', 'Negro blood,' [...] 'the passionate Latin'– all represent a belief in essence. A mysterious mana (for good or ill) resides in a group, all of its members partaking thereof". Drawing on this idea, psychological essentialism is nowadays viewed as fertile ground for prejudice and stereotyping, shifting both phenomena from deviant to ordinary manifestations in social cognition (Fiske & Russell, 2010). Most studies focus on particular social categories: Essentialist reasoning has been identified, amongst others, in respect of race (Hirschfeld, 1996), ethnicity (Gil-White, 2001), gender (Mahalingam, 2003), sexual orientation (Duncan & Kemmelmeier, 2012; Haslam & Levy, 2006), and personality (Haslam, Bastian, & Bissett, 2004; Meyer, Leslie, Gelman, & Stilwell, 2013). Furthermore, correlations were found between essentialist beliefs and people's willingness to hold stereotypical or prejudiced assumptions (e.g. Bastian & Haslam, 2006; Diesendruck & Menahem, 2015; Mandalaywala, Amodio, & Rhodes, 2018), however, with significant differences amongst categories (especially regarding anti-gay attitudes, see Duncan & Kemmelmeier, 2012; Haslam & Levy, 2006; Haslam, Rothschild, & Ernst, 2002). Along with biological categories, social categories constitute the second major domain in the grip of

psychological essentialism. Yet, in spite of numerous communalities between both domains, their developmental trajectories are quite different. Taylor et al. (2009), for instance, found that gender and animal concepts become more differentiated with increased age. In switched-at-birth tasks, children aged five predominantly predicted behavioural properties in both animal and gender condition based on the inherent category. Older participants, on the other hand, were less likely to infer gender-linked behaviour from category membership (but see Eidson & Coley, 2014). As regards flexibility of properties, all age groups were increasingly more open to environmental influence on behaviour associated with gender than on animal-linked behaviour (Taylor et al., 2009). Differences also emerge between particular social categories: When reasoning about stability of social categories, Kinzler and Dautel (2012) found that five-year-olds essentialized language more strongly than race. Reversed results were obtained with older children and African American children (see also Mandalaywala, Ranger-Murdock, Amodio, & Rhodes, 2018). Furthermore, children consider ethnicity as a much more inductively powerful category than adults (Diesendruck & HaLevi, 2006). In this respect, Diesendruck, Goldfein-Elbaz, Rhodes, Gelman, and Neumark (2013) found that cultural input contributes to essentialist beliefs: While gender was consistently essentialized both in their North American and Israeli studies, race and ethnicity show different developmental patterns depending on "their particular cultural significance" (Diesendruck et al., 2013, p. 1916). Similar findings were obtained as regards religious categories (Chalik, Leslie, & Rhodes, 2017): Children at the age of five held stronger essentialist beliefs about religious groups than older children and adults. By adulthood, however, religious Jewish participants more readily essentialized religions – a pattern which might result from exposure to a theology with explicit essentialist conceptions. Overall, findings indicate that psychological essentialism becomes effective to different extents at different dates: Biological essentialism is documented in many ways as early as at preschool age (Gelman, 2003). Gender-related and religious essentialism emerges at around the same time, yet diminishes with age, whereas racial essentialism grows stronger over the course of development given particular cultural settings. Moreover, pre-schoolers are more rigid in their perception of language and ethnicity as a social category as compared to older children and adults.

With regard to domain specificity, the developmental evidence is compatible with the conceptions that essentialism is either a domain-general bias instantiated separately in different domains (Carey, 1996; Kelemen & Carey, 2007), or derives from separate modules (Hirschfeld, 1996, 2007; Sperber & Hirschfeld, 2004). Withal, the extension of a domain-specific mechanism from the biological to the social domain seems rather unlikely as, in this case,

increasing convergence between domains would be expected (Gelman, 2003; Taylor et al., 2009). Findings also suggest that there are many different components to essentialism (e.g. inherence, inductive potential, stability, causal determination), which do not always appear jointly together. For this reason, Haslam et al. (2000) distinguish between two essentialist dimensions: Categories are either essentialized as natural kinds (referring to their naturalness, discreteness, immutability, historical stability and their necessary features), or in terms of entitativity, "a function of Gestalt principles of similarity, proximity and common fate" (Haslam et al., 2000, p. 115), which highlights exclusivity, uniformity, informativeness and an underlying reality.[1] In this sense, essentialist beliefs might not only result from one but from two domain-general perceptual biases. Entitative categories are not necessarily considered to be natural kinds and, vice versa, natural kinds do not have to be entitative. Nonetheless, variation in essentialist beliefs might arise from this very distinction (Prentice & Miller, 2007): For biological kinds, perceptions of a category's naturalness and entitativity are closely related, whereas they are fairly different in the case of social categories (e.g. liberals are rated low on naturalness but highly with regard to entitativity, whereas men are perceived highly natural but not entitative; see Haslam et al., 2000). It is interesting for the purpose of this thesis that religious categories align with biological kinds: In their study, Toosi and Ambady (2011) found that religious groups were rated highly both on naturalness and entitativity, indicating that "religion is unique as a social identity" (Toosi & Ambady, 2011, p. 24). Gelman (2003) takes the following position on the matter of domain-specificity: In accordance with Carey (1996), she claims that psychological essentialism derives from a domain-general predisposition with different instantiations in different domains. However, rather than a single default assumption that shifts focus on the origins of a kind, multiple root capacities are assumed to give rise to different essentialist components (Gelman, 2003; Rhodes & Mandalaywala, 2017). In that sense, psychological essentialism can be seen as a by-product of several cognitive biases none of which is uniquely attributed to it. For instance, causal determinism describes people's general assumption that events and properties are caused. In the absence of external causation (e.g. for movement or behaviour in natural kinds), attention is drawn towards some sort of internal cause (i.e., an essence). Gelman (2003) also suggests that causal determinism is a precondition to Haslam's naturalness dimension, whereas entitativity derives from a default assumption about property clusters: When people observe that members of a category share multiple properties, they readily assume the existence of

[1] Note that Haslam et al.'s (2000) two dimensions coincide with Gelman's natural kind and essence assumption (Gelman, 2003; Gelman & Ware, 2012), yet naturalness and entitativity are much stronger conceptions.

further common features. In other words, property clusters make up a large part of a category's inductive potential. Given that they are richer for biological kinds than for simple artefacts, clustered properties further explain differences between the two domains (see chapter 2.1.3.3., p. 31 for a similar point). According to Gelman (2003), the belief that essences are immutable stems from people's general ability to track individual identity. More precisely, people recognize that individuals and members of a kind remain the same in spite of superficial, temporal or spatial transformation. The link between this general ability and essentialism becomes evident in another way: People tend to essentialize only those artificial objects for which they can track identity through time and space as well (e.g. artwork, objects with personal value, celebrity possessions). Boundary intensification, on the other hand, follows from the general disposition to distinguish between appearance and reality. As the appearance/reality distinction is prerequisite to abstracting from superficial variation and thinking about underlying essences, atypical members can be added to a concept on the basis of such underlying properties. Accepting atypical category members also follows from labels and the fact that children defer to linguistic input given by experts (for further discussion regarding the role of language see chapter 3).

In summary, psychological essentialism describes people's implicit belief system about some underlying, informative property that is present in every member of a biological or social category (Haslam & Abou-Abdallah, 2015). Essentialist beliefs emerge early in development and come in various forms and manifestations. Against this background, essentialism is strictly speaking not domain-specific and modular, nor a fixed set of beliefs that applies in an all-or-nothing manner (Gelman, 2003; Haslam et al., 2000; Prentice & Miller, 2007). Essentialism presumably derives from convergence of several predispositions and "develops to the extent that these underlying capacities develop" (Gelman, 2003, p. 321). Moreover, it can be thought of as a metaconceptual stance which – as a placeholder notion – "does not specify particular conceptual content, [but] constrains and […] guides further knowledge acquisition" (Gelman, 2003, p. 267).

2.1.2. The Inherence Heuristic

2.1.2.1. An Outline of the Mechanism

Category representation might further be constrained by a single, more general cognitive mechanism. Drawing on Kahneman's and Tversky's seminal work on heuristics and biases (see e.g. Kahneman, 2011; Tversky & Kahneman, 1974), Cimpian and Salomon (2014) propose a broad mental short-cut which comes into effect when people intuitively reason about patterns in the world. By virtue of the so-called **Inherence Heuristic**, people identify components of an observed regularity and hold their inherent features responsible for its emergence. To state an example, people commonly assume that little girls like pink, because there is something inherently delicate and feminine about this colour. Likewise, orange juice is allegedly a breakfast beverage due to its morningly taste. External factors such as historic or customary development are largely dismissed in intuitive heuristic explanations.

In general, intuitive heuristics are defined as fast and implicit reasoning processes implemented on a low level of the cognitive architecture (usually dubbed type/system 1 processes, see Evans, 2003; Evans & Stanovich, 2013). Rather than considering all facts, people quickly generate plausible explanations based on the most prominent pieces of information (Cimpian, 2015). Heuristic outputs are default responses, they are derived in a quick and dirty fashion and rely on a limited number of rules and principles (Tversky & Kahneman, 1974). Moreover, they serve as input to high-order evaluative processes (Cimpian & Salomon, 2014). So-called type 2 processes require a high amount of cognitive effort, they are slower, more flexible and involve conscious experience and intent (Evans, 2003; Evans & Stanovich, 2013). Evaluative processes may override intuitive heuristic reasoning, but due to limited cognitive capacities, replacement predominantly takes place when there is sound reason to doubt the initial response (Cimpian, 2015). At large, humans are cognitive misers such that a great amount of judgement and decision-making builds upon highly economic intuitive heuristics (Kahneman, 2011).

The Inherence Heuristic comprises of three stages of processing (Cimpian & Salomon, 2014): As a first step, people set up a general task or question. More precisely, heuristic reasoning starts off by people targeting a pattern which they seek to explain. While at first glance this might seem trivial, it comes in fact with an important theoretical implication: People have a strong urge to derive theories from perceptual input. According to the theory theory framework (see e.g. Gopnik, 1998; Gopnik & Meltzoff, 1997), humans are equipped with a "special and distinctive set of […] representational capacities" (Gopnik, 1998, p. 102) – a theory formation system – which organizes perceptual representations in terms of their abstract causal structure. As deriving underlying structure demands a large amount of cognitive effort and, moreover,

benefits (such as predicting future behaviour) are often indirect or become obvious much later, the theory formation system is accompanied by a motivational component. An *explanatory drive* acts, similar to physiological motives, as an inherent motivation to construct and adjust theories about the world. By positing people's urge to explain patterns as a basic precondition to heuristic reasoning, Cimpian and Salomon (2014) follow a theory-based account of knowledge representation.

Once the search for an explanation is initiated by the explanatory drive, information is activated in a second step at the so-called *mental shot gun* stage (see Kahneman, 2011). This proceeds along two lines: Either an already existing, consolidated explanation is retrieved from long-term memory such that the search for explanation terminates right away, or fragments of information are retrieved in order to construct an explanation online. In the latter case, two principles are obeyed (Cimpian, 2015): Firstly, as processes at the retrieval stage are fast, shallow and highly efficient, they tend to prioritize those pieces of information which ensure least cognitive effort. This means, on the one hand, that only highly salient information is activated, usually referring to the pattern's focal components (e.g. <girls, pink> or <orange juice, breakfast>). On the other hand, not all information about components is equally easy to access. The Inherence Heuristic proposal states that information about stable and inherent features is easiest to access and, consequently, retrieved with high probability from semantic memory (Cimpian, 2015). Secondly, only information that is relevant for generating explanations is activated. For instance, it might not seem relevant to the preference of pink that girls inherently have two arms and legs. By the same means, the fact that orange juice is generally drunk from a glass would seem irrelevant to explaining why juice is consumed at breakfast. Essentially, the mental shot gun stage includes an automatic computation of relevance.

The third and final stage of processing is termed the *storyteller* stage (see Kahneman, 2011) where the output finally comes into consciousness an intuitive explanation. In this last step, the information retrieved during mental shot gun percolates up to working memory in order to construct a coherent explanatory narrative (Cimpian, 2015; Cimpian & Salomon, 2014). For this purpose, different explanatory relations are in use. Information is either assembled in causal relations (e.g. biological or psychological causation), or in a normative, value-based manner (e.g. appropriateness, optimality) which applies to both examples mentioned above: Pink is deemed to be an inherently delicate colour and, thus, *suitable* for delicate little girls. Likewise, something about orange juice makes it *optimal* as an early morning beverage. Which particular explanatory relation is selected during the storyteller process depends on how plausible the

output is with regard to context and prior knowledge (Cimpian & Salomon, 2014). It goes without saying that the storyteller stage is biased from the very beginning. Since the "retrieval stage tends to oversample inherent facts" (Cimpian, 2015, p. 7), these pieces of information are also overrepresented in working memory. As a result, explanations most likely turn out to be inherence-based as well. Only if external facts are highly accessible at the mental shot gun stage – even more accessible than inherence information – intuitive explanations may appeal to external factors. According to Cimpian and Salomon (2014), the predominance of inherent features becomes even more obvious in cases where they remain vague or undiscovered. Even though it is not straightforward which intrinsic properties of either orange juice or breakfast enter a normative relation, they ultimately explain the observed pattern in people's intuitive judgements.

2.1.2.2. Explaining Various Phenomena

The Inherence Heuristic is introduced as a domain-general cognitive mechanism that causes people to explain observed patterns with reference to inherent properties. On these grounds, it should be compatible with people's intuitive explanations in a variety of domains (Cimpian & Salomon, 2014). Indeed, first evidence originates from research in social psychology. So-called attribution theories are concerned with the question how people make causal and dispositional inferences about their own behaviour or the behaviour of others (e.g. Kelley, 1973). Studies showed that causation of an interpersonal event (e.g. Diane declines Larry's invitation, see example in Orvis et al., 1975, p. 614) is more often attributed to the intrinsic characteristics of participants, when additional information suggests a highly consistent pattern (e.g. Diane always declines Larry's invitation). However, people tend to hold the specific circumstances accountable when inconsistency with previous interactions is highlighted (e.g. Diane never declined Larry's invitation before). As predicted by the Inherence Heuristic, the presence of behavioural patterns seems to promote explanations with reference to inherent features (see also Hilton, Smith, & Kim, 1995; McArthur, 1972 for similar results).

The Inherence Heuristic can also account for another well-documented phenomenon in social psychology, namely for system justification tendencies. In order to maintain their belief in a righteous world, people tend to judge societal or political structures as inevitable, legitimate and fair (Cimpian & Salomon, 2014). They furthermore privilege explanations referring to the characteristics of people who occupy a position in the socio-political structure. More precisely, instead of blaming external systematic factors, causation of societal patterns is associated with

the alleged (inherent or developmentally acquired) flaws of an afflicted group (Napier, Mandisodza, Andersen, & Jost, 2006). Likewise, inherent virtues are believed to result in a privileged position in society. System justification based on inherent properties has been documented with respect to poverty (Cozzarelli, Wilkinson, & Tagler, 2001), gender inequality (Jost & Kay, 2005; Kray, Howland, Russell, & Jackman, 2017) or shortcomings in the aftermath of natural catastrophes (Napier et al., 2006).

Evidence for inherence-based explanations can also be drawn from feedback and motivation research. For instance, children seem to cope differently with subsequent failure depending on how feedback is phrased (Cimpian, Arce, Markman, & Dweck, 2007; Kamins & Dweck, 1999). When feedback given previously reflected general praise, children more often displayed helpless behaviour after a setback (negative affect, resignation, negative self-evaluation). Conversely, mastery-orientated reactions (neutral affect, persistence as regards the task, positive self-assessment) were more likely when a specific performance was highlighted in previous feedback. These differences can be interpreted as follows: Since general praise implies a pattern of success, children conclude that they have inherent qualities such as intelligence or talent which are stable, incontrollable and elicit achievement (entity theory of intelligence). Failure thus leads to a sense of helplessness, or even to questioning one's self-concept. At the same time, specific praise does not trigger the search for inherence-based explanations. Failure can be attributed to more fluid factors such as effort, strategy or concentration (incremental theory of intelligence, see Kamins & Dweck, 1999).

Another phenomenon can be reconciled with the Inherence Heuristic. "Nominal realism" refers to the belief that words reflect the intrinsic features of their denoted objects. In this respect, people seek to explain the patterns of word-referent-pairings. Rather than arbitrary and conventional by nature, words are considered to be objective cues, or even part of the objects themselves (Gelman, 2003). Nominal fit beliefs are predominantly found in young children with regard to category labels (e.g. Benelli, 1988; Brook, 1970; but see Rosenblum & Pinker, 1983), yet also to some extent in adults (Sutherland & Cimpian, 2015). Brook (1970), for instance, interviewed children ranging from 6 to 10 in order to access two aspects of nominal realism, namely the natural origins of names and their fixed nature. With progressing age, children started to differentiate between referents and labels and recognized their underlying arbitrary relation. At the age of six, however, children were almost exclusively convinced that names contain aspects of meaning such that "the sun has to be called the sun because it is a fiery planet and it shines" (Brook, 1970, p. 167). Nominal realism also pertains to idiomatic expressions (Cimpian & Salomon, 2014). Idioms are often accompanied by the intuition that

their meaning arises from the meaning of their components. In truth, this intuition has been shown to be merely a function of use. Keysar and Bly (1995) found that unknown idioms were perceived as transparent after they got familiarized in a learning phase. This was the case both when participants learned the original as well as the reversed meaning of an idiom. Overall, nominal fit intuitions might result from a general heuristic process which explains a linguistic pattern (broadly speaking word use) by means of its inherent semantic features. Sutherland & Cimpian (2015) addressed this possibility more thoroughly: In their study, they found that nominal fit beliefs were predicted by the general tendency to rely on inherence-based explanations. The results were obtained for children as well as for adults, even when a variety of alternative factors (e.g. intelligence, counterfactual thinking) were controlled. Furthermore, when manipulating people's tendency towards inherence-based explanations, "a downstream effect on their tendency to see words as fitting with their referents" (Sutherland & Cimpian, 2015, p. 237) emerged. In summary, the Inherence Heuristic can account for various phenomena which hitherto have been classified as unrelated. By postulating one single powerful underlying mechanism, the Inherence Heuristics proposal does not only offer a new perspective on those phenomena, but also casts light on the more general question how people perceive and represent patterns in the world.

2.1.2.3. The Inherence Heuristic vs. Psychological Essentialism

Prima facie, it is not entirely obvious what the relation between the Inherence Heuristic Hypothesis and the aforementioned essentialist framework would be. Both approaches highlight people's assumption of inherent properties as well as their inductive potential in reasoning processes. The main difference might therefore lie in the fact that the first spells out a precise reasoning mechanism, whereas the latter focuses on contents of its outcome (Cimpian, 2015). However, psychological essentialism and the Inherence Heuristic are far from being two sides of the same coin. For one thing, the Inherence Heuristic is a mechanism that accounts for far more phenomena than those which fall into the scope of essentialist reasoning (see 2.1.2.2.). At the same time, several other cognitive tendencies are suggested to endow essentialist reasoning (see 2.1.1.2.). Differences also become evident by the comparison of essential and inherent features: Essences are defined as unobservable, internal, underlying substances which do not contribute to the meaning of a concept and often remain beyond our grasp (Gelman & Ware, 2012; Medin & Ortony, 1989). While this may also be valid for inherent features, it does not capture their full range. Inherent features can equally be obvious, concrete and superficial

as long as stable characteristics of an entity under consideration are portrayed (Cimpian & Salomon, 2014). Most importantly, the Inherence Heuristic "[does] not necessarily pinpoint a single inherent feature as the source of the multiple regularities observed within a kind" (Cimpian & Salomon, 2014, p. 474). As opposed to psychological essentialism which attributes a single causal essence to each natural kind, different properties can be utilized for different inherence-based explanations.

To ask ourselves how essentialism and the Inherence Heuristic relate to each other actually means to ask ourselves how essentialist beliefs come into existence. Cimpian and Salomon (2014, p. 474) argue that "the early output of the inherence heuristic is gradually elaborated into a full-blown essentialist stance" – establishing the Inherence Heuristic as a direct precursor to psychological essentialism. Accordingly, how do inherence-based explanations convert into essentialist beliefs? As previously mentioned, several cognitive tendencies contribute to essentialist reasoning including causal determinism and induction from property clusters (Gelman, 2003; Rhodes & Mandalaywala, 2017). In a similar vein, Cimpian and Salomon (2014) suggest several cognitive biases which refine the rather inchoate Inherence Heuristic over the course of development. For instance, it has been shown that even eight-month-old infants adhere to some sort of "innards" principle that is the abstract expectation that self-propulsion and agency – both of which are properties of biological kinds – are caused by the insides of an object (Setoh, Wu, Baillargeon, & Gelman, 2013). During a familiarization trail, infants were introduced to two novel objects that both either moved and/or responded independently or remained motionless and mute. The two objects were then rotated to unravel that one of them was hollow, while the other one was sealed. Measured looking times were significantly longer for the hollow objects in the self-propelled, agentive condition. This suggests that infants detected a violation when objects which they identified as living kinds proved to lack insides (for similar findings in 14-month-old children see Newman, Herrmann, Wynn, & Keil, 2008). Additionally, Gottfried and Gelman (2005) showed that children as early as age four display a sophisticated understanding of the role of internal mechanisms. Rather than the insides themselves, they endorse an abstract internal energy as the causal force to an animal's movement (but not to static behaviour or growth). It goes without saying that essentialism reaches beyond these conceptions of insides or abstract internal energy (Cimpian & Salomon, 2014): Essences are not general causal forces, but distinct for each natural kind. Moreover, they do not only elicit self-propelled or agentive behaviour, but in fact a range of properties including non-obvious or static characteristics. The two causal biases, somehow conflated into Gelman's notion of causal determinism, can therefore be viewed as guiding

principles “from an inchoate sense that the patterns encountered within a living kind are due to some combination of inherent factors to more concrete ideas about these factors” (Cimpian & Salomon, 2014, p. 475). This means that, in certain domains, retrieval at the mental shot gun stage is eventually narrowed down to a specific type of inherent features, namely to internal and non-obvious essences.

In their study, Salomon and Cimpian (2014) tested two predictions deriving from this proposed relationship. First of all, as inherence-based thinking is the precursor to later emerging essentialism, individual differences in endorsement of essentialist beliefs (see e.g. Bastian & Haslam, 2006) should be predicted by people’s general predisposition towards inherence-based explanations. Secondly, the Inherence Heuristic might still be the foundational mechanism of essentialism in adulthood. Manipulating the extent of inherence-based intuitions should hence show effect in the endorsement of essentialist beliefs. The study showed that people who were more inclined towards inherence-based explanations of patterns relating to non-essentialized domains (e.g. traffic lights, toothpaste, weekdays) held significantly stronger essentialist beliefs about social categories. However, after statements were presented that highlighted a pattern’s external explanations (e.g. historical development, conventionality), people were less likely to agree with inherence-based explanations. A mediation analysis further revealed that decreased reliance on inherence-based explanations led to significantly reduced essentialist beliefs. In sum, these results indicate a causal relation between psychological essentialism and the Inherence Heuristic’s broader cognitive mechanism. Along with other well-known biases (e.g. representativeness bias, anchoring bias), the Inherence Heuristic might therefore represent a fundamental part of human cognition, giving rise to various psychological phenomena, in particular to causal and dispositional attribution, system justification tendencies, nominal realism as well as to psychological essentialism.

2.1.3. The Aspect Hypothesis

2.1.3.1. Principled and Statistical Connections

So far, we have seen that representations of social and natural categories are constrained by a metacognitive essentialist stance and/or an even broader inherence-based heuristic. Both constraints operate in such a way that they put emphasis on certain properties in conceptual representations. As a result, essentialist or inherent properties are foregrounded in processes of reasoning and thinking. Prasada and Dillingham (2006, 2009) propose a slightly different constraint on the representation of nominal concepts. Rather than attributing primary importance to particular properties, they point out that differences in reasoning processes might arise from different relations between a category type, its members and the properties they possess. In other words, how properties are connected to a kind within a concept might be a key to understanding why some are generalized to all category members, while others are not. But let us start from the beginning: In theory-based approaches, it is generally assumed that our conceptual system makes use of different modes of explanations in order to build intuitive theories about our environment (Prasada, 2017; Prasada & Dillingham, 2009). Accordingly, intuitive theories are restrained by an essentialist-causal stance which explains properties and behaviour in terms of underlying substances (e.g. Gelman, 2003), a teleological stance according to which objects exist to serve a purpose (e.g. Kelemen, 1999), and an intentional mode of explanation which attributes goal-directedness to actions and behaviour (e.g. Gergely, Nádasdy, Csibra, & Bíró, 1995). Prasada and Dillingham (2009) add another stance to the list: A so-called formal mode of explanation exploits a part-whole principle such that properties of category members are explained by means of the kind of thing they are (Prasada & Dillingham, 2009). All modes of explanation come along with certain representational requirements: While the essentialist stance is solely sensitive to properties that represent effects and products, the intentional stance requires agentive properties. The teleological mode, on the other hand, only adheres to properties represented as means to an end. As regards the formal mode of explanation, the connections between a kind and the properties of its instantiations come to the fore (Prasada & Dillingham, 2009): Formal explanations are only licensed by so-called principled connections which relate a kind to those properties which are present in its instantiations, yet determined by the kind in question (henceforth called k-properties). To state an example, airplanes are known to have wings by virtue of being airplanes and dogs are four-legged because they are dogs. By this means, having wings and four legs constitute k-properties to planes and dogs, respectively. Factual or statistical connections, on the other hand, establish a connection between kinds and properties that are not determined by the kind itself, but rather

by the frequency in its tokens (hereafter called t-properties, see Prasada & Dillingham, 2009). Firefighting vehicles, for instance, are known to be red, yet simply because many instantiations display this property. Instead of assuming an underlying part-whole principle, we only have statistical expectations about the prevalence of redness in fire engines (i.e. red is a t-property). It should be noted that any connection only pertains to a specific kind-property-pairing: Even though redness represents a t-property of fire engines, it certainly is a k-property of field poppy. In general, principled connections come along with some particularities: Firstly, they spark off the conceptual system's formal mode of explanation such that k-properties in instantiations of a kind are **formally explained** on the basis of the kind itself. In an experiment, Prasada and Dillingham (2006) observed that, when people evaluated explanations for why something has a given property, formal explanations were rated significantly higher for k-properties than for t-properties. Prasada, Khemlani, Leslie, and Glucksberg (2013) moreover showed that formal explanations are only suitable in relation to properties with principled connections as opposed to those with causal connections to a kind (e.g. striking generics such as "ticks carry Lyme disease", see Prasada et al., 2013, p. 407).[2] Secondly, principled along with statistical connections create **statistical expectations** such that properties are assumed to be present in every member of a given category type. Herein also lies the difference to *logically necessary* and *contingent* connections (Prasada & Dillingham, 2006): While logical necessity demands that all instantiations of a kind possess a relevant property (e.g. all men are human, all bachelors are unmarried), principled connections allow instantiations that lack their k-properties (e.g. a not particularly sour lemon, a dog with only three legs). Contingency, on the other hand, is too weak a word for the expectation that k-properties are generally true to every member of a kind. Following from this, principled connections also have a **normative** component. In other words, they invoke the assumption that all members of a given category *should* have their k-properties, if not defective and, thus, exceptions to the rule (Prasada & Dillingham, 2006). Normative expectations only apply to principled, but not to statistical connections, as was shown in another experiment by Prasada and Dillingham (2006): In a truth-value judgement task, participants rated normative statements significantly higher when they referred to the prevalence of k-properties as compared to t-properties. For instance, the mere fact that fire engines are typically red does not necessarily lead to the assumption that there is something wrong with their yellow

[2] Causal connections imply that kind membership disposes a member to have a certain property, irrespective of whether the property is actually present. Further note that different types of connections are not mutually exclusive and often co-occur (Prasada et al., 2013). In fact, Prasada and Dillingham (2009) suggest that, while t-properties only show statistical or causal connections to a kind, k-properties establish both principled and statistical connections.

counterparts in Switzerland. However, we come to this very conclusion when encountering a three-legged dog.

2.1.3.2. The Representation of Aspects

The question remains which cognitive mechanisms are required in order to represent different types of connections. In the case of statistical connections, the answer is quite straightforward, namely a general "associative mechanism in which the strength of the association varies as a function of the prevalence of the property in instances of the kind" (Prasada, 2012). Accordingly, statistical connections are represented as simple associations which potentially exist between any two things. Principled connections, by contrast, only relate to kinds and some of their respective properties. Furthermore, they do not appear to be functions of frequency as is indicated by Prasada, Khemlani, Leslie, and Glucksberg (2013): When people rated assertions reflecting the characteristics of principled connections (i.e. normativity, formal explanations), minority characteristic generics (such as "ducks lay eggs", see Prasada et al., 2013, p. 407) patterned similar to unambiguous examples of principled connections (such as "airplanes have wings," see Prasada et al., 2013, p. 420). This is remarkable since minority characteristic generics merely apply to less than half of a kind's instantiations (i.e. only to mature, female ducks; for a closer examination of generics see chapter 3.1.). Against this backdrop, Prasada and Dillingham (2009, p. 407) propose the so-called **Aspect Hypothesis** according to which representing "a principled connection between a kind and a property requires representing the property as one aspect of being that kind of thing". By this means, principled connections bear comparison with part-whole relations such that having a property is part (or aspect) of belonging to an entire kind. For representation, Prasada and Dillingham (2009) propose the following mechanism in which a property P is co-indexed with an aspect representation *a1* of a particular kind K (Prasada & Dillingham, 2009, pp. 433f.):

(1) $K_i \rightarrow a1, a2, \ldots \; P_{a1}$

(2) $\text{CANARY}_i \rightarrow a1, a2, \ldots \; \text{YELLOW}_{a1}$

Aspect representations in (1) and (2) do not represent k-properties themselves, they rather resemble variables that are bound by a property's representation. In consequence, principled connections are not part-whole relations between the concept of a k-property (e.g. YELLOW) and the concept of a kind (e.g. CANARY), but "between the *things* that are thought about via the concepts YELLOW and CANARY" (Prasada & Dillingham, 2009, p. 433). In other words,

an aspect representation of a kind is never identical with the concept of a k-property, but they both refer to the same instances. This way, k-properties are not necessary conditions for being a kind of thing, nor do instances of a kind constitute an equivalence class. Instead, we can think of them as distinct and only matching with regard to their kind membership and their possession of certain k-properties (Prasada, 2010). This is of particular importance as one can be completely ignorant of the fact that canaries are typically yellow, yet still have a conceptual representation of them. In such cases, the aspect representation remains without an index and merely represents the general expectation that canaries have principled connections to certain k-properties (Prasada & Dillingham, 2009). By the same token, aspect representations can account for normative expectations regarding the prevalence of a specific k-property in all members of a kind. If a property represents an aspect of being a kind of thing, a principle of completeness or perfection can be exploited. On that note, the specific property should generally be present in all instantiations of a kind, for without this property an instantiation of a kind would be considered aberrant. Many kinds – by virtue of being the kinds of things that they are – are represented as physical objects which interact with the world. Causal interactions may sometimes obstruct a k-property from being present or from developing in the first place (e.g. albino canaries, dogs that have been mutilated, see Prasada, 2012). The mechanism in (1) provides this possibility and licenses instantiations lacking their k-properties for external reasons. Notwithstanding, it rules out that instances of a kind lack their k-properties for internal reasons, or rather by virtue of being the kind of thing they are (Prasada, 2010).

Although the Aspect Hypothesis does not help in determining the particular k-properties of a kind, it surely elucidates how formal explanations fall into place (Prasada, 2012, 2017; Prasada & Dillingham, 2009): Given that the existence of a whole (e.g. a kind) presupposes the existence of all of its parts, identifying something as an instantiation of a kind simultaneously accounts for the existence of its aspect properties. Investigating the explanatory value of a part-whole principle, Prasada and Dilllingham (2009) asked people once again to judge formal explanations for why something has a given property. This time, however, they included another type of formal explanation which included an explicit part-whole relation between a kind and the t- and k-properties in question (e.g. "Why is that red? Because it is a *red barn* [emphasis added]", Prasada & Dillingham, 2009, p. 421). People found those explanations equally explanatory for k- and t-properties, but accepted formal explanations solely for k-properties when the part-whole principle remained implicit ("Why is that red? Because it is a *barn* [emphasis added]", Prasada & Dillingham, 2009, p. 421). These results suggest that formal explanations which, by and large, are based on a part-whole principle are only supported by

properties identified as aspects of an entire kind (for further evidence see Prasada & Dillingham, 2009).

2.1.3.3. Domains of Aspect Representation

To reiterate, different modes of explanation require different property representations. They further make use of different explanatory principles: While the causal-essentialist mode of explanation applies a causal principle only to essential properties, the teleological stance explains purposive properties by use of a teleological principle. The intentional mode, furthermore, grounds the explanation of agentive properties on a principle of intentionality. Along these lines, the Aspect Hypothesis suggests a domain-general mode of explanation which does not focus on the content of properties per se but is sensitive to how they are represented within a concept. In that sense, only principled connections license part-whole explanations inasmuch as they represent certain properties as aspects of a kind (Prasada & Dillingham, 2009). Another clear line can be drawn between the formal mode and other modes of explanation, namely with respect to domain specificity (Prasada, 2017). For the essentialist stance, it has been established that primarily biological and social concepts are explained in terms of a causal essence (Gelman, 2003; Haslam et al., 2000; Prentice & Miller, 2007). In contrast, teleological explanations are confined to those kind concepts which are assumed to perform a function (Kelemen, 1999). Only artefacts and biological parts are therefore in the scope of the teleological mode of explanation. As regards the intentional stance, behavioural properties of a rational agent are to be explained (Gergely et al., 1995).[3] The formal mode of explanation, on the other hand, is quite broad in its application (Prasada, 2017; Prasada & Dillingham, 2006): Formal explanations can similarly relate to properties of animals, plants or artefacts as well as to the characteristics of, amongst others, imaginary beings (such as ghosts, unicorns), events (e.g. funerals) or abstract mathematical concepts (e.g. geometric forms). Formal explanations have moreover been found with reference to social concepts (see e.g. Prasada & Dillingham, 2006, 2009). In light of this, Prasada and Dillingham (2009, p. 432) argue that, "[b]ecause of its domain-generality […], it is plausible that the formal mode of explanation is available at the earliest stages of development and is part of the logic of kind

[3] In this context, domain-specificity seems to be tantamount to the representational requirements of properties. However, this is not quite the case: Artefacts, for example, can possess properties represented as effects and products as well (e.g. the ergonomic shape of baseball bats, see Kelemen & Carey, 2007), but we usually do not explain them in an essentialist manner. We rather attribute causality to the intention of a creator. Moreover, biological kinds can be argued to serve a function (e.g. guard dogs or pack animals). Still, teleological explanations for properties such as barking or heavy lifting are quite unreasonable so that we trace them back to causal essences instead.

representations." Their line of reasoning can be underpinned by experimental evidence from developmental psychology: When Taylor et al. (2009) asked 5- and ten-year-old children to explain their predictions of properties in a switched-at-birth task, even the youngest participants readily produced formal explanations. Haward, Wagner, Carey, and Prasada (2018) investigated the developmental question more thoroughly by comparing normative expectations and formal explanation judgements in children by the age of four to seven and in adults. Both children and adults endorsed that something was wrong with a member of a kind when pictures depicted a particular missing k-property (e.g. a three-legged cow). As for t-properties, participants answered in the negative irrespective of their age (e.g. a square instead of a round plate). Moreover, an explanation elicitation task showed that both children and adults spontaneously produced formal explanations, both with higher probability for k- than for t-properties. The study disclosed that people steadily differentiate between principled and statistical connections throughout their lifespan, supporting the claim that a formal explanatory stance is "part of the basic machinery humans use to think and speak about kinds" (Haward et al., 2018, p. 257).

In general, the experimental evidence suggests that the formal mode of explanation applies to various content domains (Haward et al., 2018; Prasada & Dillingham, 2006, 2009; Prasada et al., 2013). Be it as it may, domain differences do exist. As noted by Prasada and Dillingham (2006), the amount of principled connections apparently varies across domains so that natural kinds appear to possess more k-properties as compared to artefacts and social categories. Furthermore, which properties have principled connections to kinds seems to follow certain regularities (Prasada, 2017): While colours are often k-properties of natural kinds, they hardly ever are when it comes to artificial kinds. In view of this, Prasada (2017) recently claimed that domain differences arise from the fact that different types of kinds are connected to different types of k-properties. Put differently, different domains can be identified via the number and types of properties connected to their kinds via principled connections. Other than that, they have no "independent characterization" (Prasada, 2017, p. 1480). But which are the types of k-properties that make up different types of kinds? Concerning this matter, Prasada (2017) argues that only properties which contribute to material explanations (i.e. what is a thing constituted of), to efficient explanations (i.e. what produced a thing) or to teleological explanations (i.e. what are the intrinsic functional properties of a thing) may potentially be represented as k-properties and, thus, be formally explained with reference to a kind (for illustration, see Fig. 1 on the next page). The three property types can be realized in different ways: The first type includes either parts or stuff properties (e.g. substance, colour, texture), whereas properties of

the efficient type specify whether a thing was produced by a member of the same category (i.e. bred) or another kind (i.e. manufactured). The third type of properties includes monadic functions which relate a kind to itself (e.g. potential of growth), dyadic functions which establish a relation between two things (e.g. perceptual capacities, goal-directed movement) and intrinsic triadic functions such as reasoning and thinking (Prasada, 2017). Crucially, this is where domain differences become manifest: While both part and stuff properties are linked to animal kinds by means of principled connections, for artefacts, only parts are represented in this way. Accordingly, artificial kinds bear less principled connections to material properties than animal kinds. This way, artefacts are not determined to possess particular stuff properties by virtue of being of a kind (e.g. "#That is made of wood because it is a chair," see Prasada, 2017, p. 1482). They rather establish causal connections such that they are merely causally disposed to consist of a certain substance, colour or texture. In the same vein, efficient and teleological properties are represented differently across domains. Whereas, for instance, both animal and artefact kinds establish principled connections to divergent properties of the efficient type, natural kinds do not do so altogether. Consider, for example, geysers and mountains. One can be completely unaware of what produced them, yet still have a grasp of their concepts. Likewise, properties of the teleological type (e.g. growth, perception, goal-directedness) do not pertain to natural kinds. Those short examples illustrate that ontological distinctions between kinds might rely on different types of properties and the amount of principled connections they have to a kind (Prasada, 2017).

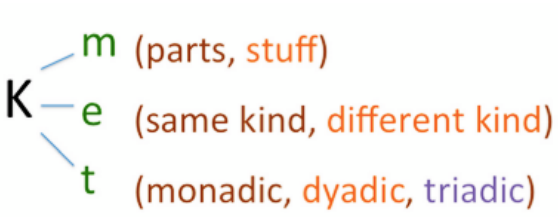

Figure 1. Types of k-properties and principled connections: (m) properties related to a material mode of explanation, (e) properties related to an efficient mode of explanation, and (t) properties related to a teleological mode of explanation (Prasada, 2017, p. 1481)

Ultimately, how can the Aspect Hypothesis be related to the two formerly mentioned internal approaches? First of all, Prasada and Dillingham (2006, 2009) emphasize that the formal mode of explanation is merely an additional way of making sense of the world such that it does not supersede an essentialist framework. However, their approach may contribute a crucial theoretical element to psychological essentialism by specifying which properties are in the scope of the essentialist stance: "The properties that are understood to be true of instances of a kind by virtue of their being the kinds of things they are (i.e., the properties for which we can provide formal explanations) are also candidates for the properties that are understood to be caused by an underlying essence in the relevant domain" (Prasada, 2017, p. 1485). According to this view, psychological essentialism presupposes that certain k-properties are represented

as aspects of a kind. With regards to the Inherence Heuristic, the Aspect Hypothesis may provide a similar contribution (Prasada, 2017): Since properties with principled connections are highly accessible within kind representations (see mechanism in 1), they stand good chances to be retrieved from semantic memory and assembled into inherence-based explanations (see chapter 2.1.2.1.). Formal explanations may even be retrieved as a whole at the *mental shot gun* stage, given that they are common in use and possibly consolidated in long-term memory. In this respect, Gelman, Cimpian, and Roberts (2018) propose that formal explanations are in fact placeholders for more specific and informative (inherence-based) explanations. In other words, the meaning of formal explanations goes beyond their literal content and incorporates a pragmatic prompt to "look further" for an inherent feature or cause – strictly speaking, for another aspect of a kind. In any case, the main difference between the Aspect Hypothesis and the Inherence Heuristic probably lies in their perspective on the explanatory nature of humans: While the Inherence Heuristic illustrates how intuitive explanations fall into place (i.e. a computational mechanism), the Aspect Hypothesis establishes how properties which are relevant to intuitive explanations are represented in the conceptual system (i.e. a representational mechanism).

2.2. Externalist Theories

Conceptual representations are constrained by broader mental structures in that "there is pressure for concepts to be consistent with whatever else [we] know" (Murphy, 2004, p. 60). So far, we have only reviewed naïve theories and cognitive mechanisms which identify the internal properties of a category as determinants of category membership. However, when it comes to social concepts, people also consider external factors to shape a category's identity: For instance, it has been claimed that people's **Belief in Social Determinism** strongly contributes to the representation of social categories (Rangel & Keller, 2011). Social determinism describes a naïve conception according to which the essential properties of a person are fundamentally influenced by external social factors. Social concepts might be constrained in yet another way: Vasilyeva, Gopnik, and Lombrozo (2018) argue that social categories are represented as occupants of positions within a social structure. As such, they are subject to constraints operating on structural positions, rather than on the actual entities. In other words, people expect particular properties in all members of a social category based on their position in the social structure (**Structural Thinking**).

2.2.1. The Belief in Genetic and Social Determinism

With respect to social categories, a particular *internalist* essentialist belief has been claimed to be the predominant or sole basis of essentialist reasoning, namely the **Belief in Genetic Determinism** (hereafter abreviated BGD, see e.g. Dar-Nimrod & Heine, 2011; Heine, Dar-Nimrod, Cheung, & Proulx, 2017; Keller, 2005; Suhay, 2017). According to BGD, a person's genome represents their underlying essence and, thus, determines individual characteristics and identity.[4] More importantly, people believe that genes define category membership and cause a social category to share distinctive properties and behaviours. Several studies have explored the role of BGD in social group perception: Keller (2005) found that endorsement of genetic deterministic beliefs strongly correlated with negative racial stereotyping, prejudice, system justification and conservative socio-political tendencies all of which are associated with essentialist reasoning in the social domain. The study moreover documented a causal effect: Participants were increasingly more dismissive of others and more biased towards their own in-group, when genetic information was priorly highlighted. Similar findings were obtained for other social categories, for instance, when people reasoned about mental illnesses (Phelan, 2005, for an extensive review see Dar-Nimrod & Heine, 2011).

In any case, the plausibility that BGD is the core element of essentialist theorizing about social categories has lately been brought into question: For one thing, people do not exclusively report genetic deterministic beliefs when explaining the characteristic properties of others. Instead, they consider multiple causal factors (see Jayaratne et al., 2009). Secondly, not everyone endorses genetic deterministic explanations to the same extent. In fact, some people reject them altogether (Keller, 2005; Rangel & Keller, 2011). If people indeed considered genes as essences across the board, "we would definitely expect that all individuals endorse BGD to a certain degree" (Rangel & Keller, 2011, p. 1058). In light of this, Rangel and Keller (2011) propose a different, *externalist* version of psychological essentialism: **Belief in Social Determinism** (henceforth abbreviated BSD) describes the conviction that a person's essence is sustainably influenced by social factors such as upbringing, education, social background or peer contact. Strictly speaking, BSD agrees with internalist theories on the assumption that people believe in an underlying, internal property – an essence or a so-called "social character" (Rangel & Keller,

[4] It should be emphasized that people's common understanding of genetics is often quite rudimentary and misconceived: "Rather than as components of a biological process for building proteins" (Heine et al., 2017, p. 145), genes are commonly perceived as simple, agentive and monocausal units – an overgeneralization of the Mendelian law of Inheritance. In view of the large overlap between this misconception and psychological essentialism, genes are rather prone to fill in for an essence placeholder. Note, however, that this is culturally dependent: Sousa et al. (2002) and Waxman et al. (2007) found that Brazilian and Menominee children rather considered blood as a possible candidate for essences.

2011, p. 1058) – which causes the observable properties of all category members. The difference towards internalist theories, however, is that BSD acknowledges the impact of external factors on the underlying social character and, therefore, on the very nature of a social category. As compared to BGD, social deterministic beliefs do not entail the content of an essence placeholder (e.g. genes), nor do they focus on the internal structure of a category itself. They rather build upon factors which are located outside of categories but constitute strong and stable constraints on their mental representation. In this regard, such factors are to be distinguished from situational factors which emerge from the immediate context and merely induce behaviour of individuals, rather than the characteristics of an entire category (see attribution theory e.g. Kelley, 1973; or situationism e.g. Ross & Nisbett, 2011).

Rangel and Keller (2011) investigated the prevalence of BSD more closely: People consistently endorsed statements reflecting genetic and social determinism with overall higher ratings for BSD. Furthermore, measures of BSD were reliable predictors of prejudice endorsement, stereotyping and discriminatory tendencies. The study moreover showed that social deterministic information elicited in-group favouritism such that participants rated their respective ingroup more positively than an outgroup. Further evidence can be gleaned from developmental studies: In switched-at-birth tasks, Taylor et al. (2009) found an incremental gap between ten-year-olds' biological and social concepts. While younger children, in general, predicted behaviour based on heredity, children at the age of ten gravitated towards environment-based predictions for gender-linked behaviour. The study further investigated whether children stick to their predictions when presented with an alternative prediction: In this respect, all age groups endorsed an environmental impact on gender-linked behaviour, but maintained their original heredity-based predictions for animals (Taylor et al., 2009). Overall, findings point to the fact that BSD is strongly implemented in people's essentialist beliefs about social categories and emerges at an early developmental stage.

Interestingly enough, BSD can account for individual differences in people's deterministic explanations as reported by various studies (e.g. Jayaratne et al., 2009; Keller, 2005; Rangel & Keller, 2011): Different people may identify different social factors when they explain behaviour or infer properties to a social group (Rangel & Keller, 2011). BSD furthermore grants people the possibility of complex explanations: Rather than committing to a single essential cause, people may consider multiple causes (e.g. social status, upbringing, socialization) which operate on a category from the outside and shape its characteristics to different extents. By contrast, genetic determinism describes a monocausal essentialist belief: Given that people's understanding of genetics is often quite rudimental, they usually attribute cause to genes in

general, rather than to individual chromosomes or to interactions between specific sequences of nucleotides (Heine et al., 2017). However, BSD does not exclude genetic determinism or any other lay theory about the nature of essences (Rangel & Keller, 2011). In fact, the theory accommodates people's general belief in the existence of a stable, internal group essence, while also presuming a relative contribution of external social factors. In that sense, BSD is a hybrid theory which shares basic assumptions with internalist as well as with externalist approaches: Associations between a social category and its inherent properties are the result of constraints which both arise from the category's internal structure as well as from stable influences from an outside source.

2.2.2. Structural Thinking

2.2.2.1. The External Structure of Social Categories

Although the BSD approach acknowledges the impact of external factors on people's beliefs about the social world, it remains silent as to the question how these constraints come into existence. More importantly, the theory lacks a specific mechanism which illustrates how external factors may shape a category's mental representation. On this matter, Vasilyeva, Gopnik, and Lombrozo (2018) propose a specific style of reasoning: People **think structurally** about social categories when they localize them within a broader social structure and identify constraints which act upon the structural elements. In that sense, representations of social categories are constrained by the "*structures* in which they are embedded" (Vasilyeva et al., 2018, p. 1736).

Let us start with the basic question of what structures are. In philosophy, Shapiro (1997, p. 73) defines a structure as an abstraction from a system which, in turn, is the "collection of objects with certain relations". Systems may appear in different domains. A family system, for instance, connects specific individuals via marital or blood relations. In the same vein, a symphony is a collection of specific tones which stand in a temporal-harmonic relation to one another. A game of chess, on the other hand, is a system of pieces which relate in a spatial arrangement and by means of possible moves. Abstracting away from the particular entities leaves us with a general **structure**, more precisely, a relational network that is common to all systems of similar type. In the case of a family system, the underlying structure consists of relations such as "parent of," "child of" or "spouse of" without further specifying the individuals of a particular family (Haslanger, 2016). Likewise, the structure of chess entails how king and rook pieces are allowed

to move across the chessboard, yet it turns a blind eye to the fact that they may be sculpted of marble, wood or plastic. Only their abstract positions – the "nodes in the structure" (Haslanger, 2016, p. 121) – are part of the relational network. In sum, structures map out "the interrelationships among the objects, [while] ignoring any features of them that do not affect how they relate to other objects in the system" (Shapiro, 1997, p. 74). Two important points follow from this definition: Firstly, there is a distinction between the actual objects of a system (e.g. the particular family members) and the abstract positions they occupy in a structure (e.g. the position of a parent, a child, a spouse, etc., see Haslanger, 2016; Shapiro, 1997). Secondly, we can define structural constraints in terms of interrelationships so that properties which are associated with a node are restrained by the relations to other structural positions. Crucially, structural constraints act upon the "structure as such, independent of any exemplifications it may have" (Shapiro, 1997, p. 83). In a family system, for example, parents are liable for their children by virtue of an underlying parent-child relationship. Responsibility for one's own children, however, constrains the entire parent node, regardless of whether it is occupied by Jimmy, John or Sue (Haslanger, 2016). By the same token, the social category "women" occupies a structural node in which its members are monetarily disadvantaged. Yet, under different circumstances, any other social category could fill this position and be similarly restrained (e.g. in a different cultural context or a different social arrangement; Vasilyeva et al., 2018). In sum, structural constraints can be defined as relations between the abstract entities of a structure (Haslanger, 2016). Relations may be causal (e.g. a seventh cord causes a resolving chord in symphonic structure), constitutive (e.g. a rook is determined by its possible moves across the chessboard) or regulative (e.g. parents ought to behave in a certain way as caregivers). Beyond these particularities, they generally constrain properties which are attributed to positions in a structure and, thus, determine the distribution of possible outcomes. Vasilyeva et al. (2018) argue that the social structure is similarly represented in the conceptual system. In this respect, they propose two different representational levels for "differentiating *kinds* […] from the *structures* in which they are embedded" (Vasilyeva et al., 2018, p. 1736): On a first level, social categories are represented in terms of their internal organization so that properties are directly associated with a social category and determined by its internal features (i.e. an essence, inherent features, aspect representations, see chapter 2.1.). On a second level, people represent society as a network of abstract positions and interrelations. Rather than attributing features to some particular category, properties are connected to structural nodes. When people think structurally about social categories, both levels of representation interact: That is to say, people locate social categories within the social structure and, thus, identify them

as **occupants** of particular social positions. As such, a social category displays the properties of its node by virtue of its structural location. There is an important addendum: Distinguishing between representations of social categories and their structural positions should allow us to distinguish properties of positions from properties of categories. More precisely, the observable characteristics and behaviours of others should either clearly be attributed to a person's internal features or the person's structural position.[5] A clear-cut distinction, however, fails to appear: Consider the example of little girls frequently wearing pink clothing (Vasilyeva et al., 2018). The property "wears pink" can be caused both by the category itself (e.g. intrinsic preferences) as well as by its structural position (e.g. stock availability in stores) such that internal and structural constraints seem to yield the same outcome. Vasilyeva et al. (2018, p. 1736) explain this fact as follows: "Structural thinking [...] is sensitive to *confounds*" between social categories and structural positions. When people think structurally about social categories, they tend to take categories for the actual nodes in the social structure. In this way, properties determined by structure are directly attributed to the category itself. The same observable properties hence become consistent with two causal models: A social category's characteristics are either caused by internal features (i.e. internalist thinking) or they result from their position in the social structure (i.e. structural thinking).

2.2.2.2. Social Categories vs. Role-Governed Categories

The importance of a confound in the structural approach emerges clearly when we compare social categories to another similarly represented kind of concepts, namely role-governed categories. A comparison is furthermore interesting since the idea of an external structure is far from new in the psychology of concepts: Barr and Caplan (1985, 1987) first distinguished between *intrinsic* and *extrinsic* features whereby the former relate to a concept in isolation (e.g. "has wings" and "has fur" is true for birds and dogs, respectively) and the latter describe relations towards other concepts (e.g. "used to work with" relates a hammer to a handyman, "covers" links clothing to people, see Barr & Caplan, 1985, p. 181, 1987, p. 398). Crucially, extrinsic features are distinct from the object itself: "If, for example, a hammer were no longer

[5] This distinction is explicit in other frameworks: In attribution theory, three different causal factors contribute to people's interpretation of the behaviours of others: Accordingly, behaviour can either be elicited by a person (i.e. their internal features), by entities (i.e. a particular stimulus), or by times (i.e. the situational setting, see Kelley, 1973). The difference between attribution theory and structural thinking, however, is that the former explains behaviour of individuals rather than properties of an entire social group. Moreover, situational factors are not structural, but emerge from the immediate context. Most importantly, persons, stimuli and situations are independent variables so that they always yield different behavioural outcomes (Kelley, 1973).

used to work with, the hammer itself would remain the same" (Barr & Caplan, 1987, p. 398). A similar distinction was drawn for entire categories and their mental representations (see e.g. Gentner, 2005; Gentner & Asmuth, 2017; Gentner & Kurtz, 2005; Goldwater, Bainbridge, & Murphy, 2016; Goldwater, Markman, & Stilwell, 2011; Markman & Stilwell, 2001): Whereas so-called feature-based categories are determined by shared internal properties such that their members display high similarity (e.g. tulip, cow), relational categories are not necessarily intrinsically similar. They are rather characterized by a common relational structure, that is their meaning is "centered on relations with other concepts" (Gentner & Kurtz, 2005, p. 154). Relational categories may be nominal (e.g. a "bridge" links two entities or places), verbal (e.g. "to play" relates an agent with a game), prepositional (e.g. "between" locates an entity in the separating space of other entities) or adjectival (e.g. "bigger" compares entities regarding amount or size). More importantly, they can be further subdivided (Gentner & Kurtz, 2005; Goldwater et al., 2011; Markman & Stilwell, 2001): While schema-governed categories name entire relational structures (e.g. robbery, visit, family), role-governed categories are characterized by playing particular roles within such structures. A "thief", for instance, designates a stealing agent, a "victim" relates to the robbed experiencer and "goods" refers to the property owned by the experiencer, yet stolen by the agent. By the same means, a "guest" characterizes a visiting individual, whereas a "host" designates the receiving party of a visit. In a nutshell, role-governed categories name the positions in a relational structure and display a range of properties by virtue of their structural location.

Although roughly equivalent to the notion of structural nodes, representations of role-governed categories and social categories are by no means the same thing. Take, for instance, the role-governed category "guest" (Goldwater et al., 2011). The category specifies the first argument of a relation "X visits Y", but apart from a common relational position, its members are highly variant as to their internal features (Markman & Stilwell, 2001): Guests may be male or female, children or adults, they may be strangers in the case of hotel guests, an acquaintance or friend in the case of a house guest, or even a dragonfly visiting one's garden (Goldwater et al., 2011). A guest thus refers to a diverse set of feature-based categories (i.e. man, woman, dragonfly, friend), all of which are joined together via the same relational role across different contexts. Accordingly, when we think of guests, we have different possible exemplifications in mind which potentially occupy the relational position in question. The opposite holds true for the representation of social positions: Much like role-governed categories, nodes in a social structure are characterized by their relations towards other positions. Furthermore, they are likewise occupied by different social categories in different cultural and societal contexts

(Vasilyeva et al., 2018). A range of possible alternatives, however, is not part and parcel of people's conceptions: When we think of a social position (e.g. the position occupied by women), we think of an exemplification (e.g. the category "woman" itself), rather than a set of different social categories which potentially fill this position.

Experimental evidence points in a similar direction: Regarding role-governed categories, Barr and Caplan (1987) found that partial membership was significantly more often assigned to categories specified via extrinsic features (e.g. vehicles, weapons, tools). Ratings for intrinsically defined categories, on the other hand, were less graded so that borderline members were either perceived as absolute or non-members (e.g. birds, mammals, vegetables). In a similar vein, Kurtz and Gentner (2001) asked participants to name as many examples as possible for a given category whereby people generated a greater number of singular examples for relational categories. As for feature-based categories, the response patterns were significantly more clustered. Findings overall suggest that people consider a variety of possible members when they reason about role-governed categories. Since their membership is determined by a common relational role rather than a number of common intrinsic features, category members can be very diverse in nature. Representations of social categories, by contrast, seem to be less varied: In their studies, Haslam et al. (2000, 2002) observed that people perceive social categories either as **natural kinds** with sharp boundaries "whose members share necessary properties or microstructures" (e.g. gender, race, ethnicity; Haslam et al., 2000, p. 120), or as **entitative categories** which indicate uniformity and exclusivity amongst category members (e.g. political groups, Jews, homosexuals, AIDS patients). Despite their resemblance with role-governed categories in denoting roles in a relational structure, social categories thus seem to be more similar to feature-based categories: Their members are perceived as invariant and defined in terms of shared internal properties. How can the notion of confounding shed light on this matter? First of all, let us assume that both role-governed categories and social categories are represented twofold in the conceptual system. Representations may once refer to a structural position and once to the occupying categories themselves. The difference between social and role-governed categories, however, is that the two representational levels are confounded in the case of social categories. When people think about a structural node in society, they expect a particular category to occupy a particular position. In other words, instead of reasoning about an abstract position, they think of a specific feature-based exemplification (Vasilyeva et al., 2018). Let us turn the tables for a moment and imagine confounds to be implemented in role-governed categories as well: In such a case, tools would be tantamount to hammers, thieves would wear masks for personal comfort or style, and we would expect guests

only to refer to male strangers who watch pay TV and use the minibar. Even though non-sensical, the examples illustrate what confounding ultimately does: The focus is shifted from a diverse set of categories towards one in particular. In this way, properties which are determined by structure are aligned with the internal features of a category. Note that by this means, a confound has the same effects as an internalist bias: It promotes homogeneity and the inductive potential of a social category (Vasilyeva et al., 2018).

The structural view does not further elaborate why confounding exclusively applies to social categories, yet the following explanation is feasible: While people are regularly faced with a diversity of role-governed categories across different contexts, variation regarding social positions is quite imperceptible. In fact, possible alternatives often only become apparent via historical or cross-cultural comparisons. For instance, in our former example, we have assigned little girls to a structural position in which pink clothing is most available to them. Back in the 19th century, however, the opposite was true. Pink – "as a watered-down, bold, dramatic red" (Frassanito & Pettorini, 2008, p. 881) – was predominantly considered to be a boy's colour. Girls, on the other hand, were primarily associated with a dark blue, in allusion to art-historical depictions of the Virgin Mary. Given that such variation often receives little to no attention, people tend to rely on the consistent information provided by their immediate contexts. Accordingly, they reason about a particular social category instead of a diverse set of possibilities.

2.2.2.3. A Novel Perspective on Formal Explanations

Up until this point, it has been rather difficult to keep the two representational levels of social categories apart. Considering the fact that they are confounded for the most part, a legitimate question to ask is why we should even distinguish between representations of social categories and their structural positions. On this note, it is worth revisiting the matter of formal explanations (see chapter 2.1.3.1.): There is wide agreement that generating theories and explanations is a key aspect of human cognition (Gopnik, 1998; Gopnik & Meltzoff, 1997; Lombrozo, 2006; Prasada, 2017; Prasada & Dillingham, 2009). Accordingly, people constantly interpret their environment by use of different modes of explanations. One of these modes – the so-called formal mode of explanation – singles out certain properties of an entity and interprets them by appealing to category membership. The properties of an entity are thus formally explained by virtue of being the kind of thing it is (e.g. "Mittens has whiskers *because she's a cat,*" Gelman, Cimpian, & Roberts, 2018, p. 56; "This carrot is crunchy *because it's a carrot,*"

Lombrozo, 2006, p. 465). Formal explanations are frequent in everyday discourse and readily produced as early as at age of four and five (Haward et al., 2018; Taylor et al., 2009). The most puzzling aspect, however, is that they are perceived informative even though they are hardly more than tautologies (Gelman et al., 2018). That is to say, stating that *this* carrot is a carrot does not add any information to our state of knowledge, yet the explanation seems acceptable and natural sounding. Against this backdrop, Gelman, Cimpian, and Roberts (2018) argued that formal explanations are more than mere references to category membership, but "actually constitute placeholders for more-specific explanations" (Gelman et al., 2018, p. 44). According to the authors, not every explanation is apt to fill in for placeholders. Only explanations relating to the inherent features of a category may be considered as possible candidates. In their study, they found that participants indeed judged inherent explanations as more specific versions of formal explanations (Gelman et al., 2018). Non-intrinsic situational explanations, on the other hand, were not conceived as restatements. Importantly, the study was only obtained with reference to natural kind concepts (i.e. animals, plants, substances) so that social categories might depart from this pattern. Vasilyeva et al. (2018), for instance, point out that formal explanations for social categories are in fact ambiguous: The statement "Smith quit her job after the baby because she's a woman" (Vasilyeva et al., 2018, p. 1736) can either prompt a search for inherent causes (i.e. appealing to Smith's female abilities or priorities) or support a structural explanation which highlights constraints associated with her structural position (i.e. a gender wage gap). Arguably, the ambiguity of formal explanations is well explained in reference to the structural view on social categories: Firstly, if representations of social categories and their positions are confounded, category labels will be ambiguous and refer to both the feature-based representations of social categories as well as their abstract positions in the social structure. This lexical ambiguity is straightforward: Consider, for example, the fact that we easily distinguish between role-governed categories and their feature-based exemplifications by using different labels (e.g. guest vs. woman, stranger, dragonfly). By contrast, in the case of social categories, we depend upon descriptions in order to explicitly mark out the very same difference. Secondly, formal explanations invoke category labels. If formal explanations indeed prompt to look for some deeper related explanation, the search will proceed along the two lines of meaning of a category label: Formal explanations refer to categories themselves and support inherent interpretations, while at the same time, they make reference to a relational position and promote structural explanations. Vasilyeva et al. (2018) investigated this possibility in further detail. In their study, adults and children from three to six years of age were introduced to illustrations of a school with separated classrooms for boys and girls. Participants were told that

each student regularly played one of two games whereby tossing a stone into either a yellow or a green bucket determined whether they played "yellow ball" or "green ball" that day. They were furthermore informed that girls tended to play yellow ball more often, whereas boys predominantly played green ball. Crucially, the buckets were depicted the same size in a non-structural condition so "that the structural factors [...] did not favor one game over the other" (Vasilyeva et al., 2018, p. 1738). In a structural condition, the yellow bucket was larger in the girl's classroom and vice versa in the boys' classroom so that a stable structural constraint on the distribution of outcomes was invoked. Participants were asked to generate explanations, evaluate three types of causal explanations (internalist, structural, incidental) and additionally rate formal explanations for why a particular girl played yellow ball a lot (e.g. "Because Suzy is a girl," see Vasilyeva et al., 2018, p. 1739). Results showed that inherent explanations were produced significantly more often in the non-structural condition, while structural explanations prevailed in the structural condition. More precisely, participants got increasingly more sensitive to the different conditions with age. Yet even the youngest children successfully distinguished between meaningful and incidental explanations. In the formal explanation task, ratings did not differ significantly across conditions or age groups, indicating that the ability to reason about structure develops early on in childhood and both internalist and structural interpretations are induced by formal explanations.

At bottom, what is the added value of a structural approach to social categories? First of all, distinguishing between categories and their respective structural position renders possible a sharp distinction between internal and external constraints on category representations: People may either think about categories and attribute causation of their properties to internal characteristics (i.e. internal constraints), or reason about an external structure by identifying causative relations towards other structural positions (i.e. structural constraints). In this regard, the structural view goes beyond merely acknowledging the impact of external factors on people's beliefs. As opposed to the BSD approach, it traces back different aspects of people's social beliefs to different levels of representation and, therefore, explains how they emerge in the first place. It should be noted that both BSD and structural thinking account for the same observation: People do not only consider essential or inherent properties when they explain the characteristics and behaviours of social groups. They further identify constraints which stably act on a social category from an external source (Rangel & Keller, 2011; Vasilyeva et al., 2018). The difference between both approaches possibly resides in their origin of discipline: While BSD stems from social psychology and describes the content of people's beliefs and their effects on social perception, structural thinking relates to cognitive psychology and aims to

disclose an underlying mechanism. In that sense, social deterministic beliefs might result from structural reasoning about social categories.

Secondly, the structural approach casts light upon some unique characteristics of social categories: As previously described (see chapter 2.1.1.2.), developmental trajectories for biological and social concepts are quite different. While biological concepts are consistently essentialized as early as at the age of four and five (e.g. Astuti et al., 2004; Gelman, 2003; Rhodes & Gelman, 2009; Sousa et al., 2002; Waxman et al., 2007), racial essentialism emerges later and grows stronger as a function of cultural context (e.g. Diesendruck et al., 2013; Kinzler & Dautel, 2012; Mandalaywala et al., 2018). Gender-related essentialism, in turn, is strongly represented in early childhood, yet diminishes over the course of time (e.g. Taylor et al., 2009, but see Eidson & Coley, 2014). From a structural perspective, this developmental pattern may in fact reflect two interacting cognitive biases: As was noted earlier, structural thinking – more precisely, the confound between categories and positions – shows similar characteristics as an internalist bias (i.e. sharp category boundaries, inductive potential, licensing formal explanations). Since most developmental studies assess children's essentialist beliefs by inquiring about such aspects of belief, experimental evidence could support both internalist and structural reasoning about social categories (Vasilyeva et al., 2018). Ultimately, the structural view calls for a re-evaluation of the experimental data as regards the questions when children begin to think structurally about social categories and how structural constraints may interact with the essentialist bias.

As a third point, the structural view reconciles social categories with another branch of research, namely relational or role-governed categories. Both types of categories are defined by their relational commonalities which "hold across space and time" (Christie, 2017, p. 769). Accordingly, abstracting a common relational structure seems to play a crucial role in representing both role-governed and social concepts. However, how does the conceptual system extract a shared relational structure? In this respect, Gentner and colleagues proposed a domain-general learning mechanism, a process of comparison, in which two representations are aligned in order to highlight their common relational features (see chapter 3.2.2.3.1. for a detailed description of the mechanism). Crucially, structural alignment is engaged in various cognitive capacities: Originally formulated as a theory of analogy and similarity (e.g. Gentner, 1983, 2003; Gentner & Markman, 1997; Markman & Gentner, 2000), it has been suggested as a foundation of relational knowledge (e.g. Gentner & Kurtz, 2005; Haryu, Imai, & Okada, 2011), processing metaphors (e.g. Bowdle & Gentner, 2005; Gentner, 1988; Gentner & Bowdle, 2008; Wolff & Gentner, 2011), goal and intention learning (e.g. Christie, 2017; Gerson & Woodward,

2012), perspective taking (e.g. Meltzoff, 2005) as well as the development of a Theory of Mind (e.g. Bach, 2014; Christie, 2017; Hoyos, Horton, & Gentner, 2015). Provided that social categories are similarly characterized by an external structure, we may assume the same mechanism to be an integral part of their mental representation as well.

2.3. Summary and Discussion

Social categories tend to be conceived as sharply defined, immutable classes and homogenous regarding their properties. Consequently, characteristic properties and behaviours are often foregrounded in processes of categorization and inferential reasoning – a circumstance which sets the stage for stereotyping and prejudice endorsement. In the previous chapter, we have encountered a number of constraints which may be implemented in the representation of social categories and give rise to those aspects of social categorization. From a theory-based perspective, such constraints are broader cognitive structures which come in the shape of naïve theories (i.e. psychological essentialism, Belief in Genetic and Social Determinism) or in form of general cognitive mechanisms, either describing processes of reasoning (i.e. inherence-based and structural thinking) or a particular means of representation (i.e. aspect representation). They may operate across various domains (i.e. Inherence Heuristic, structural reasoning, aspect representation), only relate to a small range (i.e. psychological essentialism) or apply to one domain in particular (i.e. genetic and social deterministic beliefs, category-position-confounding). Above all, constraints provide different perspectives on the question how the relevant properties of a property are associated to it in the conceptual system.

By far the most extensively studied constraint in this respect is **Psychological Essentialism** (e.g. Gelman, 2003, 2004; Gelman & Ware, 2012; Bastian & Haslam, 2006; Prentice & Miller, 2007). Psychological essentialism refers to a widespread system of belief – a conceptual bias – by which people believe in the existence of naturally given kinds. They furthermore assume that each natural kind possesses an underlying essence that causes the observable characteristics of all category members. Essences are in fact "placeholder" notions (Medin & Ortony, 1989, p. 184): People can rarely name their actual identity, yet they are generally considered to be inherited, immutable, stable and, above all, defining a kind's identity. Since they elicit the observable properties of a natural kind, their characteristics are likewise conceived as permanent, determined by birth and true for all category members. Psychological essentialism is a naïve theory that does not solely pertain to the social domain, but also to biological and

natural concepts. In consequence, some of its subtypes have been put forward, exclusively relating to the representation of social categories: According to the **Belief in Genetic Determinism** (BGD), people generally identify the human genome as the underlying essence of our social world (e.g. Dar-Nimrod & Heine, 2011; Keller, 2005). Social groups are characterized by different genetic make-ups which inevitably give rise to their typical properties and behaviours. However, BGD may not be the only element of essentialist theorizing about the social world: People also believe in **Social Determinism** (BSD) so that the essential character of a person is permanently influenced by social factors such as upbringing, education, social background or peer contact (Rangel & Keller, 2011). From this point of view, the characteristic features of a social category are still mediated via internal essences, yet outside factors likewise have a substantial impact on the identity of a social group. At bottom, a rich pool of research suggests that psychological essentialism comes in various shapes and forms and takes effect to different extents in different domains. The consequences of essentialist beliefs, however, are the same across the board: People represent certain categories in terms of an underlying, unobservable essence which maps out a naturally given order, allows to generalize the properties of a category to all of its members and, furthermore, marks sharp distinctions between categories.

In the previous chapter, different cognitive mechanisms have been outlined which contribute to social category representation by establishing particular property-category-associations. People, for instance, frequently employ an intuitive reasoning mechanism, the so-called **Inherence Heuristic**, which is set off by any observable pattern in the world (e.g. Cimpian, 2015; Cimpian & Salomon, 2014). Once a regularity is detected (i.e. a correlation between a property and a category), people identify its components and retrieve the most accessible and relevant information from semantic memory. People then construct explanations by assembling the retrieved information in a causal or normative manner. Given that intrinsic features are generally more salient to categories than other properties, intuitive explanations turn out to be disproportionately inherence-based. The Inherence Heuristic illustrates how people reason about the behavioural patterns of others by reference to their internal properties. Furthermore, it serves as a basis for a variety of unrelated psychological phenomena such as system justification tendencies, the impact of feedback on motivation, or beliefs in nominal realism (Salomon & Cimpian, 2014). Most importantly, the Inherence Heuristic may constitute the underlying computational mechanism for psychological essentialism. Accordingly, the Inherence Heuristic would be narrowed down over the course of development so that only

information about essences is eventually accessed when people intuitively explain the properties of social and natural kinds.

The fact that some features are more central to social categories than others may also result from a representational constraint. According to the **Aspect Hypothesis**, relations between properties and concepts are represented in different ways and only a particular type – that is, properties with so-called principled connections – is subject to generalizations and promotes category homogeneity (e.g. Prasada & Dillingham, 2006, 2009). Principled connections represent properties as the aspects of a kind, they invoke statistical and normative expectations as regards a property's prevalence and, moreover, license formal explanations which explain features in terms of category membership. While principled connections exist throughout all domains, differences arise regarding their amount and the property type they connect to a concept (i.e. what it is constituted of, what produced it, its intrinsic functional properties, see Prasada, 2017). For instance, representations of social categories contain a large number of principled connections to different property types. Their members are therefore expected to be homogenous with regard to a variety of properties. Artificial categories, on the other hand, bear less principled connections to properties of different property types. Fewer and different properties are thus generalized to the instantiations of an artefact. At bottom, the Aspect Hypothesis contributes to our understanding of psychological essentialism and inherence-based reasoning: In both cases, principled connections may be used to identify the relevant properties of a category so that only those properties are considered to be caused by an underlying essence or retrieved in inherence-based reasoning processes.

A representational perspective is also taken by the structural approach to social categories which assumes that properties are associated to a concept on two representational levels (Vasilyeva et al., 2018). Accordingly, social categories are represented as autonomous categories with an internal organization and, moreover, as the occupants of structural positions in society. Properties can thus either relate to the social categories themselves and be constrained by their internal features (i.e. essences, inherent properties) or to their structural position. In the latter case, they are subject to structural constraints (i.e. the relations to other structural positions). Most importantly, the two levels of representation result in different reasoning mechanisms: When people think about the characteristics and behaviours of social groups, they may either reason about categories themselves or **think structurally** about their respective properties. In other words, people either apply internalist constraints to category representations or they locate them within a social structure and identify constraints on their structural positions. However, the difference between the two reasoning styles generally goes

unnoticed because people confound social categories with their structural positions. In this way, both internalist and structural reasoning yield the same outcome: Social categories are perceived as coherent classes which bear great inductive potential as to their characteristic properties.

Chapter 2 presented a detailed review of different constraints which operate on the way the relevant properties of a social category are represented and foregrounded in processes of categorization and reasoning. Thus, a sharp distinction has been drawn between internalism and externalism: While internalist constraints generally highlight restrictions emerging from the inside of a category, externalist constraints focus on factors which originate from an external source. Crucially, differentiating between the constraints on social concepts is actually not an "either/or" question. In fact, it has been shown that some of the aforementioned constraints are closely interrelated: Psychological essentialism may emerge from different cognitive mechanisms which, amongst others, represent properties in a particular way (e.g. as the aspects of a kind via principled connections) or highlight the explanatory value of certain features in intuitive reasoning processes (e.g. inherent, essential features). Furthermore, while social determinism clearly constitutes a branch of psychological essentialism by referring to a person's unobservable essential character, people also take factors into consideration that arise from social structure (e.g. peer contact, upbringing, education). In that sense, essentialist and structural reasoning are possibly merged in social deterministic beliefs. The structural approach to social categories has a similar implication: Provided that there are two levels of representation and both levels are commonly confounded, constraints that apply on the structural level should intertwine with those relating to the category itself. In sum, there are various constraints which simultaneously operate on the representation of social categories. Regarding their effects, those constraints are fairly similar: They equally give rise to highly structured mental representations with sharp category boundaries, immutable and stable properties across a category. They fuel generalizations and inferences, and by extension, stereotypes and prejudiced attitudes. The differences, however, may become insightful when taking a closer look at how different constraints emerge and thrive in the first place. In other words, what are the means by which people learn about the relevant properties of a social category and the routes along which they generalize properties to an entire social group?

At this point, let us return to the overarching topic of this master's thesis. Assuming that all of the aforementioned constraints are likewise implemented in the representation of Jews as a social group, the questions is what enhances their impact in this particular case. Drawing on findings from two different research subjects, I argue that antisemitic metaphors are a crucial

means of constructing a sharp and highly structured representation of a Jewish social category. For one thing, a large amount of research has found an influence of generic language on internalist constraints such as psychological essentialism, inherence-based reasoning and the representation of principled connections. Given that antisemitic metaphors are nothing other than generic statements, the first part of the next chapter will give generics a closer inspection by reviewing their formal and cognitive characteristics as well as their impact on the internalist constraints of social categorization. Secondly, chapter 3.2. will elaborate how the metaphorical component of antisemitic metaphors may additionally spark off externalist constraints on social category representations due to its underlying comprehension mechanisms.

3. The Contribution of Language to Social Categorization

There is widespread agreement on the assumption that language is a key element to the transmission of social concepts and cultural knowledge across generations (Gelman & Roberts, 2017). Although people may also acquire social categories through direct observation and interaction – differences in the social domain often appear jointly with certain visual cues, social practices and behaviours, or constraining life conditions – linguistic input represents a far more effective way to learn about social reality: Language can inform us about the relevant social categories of the social world, it "provide[s] a socially transmitted system for efficiently communicating […] what belongs in those categories, and which attributes those categories possess" (Gelman & Roberts, 2017, p. 7901). Furthermore, language can impart conceptions that are merely alluded or cannot be deduced from direct experience in the first place. Language thus plays a crucial role in social and cultural learning, enabling a rich and informative representation of society as we know it.

Linguistic transmission of social knowledge may proceed in a different ways (Gelman & Roberts, 2017): People can, for instance, explicitly talk about the content of their conceptual representations. Since language is a generative system, it permits us to communicate an infinite number of messages about our fellow human beings and learn about their relevant properties in return. It goes without saying that this also applies to antisemitic metaphors: Above all, they convey negative property attributions and potentially add to people's mental representations of a Jewish social category. However, propositions are not the only way in which language contributes to social categorization. Much of what we learn about the social world remains unpronounced and rather derives from implicatures, presuppositions or other types of inferences (Gelman & Roberts, 2017). For instance, children do not need to be told that women are physically weaker than men. A reference such as "a strong woman" is sufficient to infer that strength is a relevant property dimension to a gender category and to furthermore draw the pragmatic inference that women are expected to be weaker (see e.g. Horowitz & Frank, 2016 for such contrast inferences). Social knowledge may be transmitted in yet another way by the linguistic input: According to Gelman and Roberts (2017, p. 7901), "language provides cognitive tools that aid with recall, transmission, and manipulation of concepts" (Gelman & Roberts, 2017, p. 7901). Importantly, such tools do not only help to consolidate information in long-term memory. They additionally tap into particular reasoning processes so that certain properties become increasingly relevant to the representation of a social category and more highlighted in subsequent reasoning processes. More to the point, language seems to supply

certain tools which profoundly shape the way social concepts are represented in the conceptual system, resulting in highly structured representations of a social group with sharp category boundaries. With this in mind, I argue that antisemitic metaphors have precisely these effects on people's social concepts as they piggyback not only on one but two of those cognitive tools. Put differently, beyond their negative propositional content, antisemitic metaphors bring together two components that impact the way a Jewish social category is mentally represented, namely generics and metaphors. Firstly, antisemitic metaphors may be aligned with generic statements which have previously proven to be central in the acquisition of general knowledge. Generics have furthermore been associated with the abovementioned internalist constraints on social categorization. Secondly, the comprehension of antisemitic metaphors may highlight external influences on a Jewish social category as well as its structural position in society. This way, antisemitic metaphors may push ahead the externalist constraints on social category representation based on their underlying computational processes of metaphor comprehension. All in all, I argue that antisemitic metaphors play a twofold role in the transmission of social knowledge: In an obvious manner, they convey negative associations so that people may flesh out their mental representations of a Jewish social group. The second aspect, however, may be more subtle, yet all the more pervasive: Antisemitic metaphors impact the structure of category representations such that people are increasingly prone to endorse essentialist and social deterministic beliefs about a Jewish social group, identify inherent or structural causes for their alleged characteristics and behaviours or conceive them as aspects of their social identity. In this respect, antisemitic metaphors lay the groundwork for stereotypes and prejudiced attitudes towards Jews and for further downstream consequences such as hostility, discrimination, or even war and genocide (Amodio, 2014; Lieberman et al., 2017).

3.1. Generic Language

Various academic disciplines have taken an interest in generic statements due to their exceptional characteristics: To begin with, generics have rather challenging truth conditions so that they exert a great fascination both in philosophy and linguistics (see e.g. Carlson, 2019; Krifka et al., 1995; Leslie & Lerner, 2016). At the same time, children do not seem to be troubled by them and acquire generic generalizations rather early and systematically in childhood (Hollander, Gelman, & Star, 2002; Pappas & Gelman, 1998). This paradox permits important conclusions about general cognition as well as the development of concepts, rendering generics particularly interesting to the field of cognitive science. It has furthermore

been noted that generic language has a profound impact on social cognition: "The simplicity of their surface form notwithstanding, generics are a prevalent and potent linguistic device for transmitting complicated, objectionable and typically false views about social kinds" (Wodak, Leslie, & Rhodes, 2015, p. 631). In this regard, they represent central impetus to impact people's conceptions of social categories.

In this master thesis, I assume that antisemitic metaphors share many of their features with generic statements and, as a consequence, have a similar influence on people's social concepts. Drawing on this assumption, the following chapter will take a closer look at the research subject of generics: First, we will discuss the formal aspects of generic language in order to assess their general meaning. Secondly, a cognitive perspective will be introduced which assumes that generics in truth reflect a basic default mode of cognition. In a third step, we will shift focus onto the internalist constraints on social category representations and examine how generics may enlarge their impact in the social domain. It is against this background that it becomes clear how antisemitic metaphors may shape the structure of our social concepts: Due to their status as generic statements, they promote essentialist beliefs and inherence-based reasoning and may furthermore establish principled, statistical or causal connections between a predicated property and a Jewish social category.

3.1.1. Formal Aspects

3.1.1.1. Preliminaries

Establishing an extensive analysis of generic language is a difficult endeavour given that genericity is not "a single, unified phenomenon" (Carlson, 2019, p. 247). For one thing, generic statements are varied in respect of their truth conditions:

(3) a. The dodo is extinct.
b. Ravens are black.
c. A duck lays eggs.
d. Ticks carry Lyme disease.

(Leslie & Lerner, 2016)

The sentence in (3a) can only be true if all dodos are extinct without exception. However, (3b) is true in spite of some rare exceptions in the form of albino ravens. (3c) and in (3d) are even more indulgent: Both sentences are true even though only mature female ducks lay eggs and

merely one percent of all ticks are, in fact, disease carriers. Statements such as "humans are right-handed," on the other hand, are unacceptable, though true for approximately ninety percent of humanity (Leslie & Lerner, 2016). A second major challenge for the analysis of generics follows from their variation in linguistic form. Generics may contain different subject constructions such as bare plurals (3b), definite singulars (3a) and, in some cases, indefinite singulars (3c, but see 4a). They may further contain mass nouns (4b), Latinate terms (4c), NPs with the determiner *any* (4d), or distal demonstrative DPs (4e).

(4) a. ??A dodo is extinct.
b. Gold is rare.
c. Acer rubrum (=the red maple tree) grows 40 to 60 feet tall.
d. Any lion is ferocious.
e. Those spotted owls are constantly being talked about by environmentalists.

(Carlson, 2019, pp. 255–257)

Interestingly, none of these NP constructions is exclusively marked for genericity and may occur in non-generic contexts (Carlson, 2019; Roszkowski, 2016). This is not only the case for English: Cross-linguistic comparison shows that no language in the world marks generic meaning in a unique and unambiguous way (e.g. by use of a specific generic article) so that genericity depends on a variety of variable factors (e.g. lexical semantics of the constituents, pragmatic and situational knowledge, tense and aspect marking, etc., see Behrens, 2005). Additionally, while generic sentences often display a two-part structure with a subject NP and a predicate as the relevant constituents of interpretation, this is by no means a requirement. As was pointed out by Barbara Partee in Carlson (1989), generic meaning may also relate to the grammatical object of a sentence: The dominant interpretation of (5) informs about weather forecasts and how they generally come into existence rather than "attributing a weather-forecasting function to 'all' computers" (Carlson, 1989, pp. 172f.). The subject NP is therefore interpreted existentially.

(5) A computer computes the daily weather forecast.

(Carlson, 1989, p. 172)

Considering that truth conditions and linguistic form fail to define what counts as generic language, the first question to be answered concerns formal diagnostics (Krifka et al., 1995;

Leslie & Lerner, 2016; Roszkowski, 2016): Firstly, generics do not allow for upwards entailment. That is to say, whereas non-generic sentences remain true when the subject is substituted by a superordinate term (6a), a similar replacement changes the truth value of a generic sentence (6b). Secondly, we can often distinguish between generics and non-generics by inserting an adverb of quantification. While generic sentences merely display subtle differences in meaning (7), the insertion has quite severe consequences for the interpretation of non-generic sentences (8).

(6) a. Berber lions escaped from the zoo ⇒ Lions escaped from the zoo.
b. Berber lions are extinct ⇏ Lions are extinct.
(7) a. A lion has a bushy tail.
b. A lion *usually* has a bushy tail.
(8) a. A lion stood in front of my tent.
b. A lion *usually* stood in front of my tent.

(Krifka et al., 1995, pp. 9,13)

The two diagnostic tests indicate that we are dealing with two different things here: While upward entailment requires an alternation of the NP constituent, the second test modifies the meaning of the entire sentence. Simply put, the generic meaning in (6) seems to depend on the NP, whereas in (7) and (8), genericity results from the sentence itself. To this effect, a distinction is traditionally drawn between two major types of generic language (Carlson, 2019; Krifka, 1987; Krifka et al., 1995; Leslie & Lerner, 2016): On the one hand, genericity can be located within the NP. The term *reference to a kind* relates to nominal constituents which do not refer to particular individuals or objects but rather directly to an entire sort (9). In this way, they do not allow for exceptions.

(9) a. The potato was first cultivated in South America.
b. Rice was being introduced into East Africa several centuries ago.
c. *A rat was reaching Australia in 1770 (generic reading)

(Krifka et al., 1995, pp. 2,12)

Krifka, in collaboration with Gerstner, calls this first generic type D-genericity (D for “definite” Krifka, 1987, p. 4). As illustrated by the examples (4a) and (9), only bare plurals and definite singulars but not indefinites may designate entire kinds: The sentence in (4a) contains a kind-

level predicate (i.e. “be extinct”) which is restricted to select kind-referring subjects only (Carlson, 2019; Krifka, 1987). Such predicates never occur with an indefinite singular in subject position. In (9), the definite NP and the mass noun receive a generic reading in combination with a non-stative predicate. For the indefinite singular (9c), however, there is only an existential reading available. Kind-referring NPs are not solely limited regarding their form (Krifka et al., 1995): While most generic NPs also allow for an existential reading and refer to single individuals (10a) and (10b),[6] generic reference of a definite NP can only emerge when nominals have a semantic connection to a “well-established kind” (10c).

(10) a. Bottles were standing all over the kitchen. (existential reading)
b. The Coke bottle has a narrow neck. (generic and existential reading)
c. ??The green bottle has a narrow neck. (generic reading)

(Krifka et al., 1995, p. 11)

The second type of genericity are so-called *characterizing sentences* which report regularities rather than particular events or facts. As opposed to kind-referring NPs, characterizing sentences allow for exceptions. They may relate to general normative or essential properties (11a) and (11b) or repeating events (11c). It is assumed that, for characterizing sentences, the generic meaning is either a feature of the whole sentence (Krifka et al., 1995) or a property of the VP (Mari, Beyssade, & Prete, 2013). In either case, the generic meaning does not depend on the NP’s characteristics so that all types of nominals may be subject of a characterizing sentence (I-genericity for “indefinite”, see Krifka, 1987, p. 4). Strictly speaking, characterizing sentences may even contain proper names and, thus, describe regularities about individuals (i.e. habitual sentences such as 11d).

(11) a. Milk is healthy.
b. A potato contains vitamin C, amino acids, protein and thiamine.
c. Professors drink whiskey.
d. John drinks whiskey.

(Krifka et al., 1995, pp. 3,8)

[6] There are some instances where a kind-referring NP lacks an existential reading (Carlson, 2019; Krifka et al., 1995). In so-called avantgarde interpretations (e.g. “Man set foot on the moon in 1969”, Krifka et al., 1995, p. 78), “man” can only be interpreted as reference to mankind and defies an existential reading (e.g. “*Man ate 128 pretzels in one hour in 1976”, Krifka et al., 1995, p. 83). Other examples where a kind-referring term cannot be interpreted existentially are complex NPs with explicit references to kinds (e.g. “this kind of tiger”, “each species of fish”, Krifka et al., 1995, p. 6).

Above all, generic references and characterizing sentences both have in common that they involve abstraction (Krifka et al., 1995): While kind-referring NPs abstract from individuals and objects to kinds, characterizing sentences abstract away from specific facts and events. Needless to say, the distinction between kind references and characterizing sentences is a rather technical one inasmuch as both types of genericity often co-occur in a single sentence (11a). This is also the case for antisemitic metaphors: For one thing, metaphors such as "Jews are parasites/puppeteers"[7] are, by all means, characterizing sentences. The fact that genericity is independent of the NP and emerges from the entire sentence can be illustrated by substitution: Both an indefinite and a proper name may replace the bare plural in subject position (e.g. "A Jew/John is a parasite/puppeteer"). However, antisemitic metaphors are likewise a case of D-genericity: The subject NP is a direct reference to a kind which can be selected by a kind-level predicate (e.g. "Jews are rare"). Either way, antisemitic metaphors can be identified as generic statements with both kind-referring and characterizing elements. This distinction will be particularly informative in chapter 3.1.2.2.1. where the role of generics in psychological essentialism is outlined.

3.1.1.2. Generics and their Logical Form

After covering some of the descriptive basics, let us turn towards the question how generics are formally represented on LF. One of the first influential proposals was put forward by Gregory Carlson (1977, 1980), who focused on bare plurals and their generic interpretation: First, he argues that bare plurals shall be treated in parallel with proper names such that they make direct reference to individuals. Yet, individuals denoted by bare plurals are different from individuals for proper names (for illustration, see Fig. 2). While proper names make direct reference to individual objects (O), bare plurals are single terms for entire kinds (K). The ontological

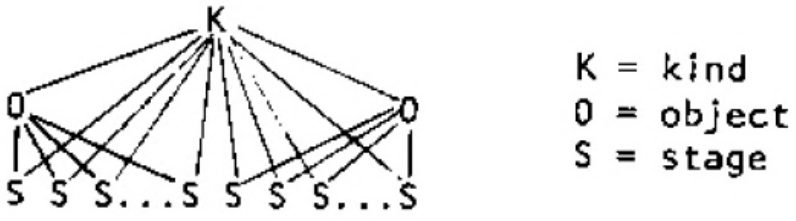

Figure 2. The hierarchic ontology of stages, objects and kinds (see Carlson, 1980, p. 69)

[7] Note that such nominal constructions merely serve as prototypical examples of antisemitic metaphors throughout the master's thesis. In actual fact, antisemitic metaphors come in a greater range than insinuated here and vary along different dimensions, including the characteristics of generic and metaphorical language as well as the particular context in which they occur.

difference may be captured as follows: Both kinds and individual objects consist of sets of so-called stages (S) which are the "spatially and temporally bounded manifestation[s] of something" (Carlson, 1980, p. 68). Both objects and kinds are identified with a number of different spatiotemporal slices of an individual as well as with the individual itself (Carlson, 1977). Kinds, however, have an additional intermediate level of representation: They are furthermore realized as individual objects. In this way, a kind ties together different objects of the same kind, ensuring that it is not existentially bound in time and space (Mari et al., 2013). Second, following Milsark (1974), Carlson (1980) distinguishes between different types of predicates whereby their applicability corresponds to the ontological levels of representation in Fig. 2: So-called *individual-level-predicates* denote permanent and stable characteristics of individuals (Krifka et al., 1995; Roszkowski, 2016). They further subdivide into *kind-level predicates* (e.g. "be extinct") and *object-level predicates* (e.g. "know French"), depending on which sort of individual they apply to. So-called *stage-level-predicates*, on the other hand, quantify over stages of an individual (e.g. "smoke"). Stage-level predicates may be predicated upon a single or several stages and may refer to both short or long stretches of time – in any case, they denote transitory events and only apply to the "parts of a whole" (Carlson, 1977, p. 448). With this distinction in mind, Carlson (1977, 1980) argues that different interpretations of bare plurals depend on the type of predicate used in a sentence rather than the meaning of the bare plural itself (Mari et al., 2013): Under all circumstances, bare plurals directly denote kinds. It is only in combination with stage-level predicates that particular stages of a kind are identified so that an event reading comes into effect (e.g. "Students are smoking"). For individual-level predicates, the denoted property is directly applied to the kind level so that a generic interpretation surfaces (e.g. "Dodos are extinct"). A similar pattern holds for proper names: We either receive an event reading (e.g. "John is smoking") or a habitual reading (e.g. "John knows French"), depending on the type of predicate and the ontological level of an individual to which the predicate applies.[8]

While the predicate distinction covers instances of D-genericity, characterizing sentences remain a challenge to Carlson's account (Carlson, 1980, 2019; Cohen, 2007; Krifka et al., 1995): Consider, for instance, a habitual sentence such as "John smokes" where a stage-level

[8] Carlson (1980) only roughly integrates definite and indefinite singulars into his account: While definite singulars may also directly denote kinds, indefinite singulars are never kind-referring expressions and cannot be selected by kind-level predicates as illustrated in (4a) and (9c). However, matters are more complicated (Mari et al., 2013): Definite singulars are not simply kind-referring variants of bare plurals, but show further constraints on their generic interpretation (i.e. a necessity for well-established kinds, 10c). Indefinite singulars, on the other hand, may receive a generic reading in sentences with reference to a taxonomic kind (e.g. "A whale—the blue whale—is becoming extinct", Mari et al., 2013, p. 27). To this day, most accounts are focusing on bare plurals, leaving other types of nominal constructions on the sideline (Leslie & Lerner, 2016).

predicate applies to the individual "John", rather than some of its particular stages. Likewise, object-level predicates may generalize over whole kinds rather than single objects (e.g. "Italians know French"). Both characterizing sentences are furthermore puzzling since they allow for exceptions so that not all John-stages are necessarily smokers and not every Italian-object must have French language skills. By contrast, D-generic statements such as "Dinosaurs are extinct" can only be true if every last dinosaur has vanished from the face of the earth. In order to account for such I-generics, Carlson (1980) introduces a generalization function in form of a covert monadic operator G'. The operator takes predicates as inputs and raises them to the next higher level of representation so that a stage-level predicate may be applied to objects and, after re-application of G', to kinds (Leslie, 2013). Importantly, while the application of G' requires at least two stages or individuals to satisfy the predicate, the operator does not inform about the precise number of stages or individuals for a generalization to hold true (Krifka et al., 1995; Leslie, 2013; Mari et al., 2013; Roszkowski, 2016). The reason for this is that, according to Carlson (1980, p. 168), establishing the exact truth conditions of a generic sentence does not fall within the remit of a semantic theory but is "a cognitive process" instead.

Carlson (1989) himself later dismissed the one-place operator G' as his analysis cannot account for cases where the internal argument receives a generic interpretation such as in (5). As an alternative, Krifka and colleagues (1987; 1995) proposed a dyadic operator which is characterized similarly to Lewis's (1975) adverbs of quantification: For one thing, the operator GEN takes sentential scope, rather than merely scoping over predicates (Collins, 2018). Secondly, GEN has a tripartite structure, consisting of a generic operator, a restrictor which determines the scope of generalization, and a matrix which picks out the property ascribed to the restricted scope (Krifka et al., 1995, p. 26; Leslie, 2008, p. 6; Mari et al., 2013, p. 66).

(12) GEN $[x_1 ... x_n]$ (Restrictor $[x_1 ... x_n]$) $\exists y_1 ... y_n$ (Scope $[x_1 ... x_n, y_1 ... y_n]$)

Crucially, the generic operator functions as an unselective quantifier so that it "bind[s] all the variables in [its] scope indiscriminately" (Lewis, 1975, p. 7). Thus, if a free variable is introduced to the scope of generalization by the restrictor (e.g. via indefinites such as bare plurals or indefinite singulars), it is bound by the generic operator irrespective of its type (Collins, 2018; Leslie, 2008, 2013; Leslie & Lerner, 2016). Variables in the matrix position, on the other hand, are bound by an existential quantifier (existential closure). A dyadic operator offers the advantage that concurrent generic readings are captured by a single structure of representation in LF. Reconsider the sentence in (5):

(5) A computer computes the daily weather forecast.

(Carlson, 1989, p. 172)

A first, weak reading indicates that computers are generally able to predict future weather conditions. In this case, the subject NP is placed in the restrictor, whereas the rest of the sentence provides the content of the matrix. A second, more dominant reading, however, generalizes about weather forecasts so that the object NP is placed in the restrictor position and the subject "a computer" is existentially bound in the matrix position. Above all, this second dominant reading can only be obtained by a dyadic operator which divides the material of an entire sentence into restrictor and matrix. By contrast, a monadic operator G' which operates on the VP level can merely yield the first unnatural interpretation (Leslie & Lerner, 2016).

At present, the standard treatment of generics follows along these two lines (Cohen, 2007; Collins, 2018; Leslie, 2013): Firstly, D-generics are generally considered as cases of direct kind predication so that a property is directly attributed to a kind-denoting term by the predicate of a sentence. This approach is underpinned by the fact that D-generics never apply to particular instances but always to the total of a kind without exception. I-generics, on the other hand, differ from direct kind predication in that they make allowance for exceptional kind members and may even express regularities of individual objects (e.g. habitual sentences). They do not generalize about kinds, but rather about facts and events. I-generics are standardly analysed in terms of an unpronounced generic operator GEN which has a tripartite structure as illustrated in (12).[9] GEN binds the free variables introduced by the restrictor and applies the predicate from the matrix to them (Carlson, 2019; Leslie, 2013; Leslie & Lerner, 2016). As opposed to D-generics, the generalized NPs of I-generics are therefore predicates themselves which provide unbound variables and, thus, restrict the scope of generalization. It is furthermore assumed that GEN is a default operator (Leslie, 2008): The operator is inserted into the logical form of a sentence whenever there is no other explicit quantificational operator available.

[9] There are few alternative conceptions: Liebesman (2011), for instance, claims that all generics – even I-generics – are direct kind predications (see also Teichman, 2016 who proposes a sophisticated kind-predicate version). According to his view, the fact that some generics allow for exceptions is based on a part-whole principle which is also true for other examples of direct predication: The sentence "This chair is wooden" is true despite some metal screws. Similar to Carlson (1980), Liebesman (2011) argues that such complexities are not semantic in nature but rather concern the metaphysical question of what it means for a kind to have a particular property (Leslie, 2013). Collins (2018), on the other hand, rejects genericity as a linguistic phenomenon altogether. Accordingly, generics do not contain a specific linguistic device such as GEN but rather emerge as "interface effects" between "independent syntactic forms [and] wider cognitive structures that support generic generalisations" (Collins, 2018, p. 61).

3.1.1.3. The Formal Meaning of GEN

While the representation of I-generics in terms of a dyadic operator is widely accepted, the interpretation of GEN remains a subject of controversial debate (Carlson, 2019; Leslie, 2008; Leslie & Lerner, 2016). Two perspectives are traditionally taken on this matter (Carlson, 1995; Roszkowski, 2016): For one thing, considering the similar logical forms of generics and quantificational statements, a first line of approach reconciles GEN with other quantificational elements. In this respect, GEN can be understood as a covert quantifier which rages over the instances of a variable (Collins, 2018). Treating I-generics as covert quantificational statements has the advantage that we can apply the same set-theoretic tools to generics as to non-generic statements in order to determine their truth conditions (Leslie, 2008; Prasada et al., 2013). A generic sentence is true if a sufficient number of "relevant individuals in the domain of the generic satisfy the predicated property" (Cohen, 2002, p. 63). The overarching question of quantificational approaches therefore is how a sufficient number of relevant individuals can be established, or rather, what kind of quantification is most appropriate (Carlson, 1980). Quantificational approaches come along with an important implication: If generic sentences can only be truly stated if sufficiently enough instances of a kind satisfy a particular property ascription, those have to be observable in some way. In other words, generics are considered to be generalizations which are grounded on observation and induction (Carlson, 1995; Cohen, 2002; Pelletier, 2010). In view of this, quantificational approaches are generally referred to as the "inductive" view on generics.

A second line of approach is the so-called "rules-and-regulations" view according to which the truth values of a generic sentence do not depend on a certain number of instances and their observed properties. On the contrary, generics are evaluated relative to general rules and regulations in the world. "Each generic sentence denotes a rule; if the rule is *in effect* […], the sentence is true, otherwise it is false" (Cohen, 2002, p. 60). Rules and regulations may be biological principles, physical laws, or social and moral norms. In any case, they are the causal forces behind the properties or events described in generic statements so that an individual or an instance of a kind occurs in a particular way because of a corresponding rule (Carlson, 1995). The difference between the two views on generics is best illustrated with an example: According to the rules-and-regulations view, the truth of a sentence such as "tigers are striped" does not depend on the fact that sufficiently enough striped tigers exist – whatever "sufficiently enough" means – but on a general (presumably) genetic rule which determines the stripedness of tigers (Cohen, 2002). In this sense, GEN is not a quantifying element, but an operator that is evaluated against some corresponding rule in the world (Carlson, 1995; Roszkowski, 2016).

Importantly, the rules-and-regulations view requires an additional ontological assumption: Both rules and regulations have to be considered as real and irreducible entities which cause the structures of the world.

It goes without saying that both the inductive and the rules-and-regulations view are umbrella terms for a series of very different approaches to the meaning of generics (for a broad review see Cohen, 2002): Within the range of inductive approaches, it has been argued that generics entail a covert universal quantifier which is restricted by context: Declerck (1991), for instance, argued that a hearer applies a generic statement only to the *relevant* members of a set whereby context and general world knowledge determines which members are of relevance. A sentence such as "[w]hales give birth to live young" is thus interpreted as "all relevant whales give birth to live young", or rather, "all female whales that are not sterile give birth to live young" (Declerck, 1991, p. 83). Another way to implement GEN as a restricted universal quantifier was proposed by so-called *normalcy-based approaches* which claim that generics quantify over the *normal* members of a kind (see also Dahl, 1975). According to Asher and Pelletier (1997, 2013), normality is best captured by a modal conditional analysis so that a generic sentence "φ is ψ" has the corresponding logical form $\forall x(\varphi > \psi)$. The consequent of the conditional ψ is only evaluated against those worlds in which a representative of φ is a normal φ. That is to say, although it is possible for instances of φ to lack ψ, this would not be the case if everything had run its normal course for the instance in question (Leslie & Lerner, 2016). Cohen (1999, 2004), on the other hand, argues that generics should be treated as *probabilistic judgements* whereby "probabilities are derived from the frequencies observed in the world" (Carlson, 2019, p. 245). A generic sentence GEN (φ,ψ) is true if the probability of φ displaying ψ is greater than 0.5 (Cohen, 1999, 2004). Generics may furthermore express relative probabilities (e.g. "Mosquitoes carry the West Nile virus"). In such case, ψ in φ must have a higher probability rate than ψ in its non-φ alternatives which are supplied by context (e.g. other kinds of insects, Leslie, 2008). In order to account for statements which are wrong despite the fact that their probability is higher than 0.5 (e.g. "People in [S]outheast Asia speak Chinese", Cohen, 1999, p. 245), Cohen suggests that generics entail a presupposition of homogeneity: If φ can be partitioned into *salient* subsets, they need to be homogenous with respect to the property ψ. In other words, speaking Chinese should roughly have the same probability for every salient subset (Cohen, 1999). Assuming that the salient partition of "people in [S]outheast Asia" (Cohen, 1999, p. 245) is one into regions, probabilities are presumably quite varied between countries as well as between urban and rural areas. Consequently, the generic sentence cannot be truly asserted.

As for the rules-and-regulations view, it has been argued that I-generics are in fact grounded in "(socially and culturally determined) beliefs" (Declerck, 1986, p. 168). Only if a speech community agrees that a property is stereotypical of an individual or a kind, a generic interpretation may arise. In this respect, Declerck (1986) argues that an operator GEN on LF is suitable to express or paraphrase a generic interpretation, however, it is by no means responsible for its emergence (for a stereotypical point of view see also Geurts, 1985). On a similar note, Heyer (1990) argues that I-generics make reference to the prototypes of a kind. Generics may therefore have exceptions since some members of a kind may be less prototypical than others. However, "[t]o decide the question whether or not a particular individual is contained in the set of typical representatives of a kind may […] require further knowledge and reasoning" (Heyer, 1990, p. 103). Importantly, both theories do not establish a relation between the generic meaning of a sentence and the properties of the instances of a kind. They rather take normative regularities as the basis of determining whether a generic sentence is true or false (Cohen, 2002).

Irrespective of their technical differences, all approaches run into problems at some point: Quantificational approaches generally struggle to explain why some generic sentences are true (e.g. "ducks lay eggs") while others are false (e.g. "ducks don't lay eggs", "ducks are female", see Leslie, 2008, p. 8; Leslie & Lerner, 2016). For instance, we might think of the latter sentences as cases of relevant quantification (e.g. "all male ducks don't lay eggs", "all female ducks are female"), however, it remains false to assert them (Krifka et al., 1995; Pelletier & Asher, 1997). Hence, relevant quantification is not capable to exclude false generics. Normalcy-based approaches are furthermore faced with the opposite problem: Generics such as "mosquitoes carry the West Nile virus" or "sharks attack swimmers" are true, even though it is neither *normal* for mosquitos to be carriers of the virus nor for sharks to attack people who swim (Asher & Pelletier, 2013; Leslie & Lerner, 2016).[10] The probabilistic account, on the other hand, has difficulties explaining what the salient partitions of a kind are. Cohen (1999, pp. 245ff.) proposes partitions regarding space (e.g. "People in [S]outheast Asia speak Chinese"), numerical scales (e.g. "People are over three years old") or gender (e.g. "Primary school teachers are female"). However, if gender indeed provided salient partitioned subsets,

[10] Asher and Pelletier (2013) try to remedy such problems by assuming that other factors such as discourse information may further restrict the universal quantifier in generic sentences (Leslie & Lerner, 2016). In the example above, a mosquito *normally* carries the West Nile virus "in the appropriate circumstances" (Asher & Pelletier, 2013, p. 331), that is, the circumstances under which the virus is acquired. Background knowledge may thus determine the appropriate circumstances of a generic sentence and enable inductive generalization even in cases where the absolute frequencies are counterintuitive to the notion of normality. Note that these considerations are similar to Declerck's (1991) proposal of relevant quantification.

we would likewise expect sentences such as "lions have manes" to be false (Leslie, 2008, p. 12). Over and above, quantificational approaches are faced with the so-called *problem of induction* (Carlson, 1995): The question at issue is how many instances of a kind need to satisfy a certain property so that a predicate can be generalized to all of its members. In other words, how many instances must be observable for a generic sentence to hold true? The problem of induction is further aggravated by the fact that no language in the world seems to pronounce a quantifying element GEN. The question, thus, is how generics can be identified as quantificational statements if no overt marking points at their inductive nature. As Liebesman (2011) concludes, the fact that generics retain their truth conditions even when rephrased as quantificational sentences gives the impression that both are two sides of the same coin. Yet, this "does not require that generics are, in any way, quantificational semantically" (Liebesman, 2011, p. 421). Rules-and-regulations approaches, on the other hand, are similarly problematic: Although the problem of induction is bypassed by the assumption that the truth values of generics do not depend on inductive inferencing but rather on a corresponding rule in effect, this is also one of the great weaknesses of the approach (Cohen, 2002). A number of questions arises: What are these corresponding rules? What does it mean for them to be in effect and, most importantly, what would the world be like if they were not? The rules-and-regulations view does not provide answers to these questions so that the truth conditions of generics are rather shallow: A generic sentence is true if a corresponding rule is in effect and false if it is not. Apart from this, it is not further determined which "states of affairs […] would make [the meaning of a sentence] true" (Cohen, 2002, p. 63). Ultimately, both inductive and rules-and-regulations approaches have their merits in explaining certain aspects of genericity. However, no approach can fully account for the overall scope of the phenomenon so that the question as regards the meaning of generic sentences is far from being settled (Lazaridou-Chatzigoga, 2019; Leslie, 2008; Leslie & Lerner, 2016).

3.1.2. Cognitive Aspects

In the past few years, generic language has been met with increased interest by cognitive scientists given that a large amount of our generic knowledge is expressed by means of generic statements. Accordingly, understanding how people evaluate generic sentences could provide insights into how generic knowledge is represented by the conceptual system (Carlson, 2010; Prasada, 2010; Prasada et al., 2013). Even though formal semantics is concerned with a very different question, namely how language relates to actual reality rather than people's subjective

knowledge structures, consulting semantic conceptions can be quite revealing (Leslie, 2008): The relevant question here is whether people understand the relation between language and the world in similar ways as proposed by semantic approaches. However, adopting a cognitive perspective on generic language is a two-way street: A cognitive point of view may, in turn, forge ahead with some of the open questions in semantics. That is to say, psychological findings may "place important constraints on what the semantics of generic sentences might look like" (Leslie & Lerner, 2016). The following chapters will explore the links between generics and cognition in further detail: Firstly, findings will be reviewed which indicate that generics are crucial means for expressing and learning generic knowledge about the world and that they are deeply implemented in general cognition (**Generics-as-Default-Hypothesis**). In the course of this review, implications for the meaning of generics will be discussed. In a second step, we will turn towards the question of how generics relate to the aspects of social categorization which have been reviewed in chapter 2. Based on the observation that generics have a major impact on social category representation, we may assume that antisemitic metaphors put to use their generic element and unleash essentialist and inherence-based reasoning about Jews as a social category. Additionally, the representation of negative properties as the aspects of a Jewish social category is promoted by their generic interpretation. In short, genericity in antisemitic metaphors is argued to be a major impetus for representing Jews as a highly structured social category with sharp category boundaries and immutable, stable properties across category members.

3.1.2.1. Generics as Defaults

One of the most puzzling questions from a cognitive point of view is why it is so hard to establish a semantic theory of generics while their acquisition seems to be a "cakewalk for language learners" (Leslie, 2008, p. 2): In their developmental study, Gelman, Goetz, Sarnecka, and Flukes (2008) analysed longitudinal data from the CHILDES database and found that generics were produced as early as at the age of two.[11] Their frequency significantly increased between the age of two and three, and by the age of four, children produced an equal amount

[11] In psychological studies, generics are more restricted than in formal conceptions. They are generally coded as such when they include an appropriate noun phrase (i.e. bare plurals, indefinite singulars, mass nouns), a present non-progressive tense and when individuating information is absent (i.e. number, possessive pronouns and deictic or specific temporal information). Definite generics and habitual sentences are therefore excluded from the scope of investigation. Furthermore, psychological studies do not explicitly distinguish between D- and I-genericity, however, both generic types can be found in the data (e.g. "I don't like *bread*" as a kind-referring NP vs. "Why *babies* can't play with children?" as a characterizing sentence, Gelman et al., 2008, p. 7,10).

of generics as their caretakers. Furthermore, children seem to systematically distinguish between generic and non-generic contexts: Pappas and Gelman (1998) examined conversations between mothers and their two- to four-year-olds when they looked at animal picture books together. The authors found that both mothers and children matched grammatical number to the number of animals depicted in the books when they uttered a non-generic sentence (e.g. "See [those] bat[s]? *They* came from the cave over there", Pappas & Gelman, 1998, p. 23). In generic sentences, however, the NP's linguistic form was independent of the depicted context such that mismatches in number frequently occurred (e.g. "See this bat? *They* live in caves", Pappas & Gelman, 1998, p. 23). These results indicate that generics do not occur by accident in child speech. Even young children distinguish between generic and non-generic meaning by use of a different distributional pattern. In fact, children seem to guide their inferences according to this difference: Graham, Nayer, and Gelman (2011) found that, when a previous demonstration of the property of a novel kind was accompanied by a generic sentence (e.g. "*Blinks* drink milk", Graham et al., 2011, p. 502), 30-month-old participants mimicked the action with the object already introduced as well as with a divergent exemplar of the same kind (i.e. an orange instead of a blue "blink"). By contrast, they picked the already known exemplar significantly more often when the sentence was non-generic (e.g. "*These blinks* drink milk", Graham et al., 2011, p. 502). Thus, children seem to understand from a very young age that generics report generalizations across kinds. Interesting results were furthermore obtained from American and Taiwanese deaf children between the age of three and four (Goldin-Meadow, Gelman, & Mylander, 2005). Despite the fact that they had not been exposed to a conventional sign language at the time of the investigation, the children readily produced generics in their own invented gesture systems (so-called "home signs", Goldin-Meadow et al., 2005, p. 110). More importantly, they did so at approximately the same rate as the hearing, age-matched control group. In summary, developmental findings suggest that generics are systematically uttered already at an early stage of childhood. Children differentiate between generic and non-generic statements and acquire them even in cases when their exposure to generic language is rather limited.

The developmental pattern conflicts with what we have learned about generics so far: For one thing, while the truth conditions of generics are fairly complex, even young children seem to master them with great ease (Gelman, 2010). Children produce generic statements considerably earlier than quantificational sentences, even though the latter have straightforward truth conditions and should therefore be acquired more easily (Leslie, 2008; Roeper, Strauss, & Pearson, 2006). What is more, genericity is not explicitly marked in any language of the world.

In light of this, we would expect it to be rather difficult for children to identify generic sentences in the first place. The early acquisition of generics raises the question how language learners successfully extract generic meaning when faced with these two challenges. Leslie (2007, 2008) answers to this question with the so-called **Generics-as-Default-Hypothesis.** According to this hypothesis, generics correspond to an innate, primitive cognitive mechanism that yields generalizations by default (see Gelman, 2010 for a similar point of view). Let us first take a look at the mechanism itself: Generally speaking, the capacity to generalize is fundamental for any form of inductive learning (Leslie, 2008; Roszkowski, 2016). Whether a child learns to keep away from hot stoves after a single touch or understands that fish do not survive in glass bowls, inductive learning requires a mechanism to abstract away from particular experiences to future items or events. Interestingly, this capacity is not restricted to human nature. All organisms, from humans to mealworms, efficiently adapt to their environment by using inductive generalizations (Baldwin, Markman, & Melartin, 1993; Gelman & Ware, 2012). In humans, however, inductive capacities are far more flexible and diverse with respect to the inferences that are drawn from a few instances and generalized to many. Such generalizations become evident quite early in childhood: Graham, Kilbreath, and Welder (2004) found that 13-month-old infants readily generalize non-obvious properties (e.g. a rattle sound) to objects of the same category by performing eliciting actions on them (e.g. shaking). Similar results were obtained with nine-month-old infants who expected non-obvious properties in objects of similar appearance (Baldwin et al., 1993) and eight-month-olds who applied an "innards" principle to agentive and self-propelled objects (Setoh et al., 2013). These early generalization skills suggest that children are innately equipped with a low-level cognitive mechanism which enables inductive reasoning and, thus, leads to the fast acquisition of generic knowledge. In other terms, children use a basic cognitive mechanism to gather information as fast and efficiently as possible about regularities in the world. Based on these considerations, Leslie (2007, p. 381) argues that generics merely "give voice to the generalizations this primitive mechanism produces". In this sense, children produce and understand generics effortlessly at an early age because they simply invoke one of the cognitive system's most basic default mechanisms.

The Generics-as-Default-Hypothesis shares some basic assumptions with formal semantic approaches: Leslie (2007, 2008) assumes that generics are represented in terms of a tripartite structure as illustrated in (12). She further agrees with the assumption that GEN is a default operator which binds the free variables in the restrictor domain and attributes the properties that are specified by the matrix to them. The difference, however, is the following: While formal approaches ground the meaning of GEN in either complex set-theoretic computations (i.e. the

quantificational view) or untraceable metaphysical realities (i.e. the rules-and-regulations view), the Generics-as-Default-Hypothesis claims that the interpretation of GEN is provided by a low-level cognitive mechanism which yields a generic interpretation by default. In that sense, the varying truth conditions of generics are "a consequence of cognition" (Lazaridou-Chatzigoga, 2019, p. 165) rather than semantics or metaphysics. What does this mean in more concrete terms? For one thing, if the primitive cognitive mechanism is as efficient as possible, it should not generalize every piece of information gathered from the environment but make a suitable selection (Leslie, 2008). For this purpose, we may assume certain cognitive principles which determine the kind of information that is generalized by the conceptual system. Among these principles are the so-called *characteristic dimensions* of a kind. Characteristic dimensions express regularities and prompt the search for information relating thereto, while leaving less important information aside. For animal kinds, they "include perhaps noises, modes of locomotion, means of reproduction and gestation, and diet" (Leslie, 2008, p. 30). In the case of artefacts and social categories, they may relate to their function or role in society. Characteristic dimensions may be innately given or provided by parental input. Whatever their actual nature and origin, they seem to be the background expectations – or rather "over-hypotheses" (see Shipley, 1983, p. 274) – which guide people's primitive inductive generalizations. According to Leslie (2007, 2008), the conceptual system is furthermore biased towards information where – instead of possessing an equally vivid alternative – the respective counterexamples simply lack the property in question (i.e. *negative* counterinstances). The reason for this is perceptual so that negative counterexamples can be overlooked quite easily, while positive properties are salient and hard to ignore. Consider, for instance, the true generic "ducks lay eggs" and the false generic sentence "ducks are female". In both cases, the property in question relates to characteristic regularities of the animal kingdom, namely, to the fact that animals reproduce in a certain way and that both males and females exist in each animal kind. However, the first example is a true generic because the counterinstances for egg-laying ducks simply fail to satisfy the property (i.e. male and immature ducks do not lay eggs). In the second case, counterinstances display a positive alternative property (i.e. male ducks are male) so that the information is not generalized to the entire kind. Similarly, a sentence such as "ducks don't lay eggs" fails to be true given that the alternative is positive (i.e. female ducks lay eggs). This first and second bias may also be implemented in the evaluation of antisemitic metaphors: Statements such as "Jews are puppeteers" or "Jews are parasites" provide information which relates metaphorically to a particular position in society – either at the top of the social structure to infer a social threat or at the bottom to dehumanize and delegitimize Jewish existence. The

process of generalization succeeds in both cases since counterinstances lack the attributed property, rather than displaying a corresponding positive alternative. In fact, it is questionable whether there are equally vivid alternatives for metaphoric ascriptions in general considering that their comprehension requires divergent processes of reasoning (see further discussion in chapter 3.2.). According to the Generics-as-Default-Hypothesis, a third cognitive principle may be responsible for the quirky truth conditions of generics (Leslie, 2007, 2008). In generalizations such as "mosquitoes carry the West Nile virus" or "sharks attack swimmers", properties considered dangerous to human kind are contained and they, furthermore, only pertain to "*good predictor[s]* of the property in question" (Leslie, 2008, p. 41). From an evolutionary perspective, this seems plausible: Even if only a small minority possesses a harmful property, this would have been a useful indicator to our ancestors and helped them to react appropriately to potential threats. At the same time, generalizations must be informative. They only hold true when other kind members at least have the predisposition to satisfy the dangerous property and if categories are not too broad or irrelevant (e.g. "animals attack swimmers", "insects carry the West Nile virus", Leslie, 2008, p. 41). Cimpian, Brandone, and Gelman (2010) tested this hypothesis more closely and found that, indeed, people were more likely to accept features which are dangerous or distinctive as the general characteristics of a kind, even when they had low prevalence. Such a cognitive principle that broadens potential danger to entire categories may likewise be reflected in antisemitic metaphors which generally invoke narratives of bodily harm or subjection.

Overall, tracing the interpretation of the generic operator GEN back to a basic cognitive mechanism comes with important advantages: By assuming that the mechanism adheres to general psychological principles, we can explain generics which prove troublesome for most formal approaches (e.g. "ducks lay eggs" vs. "ducks are female", "sharks attack swimmers"). We can furthermore account for the acquisitional pattern: A child only needs to learn "how her language partitions material into Restrictor and [Matrix], and the meaning of the nonquantificational material therein" (Leslie, 2008, p. 23). The meaning of GEN, on the other hand, is provided by means of a cognitive default. In view of this, it is also evident why GEN is unarticulated in the languages of the world: Since the least marked surface forms seem to correspond to default interpretations in general, generic generalizations are predicted to be unmarked (Leslie, 2007, 2008). Quantifiers, by contrast, require additional computational processing in order to relate the restrictor and matrix to one another. The fact that quantifiers are universally marked can thus be understood as "supplying explicit instructions to the evaluating conceptual system" (Leslie, 2008, p. 23) to perform additional computations. The

Generics-as-Default, furthermore, receives broad support from experimental research: As previously mentioned, Graham et al. (2011) found that 30-month-olds generalized a demonstrated property to a dissimilar exemplar of a kind if the demonstration was accompanied by a generic sentence. They preferred the exemplar from the demonstration in the case of a non-generic sentence. When Graham et al. (2011) conducted the experiment with 24-month-olds, they observed that the infants equally mimicked the property for the dissimilar exemplar in both conditions. In other words, they generalized the property to other members of a kind, irrespective of the linguistic input. The authors concluded that generic interpretations may step in by default when tasks are cognitively too demanding. Hollander, Gelman, and Star (2002) made a similar observation: In their study, they asked adults, three and four-year-olds about the properties of a kind whereby the questions either contained the quantifiers *some* or *all* (e.g. "Do all/some girls have curly hair?"), or were phrased as generics ("Do girls have curly hair", Hollander et al., 2002, p. 887). The results revealed a significant developmental change: When the property in question had narrow scope, both four-year-olds and adults were more inclined to affirm the generic question rather than the "all" question and they answered significantly more often with "yes" to the "some" question as compared to the generic question. The three-year-olds, on the other hand, did not distinguish between conditions. However, while the response pattern for quantifiers significantly changed with age, the generic questions were "statistically indistinguishable from those of 4-year-olds and adults" (Hollander et al., 2002, p. 893). Given that the three-year-old infants were able to distinguish between different quantifiers and generics in simple post-tests, the authors assumed that the cognitive load of the task was presumably too high and that, in such cases, infants invoke a generic default interpretation. In a follow-up study, Gelman, Leslie, Was, and Koch (2015) reduced the cognitive load by providing explicit statistical information about the property in question. When they presented pictures of different frequency levels of a property (e.g. three out of four "wugs" with big hair), even the three-year-olds distinguished between generics and quantifiers. Interestingly, children categorically interpreted quantifiers as generics when they made mistakes, so that they affirmed an "all" question, though not all category members satisfied the relevant property in the picture. Accordingly, even in less demanding tasks, "children have a default tendency to revert to generics" (Gelman et al., 2015, p. 458). This tendency was also found in adults: In Leslie, Khemlani, and Glucksberg's (2011) study, adults tended to agree with false universal statements when the corresponding generic was true (e.g. "all ducks lay eggs", see Leslie et al., 2011, p. 22). Participants did so in spite of rejecting statements such as "male ducks lay eggs" or accepting true alternatives such as "some ducks do not lay eggs" in a subsequent knowledge

test. In other words, even when participants were aware of exceptions, they accepted such universals in almost a third of the time, indicating that they followed a more basic interpretation in such cases, that is to say, a generic interpretation. Leslie and Gelman (2012) further investigated the Generics-as-Default-Hypothesis by means of a semantic memory test: In their study, they presented pictures of animals to adults and children and shared novel facts about the animals by use of generics or quantified statements (e.g. "(All/most/some) cats sweat through their paws", see Leslie & Gelman, 2012, p. 190). After completing a distracting task, participants had to recollect the novel information. An error analysis revealed that both adults and children frequently recalled "all" and "most" statements as generics, even when the generic statements were excluded from the presentation phase to avoid priming effects (c.f. Leslie, Khemlani, and Glucksberg, 2011 who found this tendency for paraphrases in a subexperiment). Children at the age of three also recollected the existential "some" statements as generics. The authors suggest that systematic errors arise because quantified statements are "cognitively more sophisticated" (Leslie & Gelman, 2012, p. 209), while generics express the cognitive system's default generalizations. Finally, evidence derives from a sentence processing study in adults (Meyer, Gelman, & Stilwell, 2011): After reading generic and quantificational sentences, participants judged whether a property was indeed characteristic of a kind whereby half of them were asked to do so as quickly as possible. Results revealed that, only under time pressure, participants endorsed universal statements about properties which are not true for all category members (e.g. "All dogs have four legs", see Meyer et al., 2011, p. 915). In fact, they rated such statements in line with generic generalizations. Furthermore, reaction times were delayed in the speeded conditions when people correctly falsified false universals as compared to verifying true generics. Participants were moreover significantly slower when they verified true universals (e.g. "All giraffes have long necks", see Meyer et al., 2011, p. 915) as compared to their true generic counterparts (e.g. "Giraffes have long necks", see Meyer et al., 2011, p. 915). This was the case in both the speed and non-speed condition so that, overall, generic generalizations were more accessible than the set-theoretical quantifiers. In summary, differences in their acquisition, distribution and processing indicate that generics are very different from explicit quantifiers. Rather than possessing quantificational meaning, generic statements seem to reflect a basic mode of generalization which is evident early on in childhood and tends to be exploited by default, particularly when the level of cognitive demand is high, or when generalizations cannot be checked against frequency specifications.

3.1.2.2. Generics and the Constraints on Social Categories

Considering that generics express a default mode of generalization that is central to the acquisition of general knowledge, we can assume that they also provide the content of a social concept during its formation. Simply put, generic generalizations are presumably stored within a category's mental representation as the properties associated to a particular social group. However, generic language contributes more than mere content to social cognition. Generics additionally influence the structure of conceptual representations, that is to say, they shape how properties are connected to a kind within its mental representation. The following chapters will take a closer look on the interplay between generics and the constraints on social concepts. In that sense, generics increase the likelihood of social categories to become essentialized, be the subject of inherence-based reasoning or bear more principled connections to their respective properties.

3.1.2.2.1. The Role of Generics in Psychological Essentialism

In her seminal work on psychological essentialism, Gelman (2003) argues that several domain-general predispositions give rise to essentialist beliefs about natural and social kinds (i.e. a tendency to distinguish between appearance and reality, to draw inferences from property clusters, to track identity over time and the belief in causal determinism). In that sense, essentialism is not an acquired set of beliefs which is extracted from linguistic input, but a capacity that develops by its own means (Gelman, 2003). This argument is particularly reasonable in light of the following findings: Gelman, Coley, Rosengren, Hartman, Pappas, and Keil (1998) analysed the content of child-directed speech during book reading sessions and found that mothers only rarely referred to the non-obvious properties of a category (e.g. insides, origin or kinship, teleological properties) or explained the mismatch between appearance and reality (e.g. "that looks like a kangaroo, but it's called an aardvark", see Gelman et al., 1998, p. 97). They further did not talk about essences so that essentialism seems to be "at least partially constructed by children themselves" (Gelman et al., 1998, p. 102). At the same time, some aspects of essentialist reasoning are already present at a preverbal stage of development: In Baldwin's et al. (1993) study, children already generalized non-obvious properties to objects of similar appearance at the age of nine months. Importantly, the study was conducted without additional linguistic cues: Children were introduced to a toy (e.g. a horn which honked when its bulb was squeezed) and then played with a second highly similar toy that was either capable to produce the respective property or not. When the property was absent, children tried to make

the toy work and performed the eliciting action significantly more often and with shorter latency as compared to the functioning toys (for similar findings see the first experiment in Graham et al., 2004). Even though findings suggest that psychological essentialism emerges independently from language due to more general cognitive biases, language may still contribute to the development of essentialist beliefs. According to Gelman (2003), linguistic input provides the critical information about which categories are to be essentialized in the first place. Two expressive functions of language are important in this respect, namely making reference to objects as members of a broader category and establishing the scope of a proposition. Importantly, "these two functions […] cannot be expressed nonlinguistically" (Gelman, 2003, p. 180), rendering essentialism an exclusively human phenomenon.[12] The expressive functions correspond to two linguistic devices, that is labels and generic noun phrases (Gelman & Roberts, 2017). Interestingly, this division reflects the distinction between D- and I-genericity in formal semantics: While labels are means to directly refer to kinds (D-genericity), generic NPs abstract away from facts or events and generalize their scope to many members of a category (I-genericity).

Let us first discuss findings regarding the influence of labels on essentialist reasoning: Graham et al. (2004) found that 13-month-olds inferred the non-obvious properties of a novel object (e.g. a rattle sound) to dissimilar objects when they were introduced by the same category label (e.g. "Look at this *blint*!", see Graham et al., 2004, p. 418). However, when labels remained absent or objects were presented with a divergent label, the infants were less likely to imitate eliciting actions. Similarly, Jaswal and Markman (2007, p. 96) found that 24-month-olds "give up a compelling, perceptually based classification in favor of an unexplained, label-based one." In the study, the researchers presented two familiar animal or artefact props to the children and enacted an associated property for each (e.g. a cat sipping milk and a dog chewing a bone). On request, the children performed the properties on a hybrid stimulus which was more similar to one of the two categories. The children were more likely to mimic the property of the dissimilar category when the hybrid was accompanied with the respective label (e.g. "dog" for a cat-like stimulus). Without labels, their choice was based on perceptual similarity. A similar labelling

[12] Animal model studies investigated this claim more closely: Cacchione, Hrubesch, Call, and Rakoczy (2016) found that apes successfully individuate objects and track their identity in terms of deeper properties, rather than superficial features: Apes engaged in a longer search when an object (i.e. banana) placed in a box turned out to be a different kind (i.e. carrot), as opposed to being superficially changed (i.e. an orange-dyed banana). They furthermore preferred a banana over a carrot slice after watching superficial transformations (i.e. coating one with the peel of the other). However, the study only tested rudimentary inferential skills which were, overall, rather fragile (i.e. only when working memory demands were low). Psychological essentialism may therefore be present in animal in a phylogenetically or ontogenetically rudimentary form, but not as a full-blown essentialist stance as illustrated in chapter 2.1.1. (see also Gelman, 2003, p. 283f.).

effect was obtained in a study with adults: Yamauchi and Markman (2000) showed that people generally inferred a missing property of an imaginary bug (e.g. horns, heads, tails) in line with the category label (inference task). They did so even when the stimulus was otherwise similar to a contrasting kind. When the missing feature was the label itself (categorization task), people were guided by similarity. The authors concluded that labels are not only features within a category's mental representation, but represent separate entities that specify "class inclusion relations" (Yamauchi & Markman, 2000, p. 792). Although findings indicate that labels – both familiar and novel ones – are central to category-based inferences throughout the human life span, it remains unclear what their actual role is as regards psychological essentialism. Some theorists have argued that essentialist beliefs develop from an overinterpretation of labels: For instance, Carey (1996, p. 193) assumes that essentialism builds on the single "default assumption that count nouns are substance sortals, that is, naming concepts that provide conditions of identity during the maximal life line of an entity". This assumption implies that labels are direct causes for representing a category as determined and unchanging in its identity. Obviously, there are labels for many different things, yet, not all of those categories are essentialized (e.g. "pet", "passenger" or artefacts, see Gelman, 2003, p. 184). In this respect, Carey and colleagues argue that it is only in the course of development that children learn about other types of sortals (e.g. phase sortals "pet", "passenger") or come to essentialize differently in different domains (recall the "design stance" for artefacts in chapter 2.1.1.2., Carey, 1996; Carey & Xu, 1999; Kelemen & Carey, 2007). Gelman (2003) takes another perspective on this matter. According to her view, labels are not the unique source of essentialism but rather direct the essentialist bias towards a target category due to their privileged status in category representation. As opposed to perceptual features, labels go beyond a superficial observational level and relate to a concept's deeper "level of explanation and cause" (Gelman, 2004, p. 407; Gelman & Roberts, 2017).

Regarding the second linguistic device (i.e. I-generics), Graham, Gelman, and Clarke (2016) found that children at the age of 30 months imitated an atypical property (e.g. drinking ketchup, eating flowers) more often for new members of a novel animal kind when both kind and property were introduced by means of a generic noun phrase. The study aimed to clarify whether generics in fact increase the inductive potential of a category or whether effects arise because non-generics restrict the scope of generalization (a possible interpretation of Graham et al.'s, 2011 findings). For this purpose, a non-generic condition and a baseline condition were included (e.g. "Look at this!", see Graham et al., 2016, p. 1356). While no significant difference was obtained between the baseline and the non-generic condition, children's imitation was

significantly more likely and more frequent in the generic condition. Previous findings thus did not arise due to the restricting effect of non-generics, but as a result of generic wording. A second experiment further showed that children did not randomly generalize to all objects available: The infants persistently inferred properties to members of the same category, but not to different category members. In another study, Gelman et al. (2010) examined "whether generics foster a more essentialized view of the target category" (Gelman et al., 2010, p. 275) that goes beyond inferring properties from a familiarization phase to other category members. In a first experiment, adults learned about a novel animal and its properties either by use of generic wording (e.g. "Zarpies like to sing"), specific labels (e.g. "'This zarpie like to sing") or no label (e.g. "This likes to sing", see Gelman et al., 2010, p. 278). Subsequently, they completed several inductive tasks about further non-obvious properties (i.e. novel induction task), the category's intrinsic predisposition (i.e. open explanation task) as well as the stability and hereditability of its properties (i.e. switched-at-birth task). Results showed that the category was essentialized more strongly in the generic wording condition. Interestingly, the effect was gradual so that the measures were significantly higher in the specific label condition as compared to the condition without labels. In a second experiment, Gelman et al. (2010) investigated the persistence of this effect in both adults and pre-schoolers: After two weeks of daily picture book sessions, children as well as their parents completed the induction tasks from the first experiment. They were tested a second time after two or three weeks. Despite the time gap, essentialist measures remained high in the generic condition for both adults and children, though the effect was weaker for children. In a third experiment, the generic input was reduced to a minimum so that the wording conditions pertained only to the task instructions and the test questions. Under these circumstances, only adults essentialized the category in the generic condition. Again, the effect was stepwise so that labels elicited significantly higher essentialist beliefs in adults as compared to the absence of labels. These findings indicate that children require more generic input in order to form an essentialist view of categories as opposed to adults who essentialized a novel category without further ado. The results additionally highlight the differences between labels and I-generics: Although hearing category labels did not encourage children, adults did essentialize category representations. In light of the discussion above, labels may not have the unique status for children as proposed by Carey (1996). As a matter of fact, I-generics seem to be even more central to the development of psychological essentialism.

So far, findings have illustrated the impact of generic language on essentialist reasoning in general. But how about its influence on essentializing social categories? Rhodes, Leslie, and

Tworek (2012) investigated this question more closely by using a similar study design as Gelman et al. (2010). In the study, adults and four-year-old children were presented with a storybook about a novel social category and their respective properties in three different wording conditions (i.e. generic, specific label, no label). Similar to previous findings, generic language fostered essentialist beliefs about the social category amongst both children and adults. Furthermore, labelling had a weak effect on adult's tendency to essentialize but not on children. This pattern was found both when the generic condition included bare plurals and indefinite singular generics, suggesting that essentialist effects were not merely the result of a singular/plural distinction. In their study, Baron, Dunham, Banaji, and Carey (2014) examined the effect of labels on the inductive potential of social categories: Children at the age of four and seven heard a story about individual members of a novel social category (e.g. "this *Lup*", "these *Nifs*") which either engaged in an antisocial or neutral action as a function of their category membership. When asked who would rather engage in further antisocial or prosocial behaviours ("This *Lup* or this *Nif?"*, Baron et al., 2014, p. 246), both age groups inferred negative behaviours significantly more often to the members of the negatively portrayed group. Older children, furthermore, made fewer positive behaviour attributions to the negative social group. In further experiments, visual cues (e.g. skin colour) indicated category membership instead of labels. In such cases, younger children failed to form a richly structured category representation and did not generalize antisocial behaviour to the target social category. These findings are insightful in two ways: Firstly, labels encouraged young children to expect negative properties in further category members, while positive behaviours were attributed by chance. This mirrors a cognitive principle which we have already encountered with respect to the quirky truth conditions of generics: According to Leslie (2007, 2008), the cognitive system is more prone to generalize harmful or dangerous properties to many members of a kind. While Leslie considers this bias to be an underlying principle of I-generics, the same principle may be implemented in children's early interpretation of specific labels. Secondly, labels fostered social categorization in young children, while visual information failed to do so. This is rather puzzling given that visual information is sufficient to form basic-level categories of animals from a very early age (for an outline see Diesendruck, 2003). The findings thus point to the fact that psychological essentialism has different manifestations across domains and that generic language may weigh differently along these lines (for domain differences regarding psychological essentialism, recall chapter 2.1.1.2): In other words, if outward perceptual information is inconclusive or ambiguous, psychological essentialism "may be more susceptible to language effects" (Gelman, 2003, p. 191).

In summary, generic language seems to guide the attention of the essentialist bias towards certain categories. People are more inclined to infer a property to other category members when they hear labels or generic characterizations, they suspect further shared properties between them and develop increased essentialist beliefs regarding a category's origin and predisposition. However, generic language only boosts essentialist conceptions. Psychological essentialism originally derives from language-independent biases which evolve into different assumptions in different domains. Such domain differences are also reflected in the impact of generic language on essentialist beliefs (Diesendruck, 2003): For the biological domain, assumptions about kinds and essences are most prevalent, cross-culturally coherent and biological categories are readily essentialized without further language input. Essentialist beliefs regarding artefacts and social kinds, on the other hand, are more varied and dependent on "their particular [personal or] cultural significance", respectively (Diesendruck et al., 2013, p. 1916; Wodak et al., 2015). In that sense, artificial and social essentialism relies more heavily on linguistic cues which define "the classes of entities that constitute [an artefact or] a culturally relevant kind" (Diesendruck, 2003, p. 777).

3.1.2.2.2. Generic Language and Inherence-Based Explanations

People tend to explain patterns in the world by reference to their inherent features, that is to say, the observed properties of an individual or category member are often attributed to its internal causal features. In chapter 2.1.2., the so-called Inherence Heuristic has been introduced as the underlying reasoning mechanism for such intuitive explanations. The heuristic mechanism identifies the components of a given pattern, retrieves their most accessible inherent features from semantic memory and assembles them either in a causal or normative way (Cimpian, 2015; Cimpian & Salomon, 2014). As a result, the output explanation identifies inherent features either as causal or as making a category member most optimal to possess the properties in question.

The impact of generic language on people's intuitive explanations has been targeted in various studies: Cimpian and Markman (2009) asked four- and five-year-olds to explain a novel property of an animal kind whereby the information was either presented in a generic format (e.g. "Why do snakes have holes in their teeth?") or in a non-generic way (e.g. "Why does this snake have holes in their teeth?", Cimpian & Markman, 2009, p. 17). In the generic condition, children produced significantly more explanations where another feature was caused by the property in question (e.g. "So they can chew food better", Cimpian & Markman, 2009, p. 18)

and fewer explanations where the property was the result of a prior behaviour, feature or event (e.g. “Maybe because a bug came in its room, and it bited his teeth” Cimpian & Markman, 2009, p. 18). In other words, they were more likely to interpret a novel property as an inherent causal feature rather than a specific effect. Interestingly, generic and non-generic wording did not make a difference when familiar properties were presented. This suggests that generic language is particularly influential when people learn new information or seek to explain novel correlation patterns in the world. Similar results were obtained for artificial objects: In their study, Cimpian and Cadena (2010) showed unfamiliar artefacts to children between four and five years of age and told them about their properties either by use of a generic or a non-generic sentence (e.g. “Dunkels are/This dunkel is sticky”, see Cimpian & Cadena, 2010, p. 63). In their explanations, older children referred significantly more often to the functional purpose of the respective property in the generic condition (e.g. “Because they’re supposed to stick to things”, Cimpian & Cadena, 2010, p. 65). The children were furthermore asked whether the object was made this way or whether something happened to it. When introduced generically, only older children were more likely to attribute the property to creator’s intent – a notion that corresponds to the essential features in biological and social kinds (recall chapters 2.1.1.2. and 2.1.3.3. regarding differences between domains). The authors concluded that the age effect was either the result of different trajectories of the explanatory stances involved (i.e. a design stance develops later than an causal essentialist bias), differences in the linguistic input (i.e. generics are more often in use for natural kinds than for artefacts, see e.g. Gelman et al., 1998; Goldin-Meadow et al., 2005) or due to the specifics of the task (i.e. novel objects vs. familiar animal kinds). In any case, causal explanations of novel facts seem to be influenced by generic language to some extent in both the biological and artificial domain at an early developmental stage. A third study investigated how generics inform the interpretation of novel properties of social categories (Cimpian & Markman, 2011): When abilities or biological traits were introduced by means of a generic sentence (e.g. "Girls are really good at a game called 'gorp'", “Boys have something called 'thromboxane' in their brains”, Cimpian & Markman, 2011, p. 475), both adults and children from four to five explained them more often with reference to inherent features (i.e. as a result of development, preference, trait, or as an intrinsic function). In the non-generic condition, the explanations were more likely to invoke accidents, diseases, training, exercise, or other external causes. Interestingly, this effect was dependent upon the nature of the category and property type: Inherence-based explanations were more frequent only in cases where social categories were “meaningful and coherent” (as opposed to context-based e.g. “boys/girls at a school”, Cimpian & Markman, 2011, pp. 487f.). Moreover, only

adults distinguished between generic and non-generic wording when the property described a lack of ability (e.g. "Girls *aren't* really good at a kind of game called 'hep'", Cimpian & Markman, 2011, p. 485). These findings may relate to a broader cognitive principle as regards generic generalizations: As proposed by Leslie (2007, 2008), the conceptual system is biased towards information where counterexamples do not possess a positive alternative property (e.g. "ducks lay eggs", recall chapter 3.1.2.1.). This principle might be more pronounced in children so that generalizations with positive counterexamples are generally not accepted by children and, in consequence, not attributed to inherent causes. Cimpian and Erickson's (2012) argumentation proceeds along similar lines: According to their view, children's causal explanations are not merely sensitive towards the difference between generic or non-generic statements. The key issue here is "whether the *beliefs* relevant to these statements are generic or nongeneric" (Cimpian & Erickson, 2012, p. 160). In other words, whether the generic generalization is believed to be valid is more central to causal attributions than the linguistic form per se. Cimpian and Erickson (2012) tested this hypothesis in their study with adults and children between four and seven: In a first step, they introduced a novel gender property by means of a generic statement, a non-generic sentence or by use of the quantifier "most". In a second step, participants were presented with a single child whose gender matched the introduction they had received in the first step. Participants were told that the child likewise possessed the property and asked to explain this correlation (e.g. "Why is this boy/girl really good at this game called 'tooki'?", Cimpian & Erickson, 2012, p. 164). The rationale of the study was the following: If inherence-based explanations are primarily sensitive to the genericity of beliefs in lieu of statements, participants would extend their causal attributions from the introduction phase to the subsequent non-generic context. Indeed, when properties were previously presented in a generic format, older children and adults understood the non-generic fact "through the lens of this belief" (Cimpian & Erickson, 2012, p. 161) and referred significantly more often to normal development, some intrinsic preference, trait or function than in the case of a non-generic introduction. This effect was only approaching statistical significance for younger children. However, all age groups produced fewer explanations with reference to accidents, diseases, exercise, or other external causes in the generic condition. Overall, these findings indicate that priorly formed generic beliefs are important to people's causal explanations.[13] Similar to essentialist beliefs, generics do not blindly prompt inherence-

[13] Interestingly, the study further revealed that generics are not the single source for inherence-based beliefs: The explanation pattern in the quantificational condition was similar to the generic condition which may be explained in two ways: Either the wide-scope quantifier "most" invoked generic beliefs about gender categories in its own way or, in line with former observations, the quantificational statements were in fact recalled as generic

based explanations. It is rather in combination with general cognitive biases and kind assumptions that they draw the attention to a property's inherent causes. Yet, if a generic statement is consistent with generic beliefs about categories and kinds, they may highlight the possibility that a property is caused by the category's deeper, inherent features.

The experimental findings which have been outlined above exclusively relate to I-generics and their influence on intuitive explanations. Another question to be answered is how D-generics and labels may inform inherence-based reasoning. In this regard, let us return to Cimpian and Salomon's (2014) assumption that the Inherence Heuristic is the underlying reasoning mechanism for a variety of psychological phenomena, including the belief that words reflect the intrinsic features of their referents (i.e. nominal realism, recall chapter 2.1.2.2.). According to nominal realism, there is a unique, true name for each natural kind (Gelman, 2003). Labels are therefore conceived as more than mere conventional form-referent-pairings, but as providers of objective cues to a category's deeper, intrinsic nature. Though nominal fit beliefs are predominantly found in young children (Benelli, 1988; Brook, 1970), some more subtle forms may also be preserved in adulthood: In their study, Sutherland and Cimpian (2015) found that 66.7 percent of participants produced a nominal fit explanation at least once (e.g. "*Yaroo* doesn't give a sound like 'chair', to imply what you do with it", see Sutherland & Cimpian, 2015, p. 233). In sum, people seek to explain linguistic patterns (i.e. the use of labels) by reference to intrinsic semantic features throughout the lifespan. In the case of social category labels, their use may incite people to ascribe an underlying truth to a labelled social group and direct the attention to their inherent causal properties.

3.1.2.2.3. Generics and the Aspect Hypothesis

While psychological essentialism and the Inherence Heuristic constrain social concepts to the effect that particular properties are more central to a conceptual representation and, thus, foregrounded in intuitive reasoning processes (i.e. as essential and inherent causes), the Aspect Hypothesis takes a different focus. Rather than emphasizing the role of certain features in belief systems and intuitive explanations as well as the consequences thereof, the question at issue is **how** the properties of a kind are generally represented by the conceptual system. According to Prasada and Dillingham (2006, 2009), there are at least two different types of connections: Firstly, statistical connections link properties to a category by means of frequencies. Such

generalizations (see e.g. Leslie & Gelman, 2012). These findings, thus, may provide further evidence for the assumption that generics express a default mode of generalization (recall chapter 3.1.2.1.).

properties are therefore determined by their prevalence in the kind's token, not by the category itself (recall chapter 2.1.3.1.). Principled connections, on the other hand, directly relate a property to a kind as one of its aspects. Broadly speaking, "there is an asymmetry between the two" (Prasada et al., 2013, p. 206): When kinds and properties involve principled connections, they are likewise related to one another via statistical connections. The same does not apply vice versa. In light of this, both statistical and principled connections promote statistical expectations regarding a property's prevalence, but it is only for principled connections that they license normative beliefs and formal explanations. Either way, both connection types invoke expectancies and inferential reasoning as regards the features of a category and its members.

Hollander, Gelman, and Raman (2009) found that generic wording may convey such kind-property-connections and that they are consulted when pre-schoolers and adults identify new members of a kind (i.e. a categorization task). In their study, a drawing of a novel kind and a particular property were presented either generically or non-generically (e.g. "This is a KEVTA. Let me tell you something about kevtas [/this kevta]. Kevtas are [/this kevta is] woolly", Hollander et al., 2009, p. 504). Two further pictures were presented whereby one depicted an instance possessing the respective property, and the other one showed an instance matching the kind in appearance or shape. When asked which of the two instances belonged to the novel category, all age groups selected the pictures with the matching property significantly more often in the generic trials as compared to the non-generic trials. In the non-generic condition, adults selected property matches below (or at) chance level, while children chose them significantly more often. In this respect, the authors argue in favour of a developmental change so that only adults tend to interpret specific labels "as implying *lack* of generality" (Hollander et al., 2009, p. 498). Interestingly enough, all age groups were significantly more likely to select the matching property in both conditions as compared to a control condition which did not pronounce the property altogether. Given that both generic and non-generic introductions alluded to the depicted property, mere reference seems to be sufficient for linking a property to a kind. However, generics promote this basic tendency more vigorously: Both adults and children shared strong expectations about the property and its prevalence in other members of a kind, whereas in the non-generic condition, only children were more likely to make this inference. Unfortunately, the study does not cast light on the exact nature of this kind-property-linkage. The question whether generics in fact establish principled or statistical connections was investigated more closely by Prasada et al. (2013): In their study, they showed that different types of generics involve different types of connections. Participants judged

several predication types[14] in combination with assertions that are supported by principled connections (e.g. "is one aspect of", "by virtue of", "are supposed to", "all normal" statements, see Prasada et al., 2013, p. 409). While majority and minority characteristics received overall higher ratings, majority statistical statements and striking property were rated rather low on these measures. In a second experiment, Prasada et al. (2013) examined which of the underlying property-kind-connections would explicitly support formal explanations: Participants rated whether formal explanations sounded naturally for different generic predication types (e.g. "Why does that (pointing to a tiger) have stripes? Because it is a tiger, and having stripes is one aspect of being a tiger", Prasada et al., 2013, p. 414). Again, formal explanations for majority and minority characteristics were broadly accepted, whereas explanations for striking properties patterned similar to those of majority statistical facts and were rated unacceptable. Based on these results, the authors concluded that only the first two generic types in fact involve principled connections. In order to assess differences between majority statistical statements and striking property generics, a third experiment explored the role of cause and disposition in people's conception. It was only for striking properties that participants affirmed causal explanations in the form of "There is something about Ks that causes them to P" (Prasada et al., 2013, p. 415), suggesting that a third type may relate properties to a kind (i.e. causal connections). Majority statistical generics, on the other hand, seem to exclusively involve statistical connections. By and large, what can we draw from such findings? First of all, generics may establish and highlight connections between a kind and its central properties. However, they are not the single source for such links: Hollander et al. (2009) found that, at least for pre-schoolers, non-generic reference to a property had a similar, yet significantly weaker effect on their conceptualizations. Furthermore, generics do not establish property-kind-connections in a unified way: As suggested in formal semantic literature (see e.g. Carlson, 2019; Pelletier, 2010), generics are by no means a coherent phenomenon. They may in fact involve different conceptual links between a property and a kind. Such connections may be statistical, principled or causal in nature and, above all, yield different inferential patterns (Prasada et al., 2013). Simply put, generics may reflect different types of generalizations and, by extension, different modes of thought "when thinking and reasoning about kinds of things" (Prasada et al., 2013, p.

[14] The generic predication types included majority characteristic generics (i.e. a property is prevalent in most members of a kind, e.g. "Tigers are striped"), minority characteristic generics (i.e. a property is prevalent in less than half of all kind members, yet central or essential, e.g. "Lions have manes"), majority statistical generics (i.e. properties with high prevalence, yet not characteristic to a kind, e.g. "cars have radios") and striking property generics (i.e. dangerous or harmful properties that are only true for a small minority of all members of a kind, e.g. "Pit bulls maul children", Prasada et al., 2013, p. 409). Importantly, all predication types can be expressed by means of a generic bare plural and, as such, were accepted as true statements in the trials.

418): While striking property generics rely on causal connections and, thus, on the essentialist stance which yields generalizations with reference to an underlying cause, majority statistical generics may reflect a statistical mode of thinking whereby the kind-property-link is mediated by statistical principles (see Prasada, 2010; Prasada et al., 2013). Majority and minority characteristics, on the other hand, rely on principled connections which invoke the so-called formal mode of explanation. Such generalizations are generally grounded in the conception that properties are true for kind members by virtue of being the kind of thing they are (Prasada & Dillingham, 2006, 2009). In sum, the conceptual system generalizes facts to kinds in different ways and these differences may precipitate in generic language (Prasada et al., 2013). As opposed to Leslie (2007, 2008) who argues for a single cognitive mechanism, the divergent truth conditions of generics may furthermore result from different types of connections and different modes of generalizations. Importantly, connection types are not mutually exclusive and can be expressed in a single generic statement. In any case, generics – in virtue of underlying kind-property-connections – employ the cognitive system's most basic reasoning mechanisms (Gelman, 2010; Leslie, 2007, 2008; Prasada et al., 2013), that is a statistical, an essentialist and a formal mode of explanation.

3.1.3. Summary and Discussion

Generics constitute a puzzling phenomenon of human language: They express generalizations about the state of the world, relate to regularities and laws, and describe habits and dispositions while, at the same time, conceding exceptions to different extents (Carlson, 2019). For many years, philosophers and semanticists have mulled over an adequate formal analysis of generic language and, hitherto, there is great disagreement on the question what generic sentences truly mean and how their quirky truth conditions come into existence. One issue has been settled thus far: As regards their representational form, theorists generally agree on the distinction between two types of genericity (see e.g. Carlson, 2019; Krifka et al., 1995; Leslie & Lerner, 2016): So-called D-generics are considered to be instances of direct kind-predication so that a property is directly attributed to a kind, not to particular members. I-generics, on the other hand, receive their meaning from a covert unselective operator GEN which binds the free variables of sentences and predicates a property to them. GEN is added to the logical form of a sentence by default whenever no alternative operator is available (Leslie, 2008). Despite agreement on this point, the interpretation of the generic operator is still under debate: One perspective taken on this matter aligns GEN with generalized quantifiers (Cohen, 2002; Collins, 2018). While the

classification of generics as quantificational statements comes with the advantage of applying regular set-theoretic tools, such approaches generally struggle with the question what the appropriate kind of quantification is. In other words, how many instances of a kind need to satisfy a property in question for a generic statement to be true and, over and above, how do we come to know the right amount of those instances (i.e. the problem of induction, see Carlson, 1995)? A second line of approach considers GEN as an operator which is evaluated relative to the order of the world. In that sense, the meaning of GEN is grounded in the ontological claim that there are underlying rules and regulations which are irreducible entities of the world and determine its state and structure. Rather than on the relation between a property and the instantiations of a kind, the truth of a generic sentence therefore depends on whether such rules are in effect (Cohen, 2002). Rules-and-regulations-approaches, however, only provide little insight into the truth conditions of generics: A generic statement is true if there is a corresponding rule in the world, and false if there is none. Apart from this, they do not provide tools to specify the exact nature of such regulations.

In recent years, cognitive scientists have entered into the discussion on generics for the following reasons: For one thing, it has been noted that generic knowledge is largely expressed by generic statements so that their evaluation may offer a window into general knowledge representation (Carlson, 2010; Prasada, 2010; Prasada et al., 2013). Additionally, generics do not pose difficulties for language learners, in spite of their complex semantics and the lack of overt generic marking. In view of this striking paradox, the so-called Generics-as-Default-Hypothesis proposes that generics corresponds to a basic learning mechanism which yields generalizations by default (Leslie, 2007, 2008). Generic generalizations do not fall into the scope of a semantic theory, nor do they correspond to the ontological structure of the world. In truth, they result from one of the most primitive mechanisms of general cognition. According to the Generics-as-Default-Hypothesis, the basic generalizing mechanism is a low-level cognitive process which operates as quick and effective as possible. The quirky truth conditions of generics are a result thereof, in so far as the mechanism is biased towards certain information: For one thing, the congnitive system is more prone to generalize along the *characteristic dimensions* of a kind (i.e. locomotion, diet, noises, reproduction, gestation, or – in the case of artificial and social kinds – function and role, see Leslie, 2008). Secondly, generalizations gravitate towards information where counterexamples lack the respective property, instead of possessing a vivid alternative. Thirdly, for evolutionary reasons, properties which are repellent or dangerous to humans are more likely to be generalized to an entire kind, even if only a small minority satisfies the property in question. Assuming a cognitive mechanism that underlies

generic statements solves a number of problems and paradoxes: First and foremost, the generic truth conditions are easily captured by use of a few general cognitive principles. Additionally, given that the underlying mechanism is considered to be innate, primitive and the basis of inferential knowledge acquisition, it is only logical that children should use and understand generic language at an early developmental stage. It is furthermore explicable why GEN remains unpronounced cross-linguistically: The Generics-as-Default-Hypothesis posits a default mode of generalization and since default interpretations are generally assumed to correspond to the least marked surface forms, unmarkedness is a predicted consequence (Leslie, 2007, 2008). Most importantly, however, a default mode of generalization may ilucidate why genericity is a powerful means of social cognition: Generic statements seem to provide the content of social concepts during their acquisition by relating to one of the cognitive system's most basic mechanisms of thought. Considering that generics are understood early on in development (Hollander et al., 2002), readily accepted even when evidence is rather tenuous (Cimpian et al., 2010) and that category information is often recalled in the form of generics (Leslie & Gelman, 2012), such generalizations are presumably stored in a social category's mental representation without further ado. However, generics do not only establish which properties are generalized to an entire social group. They furthermore influence how conceptual representations are organized in the first place. As outlined in chapter 2, several constraints are implemented in the representation of social categories which give rise to highly structured concepts, sharp category boundaries as well as stable and immutable properties across category members. In a broader sense, such constraints fuel the inductive potential of a social category and may induce stereotypes and prejudiced attitudes. Generics have been found to make a large contribution to some of those constraints: People generally hold stronger essentialist beliefs about a category when faced with generic statements (or, to a smaller extent, labels), they are more inclined to explain a property in question with reference to the category's inherent features and they set up different kind-property-connections, all of which represent the predicated property as more central and informative to a kind's concept. At large, generics highlight which social category is to be essentialized in the first place. They draw people's attention towards the possibility that a particular social group may be determined by some sort of internal cause – whether this is a single underlying essence in the case of an essentialist belief system, or different inherent properties with respect to inherence-based explanations (recall chapter 2.1.2.3. for the descent relationship between Inherence Heuristic and psychological essentialism). As for the representation of the predicated property, generics establish formal, statistical or causal connections to a social kind. They thus support a social category's

homogeneity by representing a predicated property either as an aspect of a kind (i.e. majority and minority characteristic generics), as caused by its inherent nature (i.e. striking property generics) or as highly prevalent in its tokens (majority statistical statements).

Ultimately, chapter 3.1. has so far illustrated how generics contribute to various aspects of general and social cognition. Where does this leave us regarding the overarching subject of antisemitic metaphors? In virtue of their generic component, let us assume that antisemitic metaphors have similar effects on people's social category representations: Firstly, they convey negative property attributions as regards their explicit content so that the predicated properties are more readily associated with the representation of a Jewish social category. This assumption is reinforced by the fact that, when generalizing to an entire kind, the cognitive system is particularly biased towards negative information or properties without a positive alternative. Given that the properties in antisemitic metaphors generally lack an alternative property and exclusively entail harmful predicates, they are most likely to thrive from these general cognitive principles. Furthermore, generics are "often accepted on the basis of scant evidence […] and, once accepted, they are resistant to counterevidence" (Cimpian & Markman, 2011, p. 427). In that sense, people do not seek confirmation before adding generic content to their mental representations. All in all, due to the generic component, people may be more inclined to accept the generalizations conveyed by antisemitic metaphors and expect the predicated property in most members of a Jewish category. However, we must be careful not to overestimate such effects: Conforming to egalitarian and pro-social ideals, people may be capable of countering explicit negative associations by use of self-regulatory processes and mechanisms of cognitive control (Amodio, 2014). A second effect of genericity may therefore be more central to the pervasive impact of antisemitic metaphors: Their generic nature shapes the "structure [of] our social cognition" (Wodak et al., 2015, p. 626). As generics, antisemitic metaphors boost those computational and representational constraints which highlight the intrinsic features of a Jewish social category. In consequence, people represent such a category as homogeneous regarding its internal properties and assume sharp and impermeable category boundaries towards other social groups. They assume further shared features on the basis of an essentialized representation, tend to explain the alleged properties and characteristics of a Jewish social group by reference to inherent features or represent the predicated properties as stable causes or categorical aspects. In light of this, antisemitic metaphors are straightforward vehicles for prejudices and further negative downstream consequences, not only because they impart false and offensive conceptions about Jews but also by means of their influence on the structure of a Jewish social concept.

3.2. Metaphors

The overarching aim of this master's thesis is to illustrate the impact of antisemitic metaphors on aspects of social categorization. While, thus far, we have established how such metaphors may influence the conceptual content and structure of a Jewish social category in terms of a generic component, let us now turn towards their effects as metaphors: Various theories have claimed that metaphor comprehension calls forth an external structure on the basis of which people attribute properties to a category or draw inferences about its relations towards other structural elements. Against this backdrop, I propose that antisemitic metaphors shape mental representations of a Jewish social category by highlighting external influences on the category's characteristics as well as its structural positions in the social hierarchy. In the following chapter, we will take a closer look at this proposal. More precisely, we will examine in which ways the underlying comprehension processes of metaphors additionally spark off the externalist constraints on social category representations. Bearing in mind that externalist constraints have similar effects as the aforementioned internalist constraints, antisemitic metaphors thus provide a further impetus which gives rise to a sharp and highly structured Jewish social category. As such, antisemitic metaphors reinforce their potential to further negative downstream consequences of social categorization.

In advance, it should be mentioned that research on metaphorical language is exceptionally vast and diverse. For that reason, the next chapter will begin with a general overview over different approaches "to an understanding of the nature, function, and meaning of metaphor" (Gibbs, 1992, p. 576). The focus will then shift to theories that explicitly describe underlying computational processes of metaphor comprehension. Importantly, a large portion of the experimental work on metaphor comprehension focuses on the question which theory captures most adequately the characteristics and time course of metaphors, that is, which of the theories holds true. Since answering this question reaches far beyond the scope of this thesis, the following chapter will concentrate on the theoretical literature and discuss how each of the proposed comprehension mechanisms can be linked to the externalist constraints on social concepts.

3.2.1. Preliminaries

In recent years, metaphors have caused a stir of interest amongst cognitive scientists (Bowdle & Gentner, 2005). While traditionally viewed as ornamental figures of speech, both rare and opulent in nature, this conception largely changed due to an increasing body of linguistic and psychological evidence: Metaphors are frequently used when people communicate about rather abstract concepts such as space, time and emotions (e.g. Lakoff & Johnson, 1980a; Ortony & Fainsilber, 1987), they seem to be involved in creative thinking (Cacciari, Levorato, & Cicogna, 1997) and, furthermore, constitute a fundamental part of scientific discourse (e.g. Gentner & Grudin, 1985; Kuhn, 1993). At present, there is large agreement that metaphors do not only concern the realm of poetry, but in fact help to illustrate, or even organize, more abstract conceptual knowledge (Thibodeau, Matlock, & Flusberg, 2019).

Let us first begin with some specifications (Holyoak & Stamenkovic, 2018): In general, metaphors have an underlying two-part structure "X is a Y", whereby X relates to the subject of discussion and Y to a concept characterizing the subject matter more closely. The first component of the formula is called the *target*, *topic* or *tenor*, while Y is termed the *source*, *focus*, *vehicle* or the *base* of a metaphor. Importantly, X and Y are semantically different so that metaphors establish a relation "between concepts from disparate domains of knowledge" (Bowdle & Gentner, 2005, p. 193). In technical terms, this relation is labelled a metaphorical *mapping*. Metaphors may occur in different syntactic constructions: *Nominal* metaphors correspond to the underlying two-part structure so that the source is expressed by means of a noun phrase (e.g. "The stock is a *rollercoaster*", Holyoak & Stamenkovic, 2018, p. 644). In *predicate* or *attributive* metaphors, a verb or an adjective provides the metaphoric source, respectively (e.g. "The flower *purred* in the sunshine", "the *weary* mountain", Holyoak & Stamenkovic, 2018, p. 644). As for *proportional* metaphors, the source NP is related to an additional term of the target domain (e.g. "Religion is the *opium* of the people", see Holyoak & Stamenkovic, 2018, p. 644). Furthermore, some theorists consider locative prepositions with a conventionalized meaning as metaphorical sources (e.g. "I'm feeling *up*", "He's *under* hypnosis", Lakoff & Johnson, 1980a, p. 15). It should be noted that, on the whole, conventionality is an important dimension of figurative language (Holyoak & Stamenkovic, 2018): Metaphors are novel when they are first encountered or low in frequency. Yet with repeated exposure, metaphorical extensions may change into a part of the source's literal meaning. Consider, for instance, the focal word "leech" in an antisemitic metaphor. In addition to its biological denotation, the Oxford English Dictionary specifies a leech as someone "who

'sticks to' another for the purpose of getting gain out of him" ("leech, n.2," n.d.).[15] Furthermore, metaphors may become the primary meaning of a source term, thus shifting from being a subject of psychological studies towards etymology: As for many abstract words in English, origins reach back to an opaque metaphorical source (e.g. "matter" derives from the Latin root "mater" which translates to "mother", "subjective" stems from "subiectus" which means "lying under", see Holyoak & Stamenkovic, 2018, p. 644).

Which meanings do metaphors convey? Discussions on this issue date back to Aristotle and his first scholarly definition of metaphors as "giving the thing a name that belongs to something else; the transference being either from genus to species, or from species to genus, or from species to species, or on grounds of analogy" (Aristotle, 335 BC/1971, 1457b). According to Aristotle, metaphors act on the level of words whereby one is used in place of another (Gibbs, 1994). Furthermore, metaphors are to be understood as artful comparisons which draw the hearer's attention to hitherto unnoticed similarities between the source and the target word. In Aristotle's words, metaphors are "a sign of genius, since a good metaphor implies an intuitive perception of the similarity in dissimilars" (Aristotle, 335 BC/1971, 1459a). As opposed to overt comparisons, metaphors do not disclose which aspects of the two terms are actually compared. It is for the hearer to decide whether "Juliet is the sun" relates to her warmth and beauty or rather to representing the centre of Romeo's universe. In any case, the difference between explicit comparisons and metaphors is that only the latter allow for multiple interpretations. Over the course of history, the Aristotelian view remained largely unchallenged: Many theorists treated metaphors as hidden similes and focused on the identification of those marginal aspects of meaning which similarly hold true for the source and the target word (Gibbs, 1994). In recent decades, the so-called interactionist view has departed from this classical conception of metaphors: As stated by Max Black (1962), metaphors do not merely highlight similarities between pre-existing features, but rather create them during an interaction between the target and the source. The interactionist view therefore assumes an underlying generative process which yields novel, non-literal aspects of meaning. Importantly, the interactionist view furthermore differs from the traditional comparison account by assuming that interactions take place between concepts, rather than words. Since the result of such

[15] Note that the source term "parasite" also has a conventionalized social connotation, however, its etymology is more complicated: "Parasite" stems from the Ancient Greek word "parasitos" which originally denoted a companion at table (Musolff, 2014). The pejorative sense of a sycophant already developed in Antiquity but became predominant during the 16th century when the term was borrowed into other languages. In the 18th century, "parasite" was extended to the biological domain and received its harmful connotation (i.e. transmitting infections and diseases, death). Importantly, the biological meaning has been primary ever since and, as such, serves as a rich source for metaphorical inferences about the social domain.

conceptual interactions is a "system of commonplaces" (Black, 1962, p. 40), the interactionist perspective was able to account for a hitherto unexplained characteristic of metaphorical language (Gibbs, 1994): Metaphors are open-ended in their implications and cannot be paraphrased without a severe loss in meaning. Since the 1960ties, a variety of different perspectives on metaphors have been proposed: Davidson (1978), for instance, argued that metaphors shift the hearer's attention to a limitless number of similarities, yet these effects should not be mistaken for what metaphors actually mean. Much like jokes, metaphors are a function of language use so that they may intimate more, but they "mean what the words, in their most literal interpretation, mean, and nothing more" (Davidson, 1978, p. 32). Similarly, Sperber and Wilson (2012) rejected the notion of a special metaphorical meaning. According to their view, metaphors are mere instances of a broadened lexical meaning due to considerations of relevance. Given that such considerations generally lie at the heart of human communication, metaphors rely on the same cognitive principles of language use as literal language and, thus, are neither special nor fundamentally different. At the other extreme, Lakoff and Johnson (1980a, 1980b) have argued that metaphors are most crucial to human cognition as they reflect how the conceptual system is organized. Metaphors are therefore not merely a phenomenon of language but rather the elements by which the world is represented, understood and experienced. According to the conceptual metaphor theory (CMT), many of our mental representations are structured in terms of so-called conceptual metaphors. Conceptual metaphors map a familiar, concrete source domain onto a rather abstract, unfamiliar or complex target domain (e.g. LOVE IS A JOURNEY). In consequence, we systematically relate to deeper, more concrete concepts when we talk – or reason – about an abstract target domain (e.g. "We're headed in opposite directions", "Our marriage was on the rocks", Gibbs, 2011, p. 531). A further central claim of CMT is that, due to metaphors, conceptual structures are largely embodied (Gibbs, 2011; Thibodeau et al., 2019): Most conceptual metaphors directly refer to sensorimotor experiences or physiological functions (e.g. primary metaphors such as SIMILARITY IS PROXIMITY, EXISTENCE IS VISIBILITY, Grady, 1997, pp. 283f.) and others can be reduced to such embodied experiences (e.g. the complex metaphor LIFE IS A JOURNEY entails PURPOSES ARE DESTINATIONS and CIRCUMSTANCES ARE SURROUNDINGS, Grady, 1997, p. 112). Though considerably influential, conceptual metaphor theory was met with a lot of criticism (Gibbs, 2011; McGlone, 2007; Thibodeau et al., 2019): Firstly, CMT fails to explain why metaphorical mappings are partial, that is, why certain features of the source cannot be mapped onto the target domain. As for the example THEORIES ARE BUILDINGS, we may envisage references to a theory's foundation or

construction, yet it is rather obscure what the stairways or the sprinkler system of a theory would be (McGlone, 2007). Secondly, criticism concerns circularity: Given that CMT predominantly relies on linguistic evidence, it has been pointed out that conventional metaphors are both identified as the causes and the effects of their own influence. If metaphors, however, exceeded their manifestation in language by shaping conceptual structures, this effect should likewise be manifested in non-linguistic data (McGlone, 2007).

All in all, while many philosophical and linguistic theories have pushed ahead the discussion about the nature of metaphors, only few articulated the cognitive processes which underlie metaphors and their comprehension (Holyoak & Stamenkovic, 2018). In this regard, the following chapter will review the most important theoretical perspectives: Firstly, metaphors have been considered to involve a process of comparison between the features of the source and target concept. Secondly, the so-called categorization view aligns metaphors with class-inclusion statements so that the target is assigned to an abstraction of the source concept and inherits all of its properties. Thirdly, in line with the feature-matching account, the analogy view conceives metaphors as a comparison between concepts. The difference, however, is that the relational structure of both source and target concept is compared, rather than a set of features. Importantly, the categorization view and the analogy account both offer a new perspective on the question how antisemitic metaphors may influence social categorization. As for the categorization view, similar processes may be at hand when people identify a target concept as an instantiation of an abstract source concept and when they think structurally about social categories. Analogical reasoning, on the other hand, seems to be important for abstracting a common relational structure for social categories, that is a general social structure. By highlighting relational aspects of metaphor comprehension, both theories underpin the claim that antisemitic metaphors do not only introduce novel content to a societal concept but incite people to reason in a structural way about the characteristics and behaviours of a Jewish social group.

3.2.2. Different Views on Metaphor Comprehension

3.2.2.1. Metaphors and Similarity – A Process of Feature Matching

The Aristotelian view led to the traditional understanding of metaphors as implicit comparisons which guide the listener's attention to so far unnoticed similarities (Bowdle & Gentner, 2005; Glucksberg & Keysar, 1990). What does this view imply in terms of processing? First of all, given that metaphorical statements are literally false, it has been suggested that hearers initially need to reject the "defective" literal meaning (see e.g. Searle, 1993). Secondly, the sentence must be transformed into a corresponding simile (i.e. "X is *like* a Y"). In a third step, metaphor comprehension should proceed in accordance with the comprehension of literal comparison statements.[16] While the assumption of a three-step process was dismissed due to contradictory evidence – metaphorical meaning is neither derived subsequent to literal meaning (e.g. Ortony, Schallert, Reynolds, & Antos, 1978), nor does it seem to be optional and suppressible (e.g. Glucksberg, Gildea, & Bookin, 1982) – the alignment with literal comparison statements continued to be an influential conception in parallel processing models (Johnson & Malgady, 1979; Miller, 1993; Ortony, 1979; Tversky, 1977): Following Tversky (1977), comparisons are generally assessed by means of a **feature matching process**. During comparison, people compile a subset of relevant features for each term and compute similarity by contrasting the amount of common and distinctive features. Importantly, features are not weighted equally during this matching process. According to Tversky (1977), a variety of different factors renders features more salient and, thus, more decisive for the outcome of the comparison process: Factors may include constants such as frequency and familiarity, or the informational value of a property which varies across contexts (e.g. the feature "real" is weighted low in comparisons between actual animals, but it is rather salient when comparisons involve phoenixes or centaurs). Above all, the directionality of the sentence promotes a feature's salience so that the predicate of a comparison is generally more salient than the features of the subject term. This way, differential salience explains why different features are foregrounded in different comparison statements and, furthermore, why their reversed versions may highlight distinct features. Based on these considerations, Ortony (1979, p. 162) argued that differential

[16] To clarify, comparison statements will be defined and categorized as follows (Miller, 1993): First, any comparison statement entails a copula of similarity ("like", "as", "resembles", etc.). Second, it is only for literal comparisons that two concepts of the same semantic field are related. As for similes, the two concepts to be compared originate from different semantic domains. In this regard, similes are comparable with metaphors, the main difference being that only the first "explicitly cue comparison" (Holyoak & Stamenkovic, 2018, p. 652) by use of a linguistic expression. Analogies, as a third type of comparison statements, involve four terms whereby the relations between the first and the second two are compared (e.g. "The toes are to the foot as the fingers are to the hand", Miller, 1993, p. 372).

salience is also the "principal source of metaphoricity". While metaphorical comparisons rely on matching features that are less salient in the target than in the source concept (i.e. low X/high Y), literal comparison statements involve matching salience levels (i.e. high X/high Y). People thus recognize metaphors by identifying different degrees of salience between the features of a source and a target concept. Although the so-called salience imbalance theory received partial empirical confirmation (e.g. Katz, 1982 who found that low X/high Y metaphors are judged more felicitious than the reverse), it also faced severe challenges: Glucksberg and Keysar (1990), for instance, pointed out that literal comparisons may be high X/high Y matches to speakers, yet not necessarily to hearers. Given that both literal and metaphorical comparisons aim to be informative as regards less obvious similarities between a source and a target, the hearer should perceive an imbalance of salience for both types of statements. Differential saliency may therefore be central to the production, assessment or the verification of a metaphorical comparison, yet it fails to capture their unique characteristics during comprehension (for a similar point see Gentner & Clement, 1988 who solely found a relation between salience imbalence and aptness). Secondly, it has been noted that features of metaphors are often nonidentical (Bowdle & Gentner, 2005): For example, the metaphor "Man is wolf to man" conveys that both wolves and men are predatory, though in fairly different ways (i.e. social vs. carnivorous predation). Since feature matching models generally treat nonidentical features as distinctive, processing would require rendering both features similar to each other, and thus, "introduce[…] a recursive element into [the] account" (Ortony, 1979, p. 167). A third challenge concerns another central characteristic of metaphors: Often, metaphors do not only highlight features, but attribute novel properties to the target concept. For instance, prior to hearing that "Roger [was] like a tiger in faculty meetings" (Glucksberg & Keysar, 1990, p. 7), the properties of a tiger are most likely not associated with the hearer's mental representations of Roger. Thus, the inference that he acted vigorously cannot result from a matching process between the pre-existing features of two concepts.

3.2.2.2. Attribution of Properties – Metaphors as Categorization Statements

3.2.2.2.1. Core Elements of the Categorization View

In light of the shortcomings of feature matching models, Glucksberg and Keysar (1990) proposed a different parallel processing account for the comprehension of metaphors. According to the so-called **categorization view**, metaphors are processed in line with their syntactic form, namely, as class-inclusion statements. On that note, metaphors are to be

delineated from literal comparison statements since both types of statements involve different relations between concepts: Literal comparisons relate categories of the same categorization level and appear to involve feature matches. Metaphorical mappings, on the other hand, identify the target category as an instantiation of the superordinate source concept. According to Glucksberg and Keysar (1990), treating metaphors as class-inclusion statements comes with an important advantage: It explains why reversed metaphorical statements are generally false or anomalous (e.g. "*Sleeping pills are sermons", Glucksberg & Keysar, 1990, p. 12), while literal comparison statements merely shift the focus of the matching features when they get reversed (e.g. "Bees are like hornets" vs. "Hornets are like bees", Glucksberg & Keysar, 1990, p. 7).

Let us consider some examples such as the metaphors "my job is a jail" and "men are wolves". At first glance, it is far from obvious how target and source should relate hierarchically, more precisely, in which ways *jails* and *wolves* would be superordinate to *jobs* and *men,* respectively. Moreover, while some properties of a *jail* are applicable to *jobs* (e.g. being confining and unpleasant), others are rather peculiar attributions unless the speaker is genuinely a member of prison staff (e.g. barred windows, guards). By the same token, apart from a predatory nature, most of the characteristics of a wolf cannot be attributed to men (e.g. fur, feeding behaviour, etc.). On these grounds, Glucksberg and Keysar (1990) suggest the following: The source term of a metaphor does not actually refer to the category itself but serves as a basis to derive a more abstract, "diagnostic category" (Gibbs, 1994, p. 246).[17] Rather than to a particular token, a metaphorical jail hence "refers to a type of thing" (Glucksberg & Keysar, 1990, p. 8), that is a set of confining, unpleasant and involuntary situations. The original source (i.e. a prison as an imprisonment facility), meanwhile, constitutes a prototype of the newly derived category, whereas the target concept is incorporated into the novel category as a less prototypical member. As such, it inherits the central properties of the metaphorical concept.

According to Glucksberg, McGlone, and Manfredi (1997), the relation between source and target is not a one-way street: In line with Black's (1962) interactionist view, they argue that the target contributes to the process of categorization by constraining which of the source's properties are projected onto the superordinate concept. Given that the target concept is equally familiar to the hearer, it provides information "on the sorts of properties that may be plausibly attributed to it" (Glucksberg et al., 1997, p. 52). In other words, metaphor comprehension

[17] In a similar vein, Sperber and Wilson (2012) assume that metaphor comprehension relies on such ad hoc derivations, however, they reject the idea of a specific mechanism for metaphorical language. According to the relevance-theoretical approach, metaphors are merely part of a continuum of broadened and narrowed lexical meaning, ranging from loose talk (e.g. "Here's a Kleenex!" when handing someone a tissue) and hyperbolic statements (e.g. "Joan is the kindest person on earth", Wilson & Sperber, 2012, pp. 106, 110) to metaphors.

depends on people's expectations about the ways a target concept will be characterized.[18] Take, for instance, the metaphor "My lawyer is a shark". In this example, the target concept "lawyer" narrows down the possible candidates for property attribution to those dimensions which are meaningful, interesting or relevant to the target itself (e.g. properties relating to a lawyer's ambition, temperament, skills, reputation, experience, costs, etc.). Properties, on the other hand, which are irrelevant to the characterization of a topic will not be relevant to the novel superordinate category and need to be suppressed during the derivation (e.g. a lawyer's musical abilities, weight or height). Importantly, property dimensions do not only vary regarding their content for each target concept, but also in terms of their number: While categories such as "lawyer" are limited to a small tally of property dimensions (i.e. a high constraint target), others such as "life" or "my brother" are less restrictive and can be characterized in a myriad of meaningful ways. In such cases, the target only places few constraints on the process of property attribution. At bottom, Glucksberg et al. (1997, p. 59) argue that "understanding a metaphor […] requires two kinds of world knowledge [… that] are used interactively to generate interpretations": For one thing, people must be aware of the relevant dimensions of the metaphorical topic. They furthermore need to know sufficiently enough about the source concept and its properties in order to derive an abstract ad hoc category. Both types of knowledge contribute to metaphor comprehension to different degrees: While for some source concepts, the properties projected onto the novel superordinate category are rather unambiguous (e.g. a "jail" epitomizes an unpleasant, confining situation across different metaphorical contexts), others such as *a* "snowflake" offer a variety of possible property attributions (e.g. its uniqueness, ephemerality, delicateness, the way it floats trough the air, etc., see Bowdle & Gentner, 2005). In such cases, metaphor comprehension relies heavily on the constraining dimensions of the target concept (e.g. life, youth, a child, a dancer, etc.).

[18] Note that Leslie (2008) makes a similar claim regarding the relevant properties of kinds and their role in generalizations (recall chapter 3.1.2.1.): According to the Generics-as-Default-Hypothesis, every kind involves so-called characteristic dimensions, which prompt the search for hereto related information during reasoning processes, while ignoring less characteristic information. According to Leslie (2008), characteristic dimensions constitute people's background hypotheses about the properties of a kind. In a similar way, Glucksberg, McGlone, and Manfredi (1997) claim that any target concept (even an artifical one) may contribute to the derivation of an ad hoc category by providing characteristic dimensions. Targets thus constrain metaphorical reasoning to procede along these lines of characterization.

3.2.2.2.2. Ad Hoc Categories in Metaphor Comprehension

What are the implications of the categorization account? Firstly, the categorization view suggests that metaphor comprehension does not involve some elaborate cognitive mechanism. Rather than being artful figures of speech building upon some kind of genius, metaphors draw on one of the cognitive system's most basic competences: the capacity to categorize objects and entities into more abstract, superordinate categories. Secondly, while categorization is typically conceived as a process where a number of examples gradually lead to the construction of a novel category representation (i.e. exemplar learning), in the case of metaphor comprehension, the novel concept needs to be derived in a "one-shot" manner (Holyoak & Stamenkovic, 2018, p. 647). As opposed to regular categorization, the superordinate category of a metaphor furthermore does not receive a novel unambiguous label of its own, nor is it necessarily consolidated in long-term memory. In these three respects, metaphor comprehension can be aligned with another phenomenon from the categorization literature: According to Barsalou (1983, 1991), people derive so-called *ad hoc categories* on the fly whenever they try to achieve a particular situational goal. Planning a holiday, for instance, may involve ad hoc categories such as "things to pack in a small suitcase" or "activities to do […] in Japan with one's grandmother" (Barsalou, 1991, p. 1). Importantly, such goal-derived ad hoc categories are not determined by similarity between category members. They rather tie together a heterogeneous set of categories that are characterized by their shared extrinsic function. Against this backdrop, Markman and Stilwell (2001) suggest that ad hoc categories are in fact closely linked to relational and role-governed categories (recall chapter 2.2.2.2.): All three types of categories are specified by their relationships towards other concepts, the only difference being that, in the case of ad hoc categories, these relational structures need to be constructed offhand, rather than retrieved from memory. In that sense, ad hoc categories may "be computed using the same processes that serve other role-governed categories" (Markman & Stilwell, 2001, p. 339). It should be noted that, though most goal-derived categories are newly derived for specific contexts, they have the potential to be stored as mental representations in long-term memory: Frequent air travels, for instance, require to pack a small suitcase on numerous occasions. In a similar way, the abstract superordinate category of a metaphor may be retrieved from memory when reaching a certain degree of conventionalization. However, in the case of infrequent or novel metaphors, the category is generally derived unprompted and with regard to its particular context (i.e. the target category).

Let us turn towards the underlying computational mechanism which yields goal-derived and metaphorical ad hoc categories. Both Barsalou (1991) and Glucksberg, McGlone, and Manfredi

(1997) suggest that a cognitive process named *conceptual combination* enables the derivation of categories in a one-shot manner. By means of this process, people combine or manipulate pre-existing concepts into larger mental structures (see further Murphy, 2004). Conceptual combination in its simplest form underpins the interpretation of adjective-noun or noun-noun constructions (e.g. large daddy-longlegs, curtain rail), whereas, in more complex cases (e.g. things to pack in a suitcase), mental representations of "things", "to pack", and "suitcase" are merged as well as more general knowledge about travels (Barsalou, 1991). Different members are subsequently added to the newly created category. In essence, conceptual combination proceeds top down and requires only little experience with the actual members of the derived category. As Barsalou (1991, p. 4) puts it, ad hoc categories represent "idealized knowledge about how the world should be rather than normative knowledge about how it is" (see Goldwater et al., 2011 for a similar point and for experimental confirmation as regards relational categories).

There are different strategies to combine pre-existing concepts into an ad hoc category: For one thing, conceptual combination may rely on *property transfer* whereby a concept's most salient property is attributed to another concept (Holyoak & Stamenkovic, 2018; Murphy, 2004; Wisniewski, 1996, 1997). In this regard, people construe the first concept as a reference to a specific characteristic of the second (e.g. "white things" in a game of "I spy with my little eye"). A further strategy that is employed by conceptual combination is called *hybridization* and appears to be a special case of property transfer. In hybrid combinations, a newly derived concept inherits multiple properties from both of its constituents (e.g. "zebra horse" as a new crossbred species, Wisniewski, 1996). Both property transfer and hybridization may play an important role in the interpretation of antisemitic metaphors. Take, for instance, metaphorical statements such as "Jews are vermin/parasites/leeches". In order to derive an abstract metaphorical category from these source concepts, the target's property of being a social group should suffice to create a novel category representation which epitomizes a broad conception of social parasitism (e.g. exploitive or harmful social behaviours). This interpretation seems to be the result of property transfer and does not require any further expectations about the characterization of Jews as a social category. More enriched interpretations, however, seem to be the result of a hybridization process. In this case, a number of properties – or rather, property dimensions (e.g. behaviours, social status, attitudes) – are identified for the Jewish category, constraining the properties that project from the source onto the metaphorical ad hoc category. Note that hybridization presupposes a more elaborate representation of social concepts: Only if people have naïve theories about a social group's relevant characterizations, properties can be

narrowed down to more specific attributions (e.g. harmful social behaviours, object of exploitation, devious intent). A third strategy has been identified with respect to conceptual combination (Holyoak & Stamenkovic, 2018; Murphy, 2004; Wisniewski, 1996, 1997): During *relation formation,* concepts are placed in a common relational structure so that the resulting ad hoc category emphasizes on the relational features of the concepts in question. Such a combinational process may well apply to antisemitic metaphors based on role-governed categories (e.g. "Jews are puppeteers"). As argued by Markman and Stilwell (2001), role-governed categories generally result in relational interpretations of conceptual combinations. Importantly, relation formation does not only involve the relation between the two concepts (i.e. the category "Jews" occupying the role of a "puppeteer"), but also their defining connections to other positions in the relational structure (i.e. control and manipulation of marionettes).[19]

3.2.2.2.3. Correspondences with Structural Thinking

At large, what does the categorization view and its implications reveal with respect to antisemitic metaphors and their impact on social concepts? First and foremost, assuming that metaphors are processed in line with categorization statements suggests that they reflect a fundamental cognitive capacity which organizes concepts and entities into larger hierarchical structures. A correspondence between antisemitic metaphors and categorization would thus explain and reinforce their profound influence on conceptual representations: Metaphor comprehension does not rely on a highly demanding computational process but rather on a basic cognitive function. Secondly, and more to the point, the categorization view assumes that metaphor comprehension is based on the derivation of an abstract, superordinate concept from the source category, or in broader terms, a hierarchic metaphorical structure. A metaphor is only understood successfully if the target is identified as a member of the ad hoc concept and inherits the corresponding properties by virtue of this classification. For antisemitic metaphors, this means that the social category "Jews" needs to be identified as an instantiation of an abstract

[19] Relation formation presumably applies to the metaphorical sources such as "parasite" and "leech" as well. Even though taxonomic concepts are traditionally treated as entity concepts that are defined in terms of their shared internal features (see e.g. Gentner & Kurtz, 2005), "parasite" and "leech" may in fact be relational categories due to their defining relation towards other concepts (e.g. the category of a "host"). As noted by Goldwater et al. (2016, p. 124), taxonomic categories "can have both relational components and shared attributes" so that the distinction between relational and feature-based categories is rather fuzzy at times. Arguably, people chose their strategy of conceptual combination individually such that the choice between property transfer and relation formation may largely depend upon individual differences in people's metal representations – that is, whether they represent "parasite" or "leech" in terms of relations or regarding their inherent features.

superordinate concept, exemplified by the terms “parasite”, “vermin” or “puppeteer”. As such, it displays the metaphorical concept’s properties. This line of processing runs in parallel with another reasoning style that we have come to know as an externalist constraint on social categorization (recall chapter 2.2.2.): According to Vasilyeva et al. (2018), people **think structurally** about a social category when it is identified as the occupant of an abstract position within the social hierarchy. As a result, the properties associated with the respective structural position are attributed to the social category by virtue of this relation. Importantly, properties in both metaphor comprehension and structural thinking emerge from the abstract concept or position, rather than the internal structure of the social category itself. That is to say, in both cases people modify their conceptual representations of a social group based on the category’s role in an external structure.

Another point of comparison comes to mind: Vasilyeva et al. (2018) point out that social categories are frequently confounded with their structural position so that characteristics originating from a structural position blend in with the intrinsic characteristics of a social category. Those properties thus become consistent with two causal models, either identifying internal features (i.e. internalist thinking) or the structural position (i.e. structural thinking) as causes of a social category’s properties and behaviours. Such confounds are, for instance, reflected in social deterministic beliefs (recall chapter 2.2.1.): People often hold the conviction that the identity of a social group is determined by an internal essence which, in turn, is permanently shaped by external social factors such as upbringing, education or social background. Regarding the origins of confounds, we have identified a lack of observable variation within the social structure (recall chapter 2.2.2.): While people encounter different exemplifications for role-governed categories in different contexts (e.g. man, woman, stranger, friend, dragonfly, etc. for the category “guest”) and subsume different exemplars when generating goal-derived ad hoc categories (e.g. t-shirts, tooth brush, shorts, etc. for “things to pack in a small suitcase”), variation as regards social positions and metaphorical ad hoc categories is considerably scarce. The first would require historical or cross-cultural comparisons in order to uncover different variants (e.g. matriarchal societies, changing colour attributions for gender categories), whereas the latter are derived top down merely to incorporate two category members: the target and source concept. Similar to social categories, the ad hoc categories of antisemitic metaphors should therefore invite confounding so that the properties determined by the metaphorical source are aligned with the internal features of the target category.

Needless to say, the alignment between the structural constraint on social categorization and metaphor comprehension only holds up to a certain point: For instance, when people think structurally about social categories, they do not derive anew a corresponding social structure but draw on their pre-existing representations. Furthermore, confounds may be more pronounced for social categories since they additionally share category labels with their structural position. Metaphorical ad hoc categories, on the other hand, borrow their name from the source concept so that they bear greater potential of being confounded with the prototypical source. Nonetheless, certain aspects seem to follow the same path: Firstly, in both instances, people do not think of a diverse set of category members but focus on one in particular, namely a Jewish social category that is located within a conceptual structure. Secondly, properties that are attributed to a category by means of a metaphor or a structural position are largely indistinguishable from the category's internal features: An attribution such as "Women usually quit their jobs after having a baby" may be related to inherent causes (e.g. priorities, abilities) or arise from a structural constraint (e.g. wage inequality). Similarly, properties which are attributed in the context of an antisemitic metaphor (e.g. exploitive or harmful social behaviours) may be caused by an internal motivation (e.g. personal attitudes and intent) or point to the category's subordination to a metaphorical concept (e.g. dependency on a host, professional practices). This effect seems to result from the aforementioned lack of variation. However, the categorization view offers another feasible explanation: Given that ascribed properties are always the outcome of an interaction between the source and the target concept, both may be identified as the causes of a property in question. For instance, to be in control of other members of society may give malicious pleasure to a "Jewish puppet master" (i.e. an inherent cause), but at the same time, it is the abstract metaphorical concept that enables him to act upon his internal motivation (e.g. exercising the profession of a puppeteer). Summing up, many of the characteristics of structural reasoning turn up again when we take a closer look at the comprehension mechanism as proposed by the categorization view. Such parallels underpin the assumption that antisemitic metaphors prompt people to think structurally about a Jewish social category and flout considerations about possible variation within these structures. Furthermore, merely by processing an antisemitic metaphor, people may be encouraged to align properties emerging from an external source with the inherent properties of the category. Simply put, the processes that underlie the comprehension of antisemitic metaphors may carve out a route along which people think about Jews as a social group. As a consequence, their mental representations should become increasingly rich in structure, homogeneous as regards their members and properties, and sharply delineated from other categories in the social domain.

Up until this point, we have focused on structures which consist of a target, a source and a superordinate metaphorical category. Societal structure, however, is much more complex and is most certainly not derived on the fly as is the case with metaphorical hierarchies. In this regard, another mechanism has been proposed for implementation which aligns the representations of social categories in order to extract their interconnected relationships (Christie, 2017). Crucially, structural alignment has not only been suggested as the basis of this kind of relational knowledge, but also as another approach to metaphor comprehension (e.g. Bowdle & Gentner, 2005; Gentner, 1988; Gentner & Bowdle, 2008; Wolff & Gentner, 2011). Against this background, the subsequent chapter will review on the third perspective on metaphor processing as well as its implications for the constraints on social categorization.

3.2.2.3. The Analogical View – Metaphors as Structural Mappings

3.2.2.3.1. Outlining the Structure Mapping Mechanism

In chapter 3.2.2.1., we have encountered a line of approach that – drawing on the Aristotelian view – conceptualizes metaphor comprehension in terms of a process of comparison. Another influential theory argues along this lines, however, with a very different conception of comparison in mind: Comparisons have previously been considered as a feature matching processes whereby the relative amount of shared features determines whether two concept are perceived as similar (e.g. Ortony, 1979; Tversky, 1977). The so-called **analogical view**, on the other hand, suggests that a mechanism called *structure mapping* underpins any form of comparison (Gentner, 2003; Gentner & Bowdle, 2008; Gentner & Clement, 1988; Markman & Gentner, 2000; Wolff & Gentner, 2011).[20] The difference between former comparison theories and the analogical view on metaphor comprehension is that only the latter assumes the target to be "extended by comparison to the base" (Markman & Gentner, 2000, p. 518) via the alignment of two abstract relational structures. In other words, matches are established between embedding structures, rather than the particular contents of concepts. Given that structures generally comprise of multiple concepts, structure mapping furthermore relates to entire conceptual domains. In this regard, the analogical view does not only differ from former

[20] The analogical view owes its name to the fact that structure mapping originated as a theory of analogical reasoning (see e.g. Gentner, 1983). Though starting off as a rather specific theory, structure mapping theory has made its mark in different branches of research, ranging from metaphor comprehension and the representation of relational knowledge to various aspects of social cognition (Bach, 2014; Christie, 2017; Gerson & Woodward, 2012; Meltzoff, 2005).

comparison models, but also from the categorization view which assumes a metaphorical mapping to be a transfer between single concepts.

Let us first answer the question how conceptual domains are defined within the structure mapping framework. According to Gentner (1983), the conceptual system represents different domains of knowledge in the form of networks, consisting of both nodes and predicates. While nodes correspond to the specific concepts of a knowledge domain, predicates apply to the nodes of a network. As such, they "express propositions about […] concepts" (Gentner, 1983, p. 157). A crucial distinction is drawn between different syntactic types of predicates (Gentner, 1983; Markman & Gentner, 2000): Firstly, *attributes* are formally characterized by taking only a single argument (e.g. LARGE(x)). As such, they relate to descriptive information about the individual concepts. Secondly, *functions* map a single argument onto values different than truth values and, by and large, relate to psychological dimensions (e.g. COLOUR(x) → GREY). Thirdly, predicates are *relational* when they provide multiple argument slots (e.g. COLLIDE (x,y)). Such predicates can be further subdivided into first-order predicates and second-order predicates whereby the first apply to the actual objects of a knowledge domain (e.g. COLLIDE (x,y)), and the latter take other predicates as their arguments (e.g. CAUSE [COLLIDE(x,y), STRIKE (y,z)], see Gentner, 1983, p. 157). Crucially, the structure mapping framework suggests that distinguishing between different types of predicates sheds light on different types of comparison: Two entities are literally similar to one another when their conceptual representations have both attributes and relational predicates in common. Analogical reasoning and metaphor comprehension, on the other hand, exclusively rely on an overlap between the relations of two knowledge domains. What is most striking about this distinction is the following: According to structure mapping theory, relations play a pervasive role in all sorts of comparisons (Markman & Gentner, 2000): Whenever people contrast two conceptual representations – irrespective of whether it is in the wake of an analogy, a metaphor or in order to assess literal similarity – they generally match the relational structure between two domains of knowledge. In that sense, even in the simplest cases of similarity judgements, feature matches are not sufficient and require additional structural computations. The reason why literal similarity is nonetheless easier to comprehend is therefore not due to a difference in processing, but because matches between features and relations "are all mutually supportive" (Gentner, 2003, p. 199). Conversely, the relational correspondences in analogies and metaphors are not underpinned by attribute matches and may even be subverted by dissimilar features. Take, for instance, the metaphor "Blood vessels are aqueducts" (Gentner & Bowdle, 2008, p. 110) which highlights several relational matches between the source and the target domain (e.g. both

transport something that is needed, both connect remote parts of a system, etc.). At the same time, interpretations pay no attention to particular attributes – presumably, because they are largely conflicting (e.g. blood vessels are elastic, fragile, red; aqueducts are solid, durable, grey, etc.). Concerning this matter, Gentner (1988) found that children aged five mainly interpret metaphors such as "A tire is a shoe" in terms of shared features (e.g. "Sometimes your shoe is black and the tire is black", Gentner, 1988, p. 52). This suggests that metaphor comprehension is initially guided and supported by feature similarity, thus resembling the easier case of literal comparisons. It is only at the age of nine that children develop a relational focus (e.g. "You can go places on both", Gentner, 1988, p. 52) – a preference that is characteristic for metaphor comprehension in adults (Gentner & Clement, 1988).

How are relational matches computed in the first place? In this respect, the structure mapping mechanism is assumed to comprise of two interrelated components, namely a process of alignment and a process of projection. In a first step, two conceptual structures are aligned in order to establish a maximum number of relational communalities. Importantly, such structural alignments do not proceed arbitrarily: Relations can only be matched when they are semantically similar, that is, if they have at least some "semantic subcomponents in common" (Markman & Gentner, 2000, p. 505). More importantly, the process of structural alignment adheres to a number of structural constraints (Bowdle & Gentner, 2005; Gentner, 1983, 2003; Gentner & Bowdle, 2008; Gentner & Clement, 1988; Gentner & Markman, 1997; Markman & Gentner, 2000; Wolff & Gentner, 2011): For one thing, people can only align conceptual domains which are **structurally consistent**. This means that individual predicates of the target domain must correspond to no more than one predicate within the source domain (i.e. one-to-one-correspondence). Structural consistency is established in a second way: Provided that relational predicates correspond in a one-to-one manner, their arguments are likewise placed in correspondence (i.e. parallel connectivity). In this way, an alignment starts out as a series of local matches that becomes increasingly more global. Note that parallel connectivity does not call for identity between single concepts: "The corresponding objects in the base and target don't have to resemble each other; object correspondences are determined by roles in the matching relational structures" (Gentner & Clement, 1988, p. 313). Consider, for instance, Rutherford's famous analogy between the solar system and the hydrogen atom. The interpretation is mainly guided by shared relations between sun (x) and planets (y), nucleus (x) and electrons (y) (e.g. MORE MASSIVE THAN (x,y), REVOLVES AROUND (y,x), ATTRACTIVE FORCE (x,y), see Gentner, 1983, p. 163). Conversely, the objects of both domains are rather different regarding their attributes in the sense that "there is no attempt to

match the nucleus with the sun in color, size, or temperature“ (Gentner & Clement, 1988, p. 313). In cases where another relation is the argument of a matched relational predicate (i.e. a second-order predicate), however, parallel connectivity invites for re-representation: More precisely, non-matching first-order predicates (e.g. BIGGER (x, y) and DARKER (x,y)) can be reconstructed in order to establish a higher degree of structural consistency (e.g. GREATER (SIZE (x), SIZE (y)) and GREATER (SHADING (x), SHADING (y), see Gentner & Kurtz, 2005, pp. 163f.). A second structural constraint on the process of structural alignment is the so-called **systematicity principle** according to which people favour deep relational matches over isolated matching predicates. The systematicity principle thus acts as a “selection filter” (Gentner & Clement, 1988, p. 345) whereby the amount of relational matches relevant to an analogy or metaphor is narrowed down to the richest, most fruitful contenders. Let us consider the example “Socrates was a midwife” (Bowdle & Gentner, 2005, p. 196):

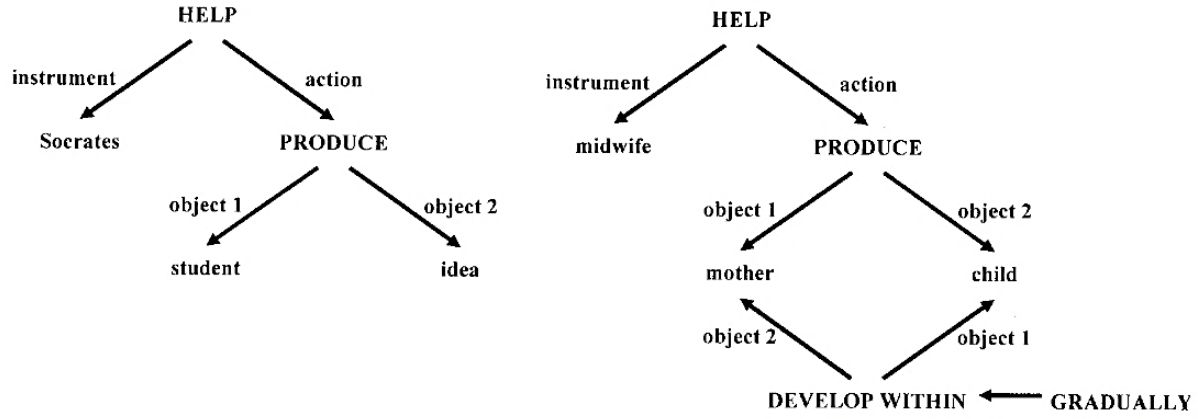

Figure 3. Structural representations of the target and the source domain for the metaphor "Socrates was a midwife" (Bowdle & Gentner, 2005, p. 196)

The metaphor establishes a system of relations consisting of the second-order predicate “help”, a first-order predicate “produce” as well as their respective arguments. Any interpretation will revolve around this interconnected system (e.g. Socrates helped his students to realize new ideas, ideas develop within the students, etc.), whereas single relations – even though relational matches as well – are irrelevant to the comprehension of the metaphor (e.g. both Socrates and midwifes have mothers on their own, they breath air, etc.). Systematicity moreover accounts for the fact that interpretations are often coherent and relate to causes: Since both causal and coherent information tends to be represented in terms of higher-order predicates, people are more likely to select those structural matches during the alignment process (Gentner & Bowdle, 2008).

Once the maximum number of relational communalities is established between two domains, the second process of the structure mapping mechanism is set in motion (Bowdle & Gentner,

2005; Gentner, 2003; Gentner & Bowdle, 2008; Gentner & Markman, 1997; Markman & Gentner, 2000; Wolff & Gentner, 2011): At this stage, people project further elements from the source to target domain. Importantly, the projection process also abides by the systematicity principle (Gentner & Markman, 1997; Markman & Gentner, 2000). Rather than projecting any element, only those predicates that are already linked to the common relational system of the source domain become so-called *candidate inferences*. That is to say, people do not simply infer random facts, but "instead project inferences that complete the common system of relations" (Gentner & Bowdle, 2008, p. 110). Regarding our example of midwifing Socrates, a plausible candidate inference would refer to the child's steady development within the mother's womb (for illustration, see Fig. 3). Unrelated relational elements, on the other hand, are not projected onto the target (e.g. the organization of midwifery associations). Note that inference candidates – as implied by the term itself – are mere candidates for metaphor and analogy interpretations (Markman & Gentner, 2000): Before they become final output interpretations, people need to check them against what they generally know about the target domain. In this way, people may choose different inferences depending on their individual representations of knowledge.

Above all, structural alignment and inference projection are dynamic, interrelated processes (Bowdle & Gentner, 2005; Gentner & Hoyos, 2017). As a result, the structure mapping mechanism can account for a variety of comparisons and their different interpretations: A comparison may highlight the shared relations between two domains when the representations of a target and a source are equally rich in structure. In such cases, interpretations predominantly rely on the process of structural alignment. At the same time, some comparisons seem to enlarge or even alter our knowledge domains. In these cases, elements connected to the structure of the source domain are absent within the target's common relational network and get projected to this less systematic domain (Markman & Gentner, 2000). It should be pointed out that both structural alignment and inference projection may yield a number of different interpretations. The structure mapping mechanism can therefore capture the fact that the very same comparison is open to different interpretations. People may generate a variety of inference candidates or highlight different shared relations, yet they ultimately favour one over the others based on how richly their individual domains of knowledge are structured (Markman & Gentner, 2000).

3.2.2.3.2. Structure Mapping and Relational Categories

The structure mapping mechanism is mainly dependent “on syntactic properties of the knowledge representation, and not on the specific content of the domain” (Gentner, 1983, p. 168). The output of this comparison process does not rely on the internal features of particular concepts, but rather on relations between concepts, matching across different conceptual domains. Accordingly, relations determine whether two things are perceived as similar or not. This characterization is reminiscent of a certain type of categories which we have encountered in chapter 2.2.2.2.: In the categorization literature, a general distinction is drawn between so-called feature-based categories and relational categories (see e.g. Gentner, 2005; Gentner & Asmuth, 2017; Gentner & Kurtz, 2005; Goldwater et al., 2016, 2011; Markman & Stilwell, 2001). While the first are specified by their internal properties so that category members are intrinsically similar, the latter display common relations to other concepts. Features and attributes, on the other hand, are less decisive for the identity of relational categories and often vary considerably. Importantly, this distinction of categories raises a crucial question with respect to conceptual development: Particularly in view of the fact that relational categories are acquired later in childhood or start off as feature-based representations (e.g. "uncle" refers to a bearded man instead of a position in the family system, Gentner & Asmuth, 2017), it is important to ask how children acquire relational categories. In contrast to feature-based categories, relational categories cannot be derived from direct experience. Children may figure out what an apple is by looking at an exemplar, but they “cannot tell which things are gifts, […] or weapons, without knowing something about their relations to other entities” (Gentner, 2005, p. 250). The critical question, therefore, is how children learn to identify the members of a relational category if not directly from the features of the exemplars in their environment. In this regard, Gentner and colleagues (Gentner, 2005; Gentner & Asmuth, 2017; Gentner & Kurtz, 2005) have argued that the structure mapping mechanism provides the basis for acquiring relational categories. In that sense, structure mapping is not only proposed as a computational mechanism which steps in whenever metaphors, analogies or similarity statements are processed, but as a more general learning mechanism by means of which children extract defining relations. Structure mapping may underpin category acquisition in several ways (Gentner, 2005): For one thing, the mechanism highlights the common relations of two representations which cannot be observed in one representation alone. It further enables inferences from one representation to another. Thirdly, the mechanism alters conceptual representations to yield more global and consistent relational matches (i.e. re-representation). In all these ways, representations become increasingly more uniform as regards their relations

such that children can successfully abstract away from particular objects and reuse the resulting structural information to identify new category members in different contexts and semantic domains (Gentner, 2005). It should be emphasized that structure mapping evolves progressively (Gentner, 2003; Gentner & Asmuth, 2017; Gentner & Hoyos, 2017; Gentner & Kurtz, 2005): Children first align rather similar concepts so that initial comparisons turn out to be concrete, bristling with specific attribute matches (e.g. "cat" and "dog" are aligned to derive the category "pet", Gentner & Asmuth, 2017). Starting from such concrete abstractions, less similar concepts are aligned (e.g. a pet python) to the effect that the extracted communalities become more abstract and global.[21] Interestingly enough, such far alignments are no longer guided by feature similarity. They seem to be prompted by linguistic input instead (i.e. labels, relational language, explicit comparisons): For instance, Gentner, Anggoro, and Klibanoff (2011) found in their study that children can successfully identify members of a novel relational category as early as age three (e.g. a "dax" referring to cutting tools) after aligning highly similar exemplars (e.g. different knifes) and less similar category members during a training phase (e.g. an axe, a saw). However, they only succeeded when the training was complemented by relational language (e.g. "The knife is the *dax* for the watermelon", Gentner et al., 2011, p. 1176).

There is evidence from other fields than categorization research that children perform structural alignments in order to extract relational information: Haryu et al. (2011), for instance, found that three- and four-year-old children were more likely to extend a novel action verb to a different context "based on the sameness of action" (Haryu et al., 2011, p. 683) when the object in the novel context was perceptually similar. When first given the opportunity to align a series of contexts with similar objects, children were even able to extent the verb meaning to dissimilar contexts. Put differently, repeated alignment enabled children to derive more global abstractions for the relational meaning conveyed by verbs. In a further study by Gerson and Woodward (2012), seven-month-old infants correctly imitated a goal-directed action that was executed by a claw (e.g. grabbing one out of two toys) after they had aligned the tool's grab with their own actions during a familiarization phase. When alignment was not possible (e.g. children were familiarized with the toy and the claw separately, or they were not encouraged to grasp for the toy themselves), children directed their actions to one of the two toys by chance. In other words, matching their own intentional structure (i.e. the relation between an object and the self) with the actions of another agent licensed infants to derive the correct goal of an action.

[21] Importantly, progressive alignment also becomes evident in the development of metaphor comprehension: Gentner (1988) found that early interpretations predominantly refer to concrete feature similarities, whereas children displayed a more abstract, relational understanding by the age of nine. This developmental shift may also reflect that structure mapping begins with rather concrete comparisons, before matches become more abstract and relational.

Finally, structural alignment seems to be implemented in children's understanding of false beliefs: Hoyos, Horton, and Gentner (2015) conducted a series of ToM-tests with five-year-olds before proceeding with an *unexpected content task.* The task depicted a variety of characters that either held true or false beliefs about the content of a box. In the critical condition, beliefs were illustrated as thought bubbles and children were encouraged to compare those mental states with each other as well as with the real content of the box. In two further control conditions, children either compared only the items and characters of a scene, or directly performed a series of posttests. Results showed that only the children in the critical condition performed significantly better in the ToM-posttests, suggesting that the alignment of mental states eased their understanding of false beliefs. A second experiment showed that alignment is generally facilitated by similarity: When situations during the unexpected content task comprised of similar characters and objects, children achieved higher posttest results, even though they were less encouraged by the experimenter to compare mental states. In view of such experimental findings, the structure mapping mechanism has been suggested as a domain general learning mechanism, not only for the acquisition of relational categories, but with regard to a variety of knowledge domains and cognitive capacities. At bottom, structure mapping is crucial whenever comparisons between conceptual structures represent a starting point for further knowledge acquisition (e.g. relational categories, verb learning) or when relational information is relevant and informative to a current task (e.g. perspective taking, imitation etc., see further Bach, 2014; Meltzoff, 2005).

3.2.2.3.3. Implications for Social Categorization

Where does this leave us regarding the impact of antisemitic metaphors on social categorization? Before addressing this question, let us recall that social categories share some of their central characteristics with a particular subtype of relational concepts (see chapter 2.2.2.2.): Just like role-governed categories, they refer to particular positions within a relational structure or, more precisely, to positions in society. Social categories furthermore possess a range of properties by virtue of their structural location. In the same ways as thieves display criminal intent due to their role within the schematic structure of a robbery, women are paid less owing to gender inequality and lower social strata are more susceptible to illness due to their limited access to the health care system. In light of such parallels, it can be assumed that the development of social category representations similarly depends on the structure mapping mechanism that is generally assumed for role-governed categories. However, structure mapping

does not only contribute to the acquisition of single social or role-governed categories: Given that, due to the systematicity principle, the mechanism prefers large systems of relations and, moreover, extends or refines them in the course of further comparisons, structure mapping also provides the grounds for generating entire knowledge domains. In other words, forming a rich and informative representation of the social structure on the basis of which people can acquire individual social categories also relies heavily upon the general learning mechanism of structure mapping.

It has been argued that one way to invite comparisons between representational structures is through language. Language – whether it is in the form of relational language, explicit comparison statements or labels referring to dissimilar objects – serves as a powerful tool to direct people's attention towards the relations within their conceptual domains. Such linguistic invitations are particularly important when object similarity does not cue structural comparison. In cases where matches between observable attributes and relational matches do not mutually support one another, the comparison process is sparked off by language alone. It is against this background that we can answer the question what the effects of antisemitic metaphors on social categorization may be: From the analogical point of view, metaphors are implicit comparisons that rely on the exact same mechanism as social category learning. In that sense, antisemitic metaphors are nothing other than linguistic prompts to such relational learning processes, or even instantiations of the same: During metaphor comprehension, people identify Jews as a social category and align the social structure with the structure of the source domain. Depending on the particularities of each antisemitic metaphor, relations between a Jewish social category and other social concepts are matched with the relations between species (i.e. zoomorphism-based metaphors), or with the relations between the participants of a puppet play (i.e. the puppeteer metaphor). In a second step, inference candidates project from the source to the target domain, pertaining for example to harmful consequences for the host category or to the subjection of the marionettes. Such inferences may become persistent add-ons of a Jewish social category when the output interpretations are stored in long-term memory. Notwithstanding, we should keep a sense of proportion here: Not every encounter with an antisemitic metaphor will automatically result in category learning. In fact, people often hold egalitarian beliefs or try to comply with pro-social norms so that effects can be mitigated by use of self-regulatory processes and cognitive control (Amodio, 2014). While adding relations to the conceptual representation of the social domain is not necessarily the immediate effect of processing an antisemitic metaphor and rather applies to extreme cases or repeated exposure, we can mark out a more subtle, yet pervasive impact on social categorization: By and

large, the structure mapping mechanism highlights the relational structure of two domains under comparison. At the very least, people are thus prompted to locate Jews as a social category within their mental representations of the social domain. Put differently, in order to understand an antisemitic metaphor, people need to identify the defining relations towards other social groups before they can match them to the structure of the source domain – a reasoning process corresponding to the structural thinking constraint on social categorization (see chapter 2.2.2.). In this regard, antisemitic metaphors do not just contribute to social categorization by making available negative associations (i.e. conceptual content), but they substantially influence how a Jewish social group and their alleged properties are mentally represented in the conceptual system. Antisemitic metaphors achieve this effect by highlighting the deterministic nature of the social structure in the course of their comprehension.

3.2.3. Summary and Discussion

The fascination with metaphor reaches back to ancient times and has produced a large body of scientific research over the years. Originating from Aristotle's treatise on poetry, metaphors have traditionally passed for ornaments of speech which guide the hearer's attention towards hitherto unnoticed communalities between a metaphoric source and target. In this vein, metaphors are considered to be implicit comparison statements so that their success depends upon the hearer's talent to unravel hidden and insinuated similarities. In recent years, however, the understanding of metaphors has radically changed. Corroborated by linguistic and psychological data, it has been argued that metaphors are "not just nice" (Ortony, 1975, p. 45) but in fact constitute an important means of communication and thought: Aside from being frequent in everyday language (e.g. Lakoff & Johnson, 1980a; Ortony & Fainsilber, 1987), metaphors often serve as explanatory models to illustrate complex issues in science (e.g. Gentner & Grudin, 1985; Kuhn, 1993), they are crucial in creative thought (Cacciari et al., 1997) and, furthermore, their interpretation has been found to be immediate and mandatory, rather than a matter of choice (e.g. Glucksberg et al., 1982; Ortony et al., 1978). In light of these findings, metaphors have been exempted from their longstanding status of exceptional poetic devices and are nowadays treated as a central phenomenon in human cognition.

While there is general agreement over this particular point, the question as regards metaphorical meaning is still a source of contention: Some theorist have rejected the existence of a special metaphorical meaning altogether and conceptualized metaphors as a function of language use (e.g. Davidson, 1978; Wilson & Sperber, 2012). At the other extreme, representatives of the so-

called conceptual metaphor theory have claimed that metaphorical mappings represent an organizational tool of the conceptual system (e.g. Gibbs, 1994, 2011; Lakoff & Johnson, 1980a, 1980b): From this point of view, complex or unfamiliar domains of knowledge can only be understood through the lenses of concrete, familiar and often embodied source domains. Yet another controversy is connected to the question of the nature of metaphors: In the preceding chapter, we have focused on the ongoing issue of how their interpretations arise in the first place. That is to say, which are the computational mechanisms that enable metaphorical reasoning? Three lines of approach have predominated discussions on this matter over the last decades: Among the first to spell out an explicit mechanism of metaphor comprehension are so-called feature matching models (e.g. Johnson & Malgady, 1979; Miller, 1993; Ortony, 1979; Tversky, 1977). Such models argue that metaphors elicit a process of comparison whereby shared features between two concepts are identified. Metaphoricity emerges due to the fact that the matching features are far more salient in the source as compared to the target concept (Ortony, 1979). Beyond that, metaphors and similarity statements are identical. A second approach is known as the categorization view according to which metaphors are equivalent to class-inclusion statements (e.g. Glucksberg & Keysar, 1990; Glucksberg et al., 1997). With regard to the underlying computational mechanism, the categorization view proposes that metaphor comprehension starts off with the derivation of an abstract superordinate concept from the source category, while the target concept supplies the relevant property dimensions. Subsequently, the target is subordinated to the abstract metaphorical concept and inherits the respective properties as a result of the categorization process. As for the third perspective on metaphor comprehension, metaphors are lined up with analogical reasoning (Gentner, 1988; Gentner & Bowdle, 2008). On this note, metaphors are conceived as implicit comparisons between entire domains of knowledge, drawing on the so-called structure mapping mechanism. During structure mapping, the relational structures of the domains under comparison are aligned in order to yield the maximum relational overlap. At a second stage, inferences can be drawn from the more systematic source to the less systematic target domain. Interpretations of a metaphor may therefore relate to relational matches between two conceptual domains as well as to several inferences that are projected from one domain onto the other – it is due to people's individual domains of knowledge and their particular structure that one is ultimately favoured and serves as the dominant output interpretation (Markman & Gentner, 2000).

As is often the case, accounts of metaphor comprehension have been challenged on various grounds: The feature-matching view was soon rejected in view of its difficulties with non-identical property matches and novel property attributions. Furthermore, it has been noted that

the distinction between similarity and metaphors is rather flawed when purely based on saliency differences (Gentner & Clement, 1988; Glucksberg & Keysar, 1990). Essentially, the rejection of feature matching models exemplifies a general conviction in recent metaphor research: Matches between the internal features of two concepts do not suffice to fully capture metaphor comprehension so that further relational considerations are required. The categorization view, on the other hand, has run into a different problem (Bowdle & Gentner, 2005): Assuming that each metaphor creates a novel ad hoc category and the same source concept may occur in different metaphorical statements, an infinite number of possible metaphorical categories should be available for each source term (e.g. unique, ephemeral, delicate, white or wintry things, falling softly from the sky, changing the sight of the landscape, etc. for "X is a snowflake"). Hearers would have to derive these different possibilities and keep them active during metaphor comprehension until "the target concept is scanned for dimensions of applicability" (Bowdle & Gentner, 2005, p. 195). Additionally, once a suitable possibility is selected, the superfluous ad hoc categories would need to be suppressed – suggesting that metaphor comprehension is in fact a highly demanding computational task. Importantly, this very argument has been submitted as a point of criticism against the analogical account: While structure mapping involves entire conceptual domains so that large portions of our cognitive resources should be in use during metaphor comprehension, the categorization view has been favoured because interactions between two single concepts are "computationally less demanding" (Holyoak & Stamenkovic, 2018, p. 464). Empirical evidence also pleads for the categorization view: For instance, it has been found that metaphors are not processed in accordance with comparison statements, but rather along the lines of subordinate-superordinate relations (Glucksberg et al., 1997). It has furthermore been found that people prefer the grammatical form of a categorization statement for familiar metaphorical statements (i.e. "X is a Y") as opposed to the overt comparison form (i.e. "X is like a Y"). However, the pattern was reversed for novel figurative statements (Bowdle & Gentner, 2005). Bowdle and Gentner (2005) additionally showed that conventional metaphors are processed faster when phrased as categorization statements in lieu of overt comparisons. At the same time, novel metaphors were more rapidly understood as comparison statements. In light of such findings, Bowdle and Gentner (2005; 2008) have proposed a theory of metaphor comprehension which incorporates both the categorization and the analogical view: According to the so-called **career of metaphor hypothesis**, metaphors rely on different mechanisms depending on their degree of conventionalization: While novel metaphors are processed as implicit comparisons by use of the structure mapping mechanism, familiar and conventional metaphors are instances of a

categorization process. The difference towards Glucksberg and Keysar's (1990) view, however, is that the abstract metaphorical source is not newly derived but consolidated in semantic memory instead. The career of metaphor hypothesis was immediately brought into question due to empirical evidence (Holyoak & Stamenkovic, 2018): Various studies demonstrated that aptness of a metaphor – rather than its conventionality – serves as a better predictor for different experimental patterns (e.g. Glucksberg & Haught, 2006; Jones & Estes, 2005; Pierce & Chiappe, 2008). With regard to Bowdle and Genter's (2005) results, it has been argued that novel metaphors are often only created for experimental purposes and are therefore less apt to begin with. In light of this, the effects of aptness may be easily confounded with conventionality. At large, this line of reasoning points at one of the major difficulties in metaphor research (Holyoak & Stamenkovic, 2018): Metaphors are by no means a unified phenomenon. They vary along various dimensions (e.g. familiarity, aptness, concreteness), come in different syntactic forms (e.g. nominal, verbal, proportional, attributional) and occur in a variety of contexts (e.g. everyday language, scientific models, political discourses, poetry or other literary genres). What is more, their interpretation sometimes relies on matching features and similar relations, whereas in other times, they transfer properties from the source to the target concept. Occasionally, interpretations may even be a conglomeration of all of these elements. What lessons can we draw from this? It can only be indicated here that „one size does not fit all" (Glucksberg & Haught, 2006, p. 938). That is to say, none of the abovementioned theories can account for the entire broad spectrum of metaphors including their different characteristics and contexts: Complex – that is proportional, poetic or unapt – metaphors are best explained by reference to the analogical view, whereas simple ones seem to behave more like class-inclusion statements, supporting the categorization view. Above all, it appears that the career of metaphor hypothesis is right in one assumption: Metaphor comprehension relies different computational processes, contingent upon the individual characteristics of metaphors and the context in which they occur. Relatedly, Gibbs (1992) put forward an argument for theoretical plurality: From his perspective, different metaphor theories describe different temporal stages of metaphor comprehension: While the analogical and the feature matching view primarily concentrate on the *products* of metaphor comprehension (i.e. recognition, elaborate interpretations), the categorization view is more concerned about "the immediate, moment-by-moment *process* of creating meaning" (Gibbs, 1992, p. 576, see also 1994). Thus, different theories of metaphor are not necessarily mutually exclusive, they simply relate to different aspects and time scales of a non-uniform phenomenon.

Finally, let us reiterate what the different computational mechanisms imply for the impact of antisemitic metaphors on social category representations. According to the categorization view, a target concept is integrated into an abstract metaphorical hierarchy. In most cases, such structures are generated in a one-shot manner: People derive a metaphorical ad hoc category from the source concept while adhering to the characteristic dimensions of the target so that only those properties are preserved which may be applied to the target concept. A first subtle influence of antisemitic metaphors on social categorization may therefore be that people access their expectations about the intrinsic characteristic dimensions of a social category (e.g. behaviours, social status, attitudes). As a result, such dimensions would become more central in further reasoning processes. However, the effects of metaphorical reasoning reach beyond merely highlighting the internal property dimensions of the social domain: The categorization view claims that a target receives its properties from the metaphorical source concept by virtue of a subordinate-superordinate relation. In the case of antisemitic metaphors, properties are ascribed to a Jewish social category as a result of its relation towards an abstract metaphorical concept. This line of processing runs in parallel with an externalist constraint on social category representation, namely the structural reasoning account: When people think structurally about social categories, they locate them within the social structure and attribute properties to them based on their connections towards other structural elements. Against this background, it can be assumed that antisemitic metaphors represent linguistic invitations to this particular style of reasoning: People may continue to think about the alleged characteristics and behaviours of a Jewish social category in terms of its position within the social structure. The analogical view allows for a similar conclusion: Provided that metaphor comprehension relies on a comparison process between the embedding structures of two concepts, metaphors draw the attention towards the location of a source and a target concept within our domains of knowledge. By this means, any interpretation as concerns antisemitic metaphors can only arise when Jews as a social group are initially located within the social structure and the constraining relations towards other social categories are identified. Given the increased awareness of the social hierarchy's deterministic nature, people may subsequently be more inclined to reason structurally about a Jewish social category. Note that the structural constraint on social categorization comes with an important particularity: Social categories are often confounded with their structural positions so that properties originating from the social structure are lined up with the social category's intrinsic features. The reason for this has been traced back to a lack of variation: People have trouble to keep the two representational levels apart when categories invariably occupy the same position within an abstract structure. Antisemitic

metaphors are no help in this respect. From the categorization point of view, they do not provide variation since a metaphorical structure is generated on the fly and comprises solely of a target, a source and a superordinate metaphorical category. In other terms, metaphorical structures only arise out of a specific context so that their elements appear rather fixed. There is a second way in which the categorization view casts light on the question how antisemitic metaphors can reinforce confounding: Properties of the ad hoc metaphorical concept are generally derived from an interaction between the source and the target concept. Accordingly, when such properties are attributed to the target during metaphor comprehension, they may causally relate to both the external source (i.e. the metaphorical ad concept) and to the category's internal features (i.e. its characteristic dimensions). Simply put, assuming that a metaphorical category is derived from an interaction between the source and the target concept, renders the lines between intrinsic characteristics and extrinsic features rather blurred. The analogical account likewise supports confounds between social concepts and their positions: From this point of view, metaphors do not invoke flexible structural representations with changing categories and positions. Instead, antisemitic metaphors rely on people's prior representation of the social domain and positions are merely highlighted during metaphor comprehension. Importantly, the analogical view assumes that people can extend or re-represent their conceptual domains during metaphor comprehension, however, these processes only pertain to relational predicates: While representations of the social structure may become richer and more informative as more relations are projected onto the target domain or as predicates are modified in order to yield a larger system of common relations, the actual objects of the social structure – that is the concepts that occupy positions – remain unchanged. In this regard, processing an antisemitic metaphors is yet another instance where people reason about rather fixed structural objects, thus confusing a particular social concept with its respective structural position.

Ultimately, both the analogical account and the categorization view offer perspectives on how antisemitic metaphors may promote structural reasoning about the social domain: Each of the theories on metaphor comprehension assumes that metaphors give rise to an external structure – either by retrieving an entire knowledge domain or by deriving an ad hoc subordinate-superordinate relation. As such, both theories imply that people engage in structural considerations about the target concept. Furthermore, the two views propose underlying comprehension mechanisms which highlight relations giving rise to the ascribed properties of a target category. Put differently, when people process a metaphor, they make inferences about a target concept or attribute properties based on the target's structural position and its connections to other concepts. At the same time, such properties and inferences are ascribed to

the target concept itself rather than to the abstract structural position. In this way, extrinsic properties emerging from the external structure are aligned with the target's internal characteristics. On these grounds, it can be argued that antisemitic metaphors do not only provide conceptual content in the form of negative associations. They additionally shape the structure of our social representations by supplying a particular style of reasoning: Comprehension of an antisemitic metaphor requires that people locate a Jewish social category within an external structure, identify the structure's constraining relations and, along these lines, infer properties to an entire Jewish social group. Even when the actual interpretation of an antisemitic metaphor can deliberately be countered by a person's desire for parity, the effects of structural reasoning are more subtle and rather difficult to mitigate: Representations of a Jewish social group become increasingly more homogeneous as regards their defining properties, they include sharp and immutable category boundaries towards others in the social structure and, furthermore, bear great inductive potential for all category members. In that sense, antisemitic metaphors are powerful instruments to make changes to our social knowledge and, thus, provide the foundation for negative downstream consequences of social categorization such as stereotyping, prejudice endorsement as well as antisemitic discrimination, hatred and dehumanization.

4. Conclusion

The overarching goal of this master's thesis was to mark out the ways in which antisemitic metaphors exert a fundamental impact on our mental representation of a Jewish social category. By analysing the cognitive effects of the two linguistic devices combined in antisemitic metaphors – that is, generic and metaphorical language – it has been shown that they provide two separate roots along which people further reason and think about Jews as a social group. In that sense, antisemitic metaphors play a crucial role in creating and maintaining biases against Jewish people and foster further detrimental by-products of social categorization.

Firstly, it has been outlined how generics correspond to one of the cognitive system's most basic, low-level mechanisms which generalizes information by default. As such, generic language gives a strong impetus for quick and efficient category learning so that the associations established between a property and a Jewish social group by means of an antisemitic metaphor are more readily accepted and added to a mental representation. Once consolidated, such forms of generic knowledge are no longer checked against further

information so that they become fairly robust to contrary evidence. However, assuming that most people are capable of mediating the content of antisemitic metaphors via regulating mechanisms and processes of cognitive control, a second, more implicit effect of generic language comes to the fore: Generics generally promote so-called internalist constraints on social category representation which highlight the restrictive nature of the category's inherent features: When faced with generic language, people are more prone to endorse essentialist beliefs, they prefer to explain characteristics and behaviours with reference to inherent causes and tend to relate properties and categories by use of different types of connections, all of which represent a property as a central and informative aspect of a social kind. These well-documented effects on people's category representations support the assumption that antisemitic metaphors have a profound impact on the structure of a Jewish social category, rendering its members increasingly homogeneous as regards their properties, and the boundaries towards other social groups sharp and impermeable.

Secondly, we have seen that antisemitic metaphors influence conceptual structure by virtue of being the things that they are, namely metaphors. It has been mapped out how different computational mechanisms of metaphor comprehension may endow externalist reasoning about the social domain. Accordingly, metaphors influence a Jewish social concept by boosting so-called externalist constraints on social categorization which focus on restrictive factors originating from an outside source. Externalist constraints crop up as social deterministic beliefs but, more importantly, they include a particular style of reasoning called structural thinking. Parallels between such a structural constraint and the different mechanisms of metaphor comprehension suggest that antisemitic metaphors supply a second way of reasoning about a Jewish social category: Arguably, encountering an antisemitic metaphor prompts people to think about a Jewish category with regard to its location in the social hierarchy and ponder on the deterministic nature of society on the category's social character. As a result, a Jewish social category is represented with an increasingly rich structure, sharp boundaries towards other elements in the social hierarchy, and immutable and stable properties across category members. Based on such highly structured representations, people may draw further inferences and generalizations, thus becoming more susceptible to the negative downstream consequences of social categorization.

Crucially, distinguishing between the impact of generic language and the effects of metaphorical reasoning is not an "either/or" question. We can rather assume that genericity and metaphoricity simultaneously unfold their influence on the representation of a Jewish social category. However, the two components may be weighted differently for each particular

instance of an antisemitic metaphor: Effects of the generic component may, for instance, be more pronounced in cases where a metaphorical interpretation merely underpins another generic property ascription (e.g. when the metaphorical source is an attributive adjective). Metaphoricity, in turn, may be more decisive in purely metaphorical property ascriptions (e.g. in standard nominal metaphors as discussed above). At the same time, the effects of the metaphorical component may decrease when a metaphorical source becomes more conventionalized and opaque over the course of time. In such cases, genericity may step forward so that the corresponding internalist constraints on social categorization outweigh the effects of metaphoricity and structural thinking.

At last, two characteristics of antisemitic metaphors which call for further investigation shall be mentioned: For one thing, antisemitic metaphors rarely come in the form of single and isolated statements. They are usually embedded in rich linguistic or non-linguistic contexts and carry a variety of hostile and offensive interpretations without committing to one in particular. In this respect, considerations about context as well as the pragmatic function of metaphors should contribute to a more comprehensive understanding of the phenomenon (Holyoak & Stamenkovic, 2018). Secondly, while in this thesis, only aspects of "cold" cognition have been discussed, an extensive analysis should also take into consideration the role of emotion and affect. Metaphors typically arouse an affective response and, in the case of antisemitic metaphors, such responses are highly negative: People may experience a threat towards the distorted image of a Jewish social group, they may fear to appear prejudiced or be anxious about not living up to their own egalitarian and pro-social beliefs. In any case, such affective responses have been identified as central elements of implicit prejudice endorsement (for a review e.g. Amodio, 2014) and as an important means to modulate memory and cognitive functions (for a review see e.g. Tyng, Amin, Saad, & Malik, 2017). Investigating the effects of emotional valence and arousal would thus give more insight into how antisemitic metaphors impart social knowledge representations, how they may influence intuitive judgements and behaviours and, furthermore, prepare the ground for stereotypes, prejudices and further negative downstream consequences in each and every one of us.

References

Ahn, W. K., Flanagan, E. H., Marsh, J. K., & Sanislow, C. A. (2006). Beliefs About Essences and the Reality of Mental Disorders. *Psychological Science*, *17*(9), 759–766. https://doi.org/10.1111/j.1467-9280.2006.01779.x

Ahn, W. K., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal Status as a Determinant of Feature Centrality. *Cognitive Psychology*, *41*(4), 361–416. https://doi.org/10.1006/cogp.2000.0741

Allport, G. (1954). *The Nature of Prejudice*. Reading, MA: Addison-Wesley Publishing Company.

Amodio, D. M. (2014). The Neuroscience of Prejudice and Stereotyping. *Nature Reviews Neuroscience*, *15*(10), 670–682. https://doi.org/10.1038/nrn3800

Aristotle (1924). Metaphysics. In W. D. Ross (Ed.), *Aristotle's Metaphysics (Vol. 2)*. Oxford, UK: The Clarendon Press (Original work published 350 BC).

Aristotle (1971). Poetics. In J. A. Smith & W. D. Ross (Eds.), *The Works of Aristotle (Vol. 11)*. Oxford, UK: The Clarendon Press (Original work published 335 BC).

Asher, N., & Pelletier, F. J. (2013). More Truths About Generic Truth. In A. Mari, C. Beyssade, & F. Del Prete (Eds.), *Genericity* (pp. 312–333). https://doi.org/https://doi.org/10.1093/acprof:oso/9780199691807.003.0012

Astuti, R., Solomon, G. E. A., Carey, S., Ingold, T., & Miller, P. H. (2004). Constraints on Conceptual Development: A Case Study of the Acquisition of Folkbiological and Folksociological Knowledge in Madagascar. In *Monographs of the Society for Research in Child Development* (Vol. 69). https://doi.org/doi: 10.1111/j.0037-976X.2004.00296.x.

Atran, S. (1998). Folk Biology and the Anthropology of Science: Cognitive Universals and Cultural Particulars. *Behavioral and Brain Sciences*, *21*(4), 547–609. https://doi.org/10.1017/S0140525X98001277

Atran, S., Estin, P., Coley, J. D., & Medin, D. L. (1997). Generic Species and Basic Levels: Essence and Appearance in Folk Biology. *Journal of Ethnobiology*, *17*(1), 17–43.

Atran, S., Medin, D. L., & Ross, N. (2002). Thinking About Biology: Modular Constraints on Categorization and Reasoning in the Everyday Life of Americans, Maya, and Scientists. *Mind & Society*, *6*, 31–61. https://doi.org/https://doi.org/10.1007/BF02513147

Bach, T. (2014). A Unified Account of General Learning Mechanisms and Theory-of-Mind Development. *Mind and Language*, *29*(3), 351–381. https://doi.org/10.1111/mila.12055

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infants' Ability to Draw Inferences About Nonobvious Object Properties: Evidence From Exploratory Play. *Child Development*, *64*(3), 711–728. https://doi.org/https://doi.org/10.2307/1131213

Baron, A. S., Dunham, Y., Banaji, M., & Carey, S. (2014). Constraints on the Acquisition of Social Category Concepts. *Journal of Cognition and Development*, *15*(2), 238–268. https://doi.org/10.1080/15248372.2012.742902

Barr, R. A., & Caplan, L. J. (1985). Two Kinds of Feature? A Test of Two Theories of Typicality Effects in Natural Language Categories. *Proceedings of the Seventh Annual Conference of the Cognitive Science Society*, 181–189. Hillsdale, NJ: Erlbaum.

Barr, R. A., & Caplan, L. J. (1987). Category Representations and Their Implications for Category Structure. *Memory & Cognition*, *15*(5), 397–418. https://doi.org/10.3758/BF03197730

Barsalou, L. W. (1983). Ad Hoc Categories. *Memory & Cognition*, *11*(3), 211–227. https://doi.org/https://doi.org/10.3758/bf03196968

Barsalou, L. W. (1991). Deriving Categories to Achieve Goals. *The Psychology of Learning and Motivation*, *27*, 1–64. https://doi.org/10.1016/S0079-7421(08)60120-6

Bastian, B., & Haslam, N. (2006). Psychological Essentialism and Stereotype Endorsement. *Journal of Experimental Social Psychology*, *42*(2), 228–235. https://doi.org/10.1016/j.jesp.2005.03.003

Behrens, L. (2005). Genericity From a Cross-Linguistic Perspective. *Linguistics*, *43*(2), 275–344. https://doi.org/10.1515/ling.2005.43.2.275

Benelli, B. (1988). If It Is a Dog, Can It Be an Animal? The Role of Metalinguistic Knowledge in the Acquisition of Linguistic Superordination. *Journal of Psycholinguistic Research*, *17*(3), 227–243. https://doi.org/10.1007/BF01686357

Black, M. (1962). *Models and Metaphors: Studies in Language and Philosophy*. Ithaca, NY: Cornell University Press.

Bowdle, B. F., & Gentner, D. (2005). The Career of Metaphor. *Psychological Review*, *112*(1), 193–216. https://doi.org/10.1037/0033-295X.112.1.193

Brook, J. S. (1970). A Test of Piaget's Theory of "Nominal Realism." *The Journal of Genetic Psychology: Research and Theory on Human Development*, *116*(2), 165–175. https://doi.org/10.1080/00221325.1970.10533916

Bursten, J. R. (2014). Microstructure Without Essentialism: A New Perspective on Chemical Classification. *Philosophy of Science*, *81*(4), 633–653. https://doi.org/10.1086/678043

Cacchione, T., Hrubesch, C., Call, J., & Rakoczy, H. (2016). Are Apes Essentialists? Scope and Limits of Psychological Essentialism in Great Apes. *Animal Cognition*, *19*(5), 921–937. https://doi.org/10.1007/s10071-016-0991-4

Cacciari, C., Levorato, M. C., & Cicogna, P. (1997). Imagination at Work: Conceptual and Linguistic Creativity in Children. In T. B. Ward, S. M. Smith, & J. Vaid (Eds.), *Creative Thought: An Investigation of Conceptual Structures and Processes* (pp. 145–177). https://doi.org/https://doi.org/10.1037/10227-007

Carey, S. (1996). Cognitive Domains as Modes of Thought. In D. R. Olson & N. Torrance (Eds.), *Modes of Thought: Explorations in Culture and Cognition* (pp. 187–215). Cambridge, MA: Cambridge University Press.

Carey, S. (2004). Bootstrapping and the Origin of Concepts. *Daedalus*, *133*(1), 59–68. https://doi.org/10.1162/001152604772746701

Carey, S., & Xu, F. (1999). Sortals and Kinds: An Appreciation of John Macnamara. In R. Jackendoff, P. Bloom, & K. Wynn (Eds.), *Language, Logic, and Concepts: Essays in Memory of John Macnamara* (pp. 311–335). https://doi.org/https://doi.org/10.7551/mitpress/4118.003.0016

Carlson, G. (1977). A Unified Analysis of the English Bare Plural. *Linguistics and Philosophy*, *1*(3), 413–457. https://doi.org/https://doi.org/10.1007/BF00353456

Carlson, G. (1980). *Rererence to Kinds in English*. New York, NY, London: Garland Publishing Inc.

Carlson, G. (1989). On the Semantic Composition of English Generic Sentences. In G. Chierchia, B. H. Partee, & R. Turner (Eds.), *Properties, Types and Meaning. Studies in Linguistics and Philosophy* (pp. 167–192). https://doi.org/https://doi.org/10.1007/978-94-009-2723-0_5

Carlson, G. (1995). Truth-Conditions of Generic Sentences: Two Contrasting Views. In G. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 224–237). University of Chicago Press.

Carlson, G. (2010). Generics and Concepts. In F. J. Pelletier (Ed.), *Kinds, Things, and Stuff: Mass Terms and Generics*. https://doi.org/10.1093/acprof:oso/9780195382891.003.0002

Carlson, G. (2019). Genericity. In P. Portner, K. Von Heusinger, & C. M. Eds (Eds.), *Semantics. Noun Phrases and Verb Phrases* (pp. 232–273). https://doi.org/https://doi.org/10.1515/9783110589443-008

Chalik, L., Leslie, S.-J., & Rhodes, M. (2017). Cultural Context Shapes Essentialist Beliefs About Religion. *Developmental Psychology*, *53*(6), 1178–1187. https://doi.org/10.1037/dev0000301

Christie, S. (2017). Structure Mapping for Social Learning. *Topics in Cognitive Science*, *9*(3), 758–775. https://doi.org/10.1111/tops.12264

Cimpian, A. (2015). The Inherence Heuristic: Generating Everyday Explanations. In R. Scott & S. Kosslyn (Eds.), *Emerging Trends in the Social and Behavioral Sciences* (pp. 1–15). https://doi.org/10.1002/9781118900772.etrds0341

Cimpian, A., Arce, H. M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle Linguistic Cues Affect Children's Motivation. *Psychological Science*, *18*(4), 314–316. https://doi.org/10.1111/j.1467-9280.2007.01896.x

Cimpian, A., Brandone, A. C., & Gelman, S. A. (2010). Generic Statements Require Little Evidence for Acceptance but Have Powerful Implications. *Cognitive Science*, *34*(8), 1452–1482. https://doi.org/DOI:10.1111/J.1551-6709.2010.01126.X

Cimpian, A., & Cadena, C. (2010). Why Are Dunkels Sticky? Preschoolers Infer Functionality and Intentional Creation for Artifact Properties Learned From Generic Language. *Cognition*, *117*(1), 62–68. https://doi.org/10.1016/j.cognition.2010.06.011

Cimpian, A., & Erickson, L. C. (2012). The Effect of Generic Statements on Children's Causal Attributions: Questions of Mechanism. *Developmental Psychology*, *48*(1), 159–170. https://doi.org/10.1037/a0025274

Cimpian, A., & Markman, E. M. (2009). Information Learned From Generic Language Becomes Central to Children's Biological Concepts: Evidence From Their Open-Ended Explanations. *Cognition*, *113*(1), 14–25. https://doi.org/10.1016/j.cognition.2009.07.004

Cimpian, A., & Markman, E. M. (2011). The Generic/Nongeneric Distinction Influences How Children Interpret New Information About Social Others. *Child Development*, *82*(2), 471–492. https://doi.org/DOI: 10.1111/j.1467-8624.2010.01525.x

Cimpian, A., & Salomon, E. (2014). The Inherence Heuristic: An Intuitive Means of Making Sense of the World and a Potential Precursor to Psychological Essentialism. *Behavioral and Brain Sciences*, *37*(5), 461–480. https://doi.org/10.1017/S0140525X13002197

Cohen, A. (1999). Generics, Frequency Adverbs, and Probability. *Linguistics and Philosophy*, *22*(3), 221–253.

Cohen, A. (2002). Genericity. In T. E. Zimmermann & F. Hamm (Eds.), *Semantics. Linguistische Berichte* (pp. 59–89). Hamburg: Buske.

Cohen, A. (2004). Generics and Mental Representations. *Linguistics and Philosophy*, *27*(5), 529–556. https://doi.org/https://doi.org/10.1023/b:ling.0000033851.25870.3e

Cohen, A. (2007). The Information Structure of Bare Plurals in English and Italian. In K. Schwabe & S. Winkler (Eds.), *On Information Structure, Meaning and Form: Generalizations Across Languages* (pp. 509–521). https://doi.org/10.1075/la.100.27coh

Coley, J. D., Medin, D. L., & Atran, S. (1997). Does Rank Have Its Privilege? Inductive Inferences Within Folkbiological Taxonomies. *Cognition*, *64*(1), 73–112. https://doi.org/10.1016/S0010-0277(97)00017-6

Collins, J. (2018). Genericity Sans Gen. *Mind and Language*, *33*(1), 34–64. https://doi.org/10.1111/mila.12168

Cozzarelli, C., Wilkinson, A. V., & Tagler, M. J. (2001). Attitudes Toward the Poor and Attributions for Poverty. *Journal of Social Issues*, *57*(2), 207–227. https://doi.org/10.1111/0022-4537.00209

Dahl, Ö. (1975). On Generics. In E. L. Keenan (Ed.), *Formal Semantics of Natural Language* (pp. 99–111). https://doi.org/https://doi.org/10.1017/cbo9780511897696.009

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic Essentialism: On the Deceptive Determinism of DNA. *Psychological Bulletin*, *137*(5), 800–818. https://doi.org/10.1037/a0021860

Davidson, D. (1978). What Metaphors Mean. *Critical Inquiry*, *5*(1), 31–47. https://doi.org/https://doi.org/10.1086/447971
Declerck, R. (1986). The Manifold Interpretations of Generic Sentences. *Lingua*, *68*, 149–188. https://doi.org/https://doi.org/10.1016/0024-3841(86)90002-1
Declerck, R. (1991). The Origins of Genericity. *Linguistics*, *29*(1), 79–102. https://doi.org/10.1515/ling.1991.29.1.79
Diesendruck, G. (2003). Categories for Names or Names for Categories? The Interplay Between Domain-Specific Conceptual Structure and Language. *Language and Cognitive Processes*, *18*(5/6), 759–787. https://doi.org/10.1080/01690960344000116
Diesendruck, G., & Gelman, S. A. (1999). Domain Differences in Absolute Judgments of Category Membership: Evidence for an Essentialist Account of Categorization. *Psychonomic Bulletin and Review*, *6*(2), 338–346. https://doi.org/10.3758/BF03212339
Diesendruck, G., Goldfein-Elbaz, R., Rhodes, M., Gelman, S. A., & Neumark, N. (2013). Cross-Cultural Differences in Children's Beliefs About the Objectivity of Social Categories. *Child Development*, *84*(6), 1906–1917. https://doi.org/10.1111/cdev.12108
Diesendruck, G., & HaLevi, H. (2006). The Role of Language, Appearance, and Culture in Children's Social Category-Based Induction. *Child Development*, *77*(3), 539–553. https://doi.org/https://doi.org/10.1111/j.1467-8624.2006.00889.x
Diesendruck, G., & Menahem, R. (2015). Essentialism Promotes Children's Inter-Ethnic Bias. *Frontiers in Psychology*, *6*, 1–8. https://doi.org/10.3389/fpsyg.2015.01180
Duncan, M. L., & Kemmelmeier, M. (2012). Attitudes Toward Same-Sex Marriage: An Essentialist Approach. *Analyses of Social Issues and Public Policy*, *12*(1), 377–399. https://doi.org/10.1111/j.1530-2415.2012.01290.x
Eidson, R. C., & Coley, J. D. (2014). Not So Fast: Reassessing Gender Essentialism in Young Adults. *Journal of Cognition and Development*, *15*(2), 382–392. https://doi.org/10.1080/15248372.2013.763810
Ellis, B. D. (2014). *The Philosophy of Nature. A Guide to the New Essentialism*. https://doi.org/https://doi.org/10.4324/9781315710624
Estes, Z. (2003). Domain Differences in the Structure of Artifactual and Natural Categories. *Memory and Cognition*, *31*(2), 199–214. https://doi.org/10.3758/BF03194379
Evans, J. S. B. T. (2003). In Two Minds: Dual-Process Accounts of Reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459. https://doi.org/10.1016/j.tics.2003.08.012
Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685
Fiske, S. T., & Russell, A. M. (2010). Cognitive Processes. In J. F. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The Sage Handbook of Prejudice, Stereotyping and Discrimination* (pp. 115–130). https://doi.org/10.1007/978-3-319-69188-6_8
Frassanito, P., & Pettorini, B. (2008). Pink and Blue: The Color of Gender. *Child's Nervous System*, *24*(8), 881–882. https://doi.org/10.1007/s00381-007-0559-3
Gelman, S. A. (2003). *The Essential Child: Origins of Essentialism in Everyday Thought*. https://doi.org/10.1093/acprof:oso/9780195154061.001.0001
Gelman, S. A. (2004). Psychological Essentialism in Children. *Trends in Cognitive Sciences*, *8*(9), 404–409. https://doi.org/10.1016/j.tics.2004.07.001
Gelman, S. A. (2010). Generics as a Window Onto Young Children's Concepts. In F. J. Pelletier (Ed.), *Kinds, Things, and Stuff: Mass Terms and Generics*. https://doi.org/10.1093/acprof:oso/9780195382891.003.0006
Gelman, S. A. (2013). Artifacts and Essentialism. *Review of Philosophy and Psychology*, *4*(3), 449–463. https://doi.org/doi:10.1007/s13164-013-0142-7
Gelman, S. A., & Bloom, P. (2007). Developmental Changes in the Understanding of Generics. *Cognition*, *105*(1), 166–183. https://doi.org/10.1016/j.cognition.2006.09.009

Gelman, S. A., Cimpian, A., & Roberts, S. O. (2018). How Deep Do We Dig? Formal Explanations as Placeholders for Inherent Explanations. *Cognitive Psychology*, *106*, 43–59. https://doi.org/10.1016/j.cogpsych.2018.08.002

Gelman, S. A., Coley, J. D., Rosengren, K. S., Hartman, E., Pappas, A., & Keil, F. C. (1998). Beyond Labeling: The Role of Maternal Input in the Acquisition of Richly Structured Categories. In *Monographs of the Society for Research in Child Development* (Vol. 63). https://doi.org/https://doi.org/10.2307/1166211

Gelman, S. A., & Diesendruck, G. (1999). A Reconsideration of Concepts: On the Compatibility of Psychological Essentialism and Context Sensitivity. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. H. Miller (Eds.), *Conceptual Development: Piaget's Legacy* (pp. 79–102). Mahwah, NJ: Erlbaum.

Gelman, S. A., Goetz, P. J., Sarnecka, B. W., & Flukes, J. (2008). Generic Language in Parent-Child Conversations. *Language Learning and Development*, *4*(1), 1–31. https://doi.org/https://doi.org/10.1080/15475440701542625

Gelman, S. A., Leslie, S.-J., Was, A. M., & Koch, C. M. (2015). Children's Interpretations of General Quantifiers, Specific Quantifiers, and Generics. *Language, Cognition and Neuroscience*, *30*(4), 448–461. https://doi.org/https://doi.org/10.1080/23273798.2014.931591

Gelman, S. A., & Roberts, S. O. (2017). How Language Shapes the Cultural Inheritance of Categories. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(30), 7900–7907. https://doi.org/10.1073/pnas.1621073114

Gelman, S. A., & Ware, E. A. (2012). Conceptual Development: The Case of Essentialism. In E. Margolis, R. Samuels, & S. P. Stich (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (pp. 1–30). https://doi.org/10.1093/oxfordhb/9780195309799.013.0019

Gelman, S. A., Ware, E. A., & Kleinberg, F. (2010). Effects of Generic Language on Category Content and Structure. *Cognitive Psychology*, *61*(3), 273–301. https://doi.org/10.1016/j.cogpsych.2010.06.001

Gelman, S. A., & Wellman, H. M. (1991). Insides and Essences: Early Understandings of the Non-Obvious. *Cognition*, *38*(3), 213–244. https://doi.org/10.1016/0010-0277(91)90007-Q

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, *7*, 155–170. https://doi.org/10.1016/B978-1-4832-1446-7.50026-1

Gentner, D. (1988). Metaphor as Structure Mapping: The Relational Shift. *Child Development*, *59*(1), 47–59. https://doi.org/https://doi.org/10.2307/1130388

Gentner, D. (2003). Why We're So Smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in Mind: Advances in the Study of Language and Thought* (pp. 195–235). https://doi.org/https://doi.org/10.7551/mitpress/4117.003.0015

Gentner, D. (2005). The Development of Relational Category Knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building Object Categories in Developmental Time* (pp. 245–275). https://doi.org/10.4324/9781410612908

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure Mapping and Relational Language Support Children's Learning of Relational Categories. *Child Development*, *82*(4), 1173–1188. https://doi.org/10.1111/j.1467-8624.2011.01599.x

Gentner, D., & Asmuth, J. (2017). Metaphoric Extension, Relational Categories, and Abstraction. *Language, Cognition and Neuroscience*, *34*(10), 1298–1307. https://doi.org/10.1080/23273798.2017.1410560

Gentner, D., & Bowdle, B. F. (2008). Metaphor as Structure-Mapping. In R. W. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 109–128). https://doi.org/10.1017/cbo9780511816802.008

Gentner, D., & Clement, C. (1988). Evidence for Relational Selectivity in the Interpretation of Analogy and Metaphor. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 22, pp. 307–358). https://doi.org/10.1016/S0079-7421(08)60044-4

Gentner, D., & Grudin, J. (1985). The Evolution of Mental Metaphors in Psychology. A 90-Year Retrospective. *American Psychologist*, *40*(2), 181–192. https://doi.org/10.1037/0003-066X.40.2.181

Gentner, D., & Hoyos, C. (2017). Analogy and Abstraction. *Topics in Cognitive Science*, *9*(3), 672–693. https://doi.org/10.1111/tops.12278

Gentner, D., & Kurtz, K. J. (2005). Relational Categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, & A. B. Markman (Eds.), *Categorization Inside and Outside the Laboratory: Essays in Honor of Douglas L. Medin* (pp. 151–175). https://doi.org/10.1037/11156-009

Gentner, D., & Markman, A. B. (1997). Structure Mapping in Analogy and Similarity. *American Psychologist*, *52*(1), 45–56. https://doi.org/10.1037/0003-066X.52.1.45

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the Intentional Stance at 12 Months of Age. *Cognition*, *56*(2), 165–193. https://doi.org/10.1016/0010-0277(95)00661-H

Gerson, S. A., & Woodward, A. L. (2012). A Claw Is Like My Hand: Comparison Supports Goal Analysis in Infants. *Cognition*, *122*(2), 181–192. https://doi.org/doi:10.1016/j.cognition.2011.10.014

Geurts, B. (1985). Generics. *Journal of Semantics*, *4*, 247–255. https://doi.org/https://doi.org/10.1093/jos/4.3.247

Gibbs, R. W. (1992). When Is Metaphor? The Idea of Understanding in Theories of Metaphor. *Poetics Today*, *13*(4), 575–606. https://doi.org/doi:10.2307/1773290

Gibbs, R. W. (1994). The Poetics of Mind: Figurative Thought, Language, and Understanding. In *Figurative Thought, Language, and Understanding*. Cambridge: Cambridge University Press.

Gibbs, R. W. (2011). Evaluating Conceptual Metaphor Theory. *Discourse Processes*, *48*(8), 529–562. https://doi.org/10.1080/0163853X.2011.606103

Gil-White, F. J. (2001). Are Ethnic Groups Biological "Species" to the Human Brain? Essentialism in Our Cognition of Some Social Categories. *Current Anthropology*, *42*(4), 515–554. https://doi.org/10.2307/3596550

Glucksberg, S., Gildea, P., & Bookin, H. B. (1982). On Understanding Nonliteral Speech: Can People Ignore Metaphors? *Journal of Verbal Learning and Verbal Behavior*, *21*(1), 85–98. https://doi.org/10.1016/S0022-5371(82)90467-4

Glucksberg, S., & Haught, C. (2006). Can Florida Become Like the Next Florida? When Metaphoric Comparisons Fail. *Psychological Science*, *17*(11), 935–938. https://doi.org/10.1111/j.1467-9280.2006.01807.x

Glucksberg, S., & Keysar, B. (1990). Understanding Metaphorical Comparisons: Beyond Similarity. *Psychological Review*, *97*(1), 3–18. https://doi.org/10.1037/0033-295X.97.1.3

Glucksberg, S., McGlone, M. S., & Manfredi, D. (1997). Property Attribution in Metaphor Comprehension. *Journal of Memory and Language*, *36*(1), 50–67. https://doi.org/10.1006/jmla.1996.2479

Goldin-Meadow, S., Gelman, S. A., & Mylander, C. (2005). Expressing Generic Concepts With and Without a Language Model. *Cognition*, *96*(2), 109–126. https://doi.org/10.1016/j.cognition.2004.07.003

Goldwater, M. B., Bainbridge, R., & Murphy, G. L. (2016). Learning of Role-Governed and Thematic Categories. *Acta Psychologica*, *164*, 112–126. https://doi.org/10.1016/j.actpsy.2015.10.011

Goldwater, M. B., Markman, A. B., & Stilwell, C. H. (2011). The Empirical Case for Role-Governed Categories. *Cognition*, *118*(3), 359–376. https://doi.org/10.1016/j.cognition.2010.10.009
Gopnik, A. (1998). Explanation as Orgasm. *Minds and Machines*, *8*(1), 101–118. https://doi.org/10.1023/A:1008290415597
Gopnik, A., & Meltzoff, A. N. (1997). *Words, Thoughts, and Theories*. https://doi.org/https://doi.org/10.7551/mitpress/7289.001.0001
Gottfried, G. M., & Gelman, S. A. (2005). Developing Domain-Specific Causal-Explanatory Frameworks: The Role of Insides and Immanence. *Cognitive Development*, *20*(1), 137–158. https://doi.org/10.1016/j.cogdev.2004.07.003
Grady, J. E. (1997). *Foundations of Meaning: Primary Metaphors and Primary Scenes*. Universitiy of California, Berkely.
Graham, S. A., Gelman, S. A., & Clarke, J. (2016). Generics License 30-Month-Olds' Inferences About the Atypical Properties of Novel Kinds. *Developmental Psychology*, *52*(9), 1353–1362. https://doi.org/10.1037/dev0000183
Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-Month-Olds Rely on Shared Labels and Shape Similarity for Inductive Inferences. *Child Development*, *75*(2), 409–427. https://doi.org/https://doi.org/10.1111/j.1467-8624.2004.00683.x
Graham, S. A., Nayer, S. L., & Gelman, S. A. (2011). Two-Year-Olds Use the Generic/Nongeneric Distinction to Guide Their Inferences About Novel Kinds. *Child Development*, *82*(2), 493–507. https://doi.org/DOI: 10.1111/j.1467-8624.2010.01572.x
Haryu, E., Imai, M., & Okada, H. (2011). Object Similarity Bootstraps Young Children to Action-Based Verb Extension. *Child Development*, *82*(2), 674–686. https://doi.org/10.1111/j.1467-8624.2010.01567.x
Haslam, N., & Abou-Abdallah, M. (2015). Essentialism. In *International Encyclopedia of the Social & Behavioral Sciences* (Vol. 8, pp. 13–15). https://doi.org/https://doi.org/10.1016/B978-0-08-097086-8.24059-2
Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist Beliefs About Personality and Their Implications. *Personality and Social Psychology Bulletin*, *30*(12), 1661–1673. https://doi.org/10.1177/0146167204271182
Haslam, N., & Levy, S. R. (2006). Essentialist Beliefs About Homosexuality: Structure and Implications for Prejudice. *Personality and Social Psychology Bulletin*, *32*(4), 471–485. https://doi.org/10.1177/0146167205276516
Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist Beliefs About Social Categories. *British Journal of Social Psychology*, *39*(1), 113–127. https://doi.org/10.1348/014466600164363
Haslam, N., Rothschild, L., & Ernst, D. (2002). Are Essentialist Beliefs Associated With Prejudice? *British Journal of Social Psychology*, *41*(1), 87–100. https://doi.org/10.1348/014466602165072
Haslanger, S. (2016). What Is a (Social) Structural Explanation? *Philosophical Studies*, *173*(1), 113–130. https://doi.org/10.1007/s11098-014-0434-5
Haward, P., Wagner, L., Carey, S., & Prasada, S. (2018). The Development of Principled Connections and Kind Representations. *Cognition*, *176*, 255–268. https://doi.org/10.1016/j.cognition.2018.02.001
Heine, S. J., Dar-Nimrod, I., Cheung, B. Y., & Proulx, T. (2017). Essentially Biased: Why People Are Fatalistic About Genes. In J. M. Olsen (Ed.), *Advances in Experimental Social Psychology* (Vol. 55, pp. 137–192). https://doi.org/10.1016/bs.aesp.2016.10.003
Heit, E. (2000). Properties of Inductive Reasoning. *Psychonomic Bulletin and Review*, *7*(4), 569–592. https://doi.org/10.3758/BF03212996
Heyer, G. (1990). Semantics and Knowledge Representation in the Analysis of Generic Descriptions. *Journal of Semantics*, *7*(1), 93–110. https://doi.org/10.1093/jos/7.1.93

Heyman, G. D., & Gelman, S. A. (2000). Preschool Children's Use of Novel Predicates to Make Inductive Inferences About People. *Cognitive Development*, *15*(3), 263–280. https://doi.org/10.1016/S0885-2014(00)00028-9

Hirschfeld, L. A. (1996). *Race in the Making: Cognition, Culture, and the Child's Construction of Human Kinds*. Cambridge, MA; London, UK: MIT Press.

Hirschfeld, L. A. (2007). Folksociology and the Cognitive Foundations of Culture. *Intellectica*, *46–47*(2–3), 191–206. https://doi.org/10.3406/intel.2007.1284

Hollander, M. A., Gelman, S. A., & Raman, L. (2009). Generic Language and Judgements About Category Membership: Can Generics Highlight Properties as Central? *Language and Cognitive Processes*, *24*(4), 481–505. https://doi.org/doi: 10.1080/01690960802223485

Hollander, M. A., Gelman, S. A., & Star, J. (2002). Children's Interpretation of Generic Noun Phrases. *Developmental Psychology*, *38*(6), 883–894. https://doi.org/10.1037/0012-1649.38.6.883

Holyoak, K. J., & Stamenkovic, D. (2018). Metaphor Comprehension: A Critical Review of Theories and Evidence. *Psychological Bulletin*, *144*(6), 641–671. https://doi.org/10.1037/bul0000145

Horowitz, A. C., & Frank, M. C. (2016). Children's Pragmatic Inferences as a Route for Learning About the World. *Child Development*, *87*(3), 807–819. https://doi.org/10.1111/cdev.12527

Hoyos, C., Horton, W. S., & Gentner, D. (2015). Analogical Comparison Aids False Belief Understanding in Preschoolers. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th Annual Conference of the Cognitive Science Society* (pp. 1–6). Austin, TX: Cognitive Science Society.

Jaswal, V. K., & Markman, E. M. (2007). Looks Aren't Everything: 24- Month-Olds' Willingness to Accept Unexpected Labels. *Journal of Cognition and Development*, *8*(1), 93–111. https://doi.org/10.1080/15248370709336995

Jayaratne, T. E., Gelman, S. A., Feldbaum, M., Sheldon, J. P., Petty, E. M., & Kardia, S. L. R. (2009). The Perennial Debate: Nature, Nurture, or Choice? Black and White Americans' Explanations for Individual Differences. *Review of General Psychology*, *13*(1), 24–33. https://doi.org/10.1037/a0014227

Johnson, M. G., & Malgady, R. G. (1979). Some Cognitive Aspects of Figurative Language: Association and Metaphor. *Journal of Psycholinguistic Research*, *8*(3), 249–265. https://doi.org/10.1007/BF01067307

Jones, L. L., & Estes, Z. (2005). Metaphor Comprehension as Attributive Categorization. *Journal of Memory and Language*, *53*(1), 110–124. https://doi.org/10.1016/j.jml.2005.01.016

Jost, J. T., & Kay, A. C. (2005). Exposure to Benevolent Sexism and Complementary Gender Stereotypes: Consequences for Specific and Diffuse Dorms of System Justification. *Journal of Personality and Social Psychology*, *88*(3), 498–509. https://doi.org/10.1037/0022-3514.88.3.498

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.

Kalish, C. W. (2002). Essentialist to Some Degree: Beliefs About the Structure of Natural Kind Categories. *Memory and Cognition*, *30*(3), 340–352. https://doi.org/10.3758/BF03194935

Kamins, M. L., & Dweck, C. S. (1999). Person Versus Process Praise and Criticism: Implications for Contingent Self-Worth and Coping. *Developmental Psychology*, *35*(3), 835–847. https://doi.org/10.1037/0012-1649.35.3.835

Katz, A. N. (1982). Metaphoric Relationships: The Role of Feature Saliency. *Journal of Psycholinguistic Research*, *11*(4), 283–296. https://doi.org/10.1007/BF01067583

Kelemen, D. (1999). Function, Goals and Intention: Children's Teleological Reasoning About Objects. *Trends in Cognitive Sciences*, *3*(12), 461–468. https://doi.org/10.1016/S1364-6613(99)01402-3

Kelemen, D., & Carey, S. (2007). The Essence of Artifacts: The Construction of the Design Stance. In E. Margolis & S. Laurence (Eds.), *Creations of the Mind: Theories of Artifacts and Their Representation* (pp. 212–230). Oxford; New York, NY: Oxford University Press.

Keller, J. (2005). In Genes We Trust: The Biological Component of Psychological Essentialism and Its Relationship to Mechanisms of Motivated Social Cognition. *Journal of Personality and Social Psychology*, *88*(4), 686–702. https://doi.org/10.1037/0022-3514.88.4.686

Kelley, H. H. (1973). The Processes of Causal Attribution. *American Psychologist*, *28*(2), 107–128. https://doi.org/10.1037/h0034225

Keysar, B., & Bly, B. (1995). Intuitions of the Transparency of Idioms: Can One Keep a Secret by Spilling the Beans? *Journal of Memory and Language, 34*(1), 89–109. https://doi.org/https://doi.org/10.1006/jmla.1995.1005

Kinzler, K. D., & Dautel, J. B. (2012). Children's Essentialist Reasoning About Language and Race. *Developmental Science*, *15*(1), 131–138. https://doi.org/10.1111/j.1467-7687.2011.01101.x

Kray, L. J., Howland, L., Russell, A. G., & Jackman, L. M. (2017). The Effects of Implicit Gender Role Theories on Gender System Justification: Fixed Beliefs Strengthen Masculinity to Preserve the Status Quo. *Journal of Personality and Social Psychology*, *112*(1), 98–115. https://doi.org/10.1037/pspp0000124

Krifka, M. (1987). An Outline of Genericity. In *Forschungsbericht des Seminars für natürlich-sprachliche Systeme* (pp. 87–125). Universität Tübingen.

Krifka, M., Pelletier, F. J., Carlson, G., ter Meulen, A., Link, G., & Chierchia, G. (1995). Genericity: An Introduction. In G. Carlson & F. J. Pelletier (Eds.), *The Generic Book* (pp. 1–124). Chicago University Press.

Kripke, S. (1994). Naming and Necessity. In R. M. Harnish (Ed.), *Basic Topics in the Philosophy of Language* (pp. 192–220). Englewood Cliffs, NJ: Prentice Hall (Original work published 1972).

Kuhn, T. S. (1993). Metaphor in Science. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 533–542). https://doi.org/https://doi.org/10.1017/cbo9781139173865.024

Kurtz, K. J., & Gentner, D. (2001). Kinds of Kinds: Sources of Category Coherence. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 522–527). Mahwah, NJ: Erlbaum.

Lakoff, G., & Johnson, M. (1980a). *Metaphors We Live By*. London, UK: The University of Chicago Press.

Lakoff, G., & Johnson, M. (1980b). The Metaphorical Structure of the Human Conceptual System. *Cognitive Science*, *4*(2), 195–208. https://doi.org/10.1016/S0364-0213(80)80017-6

Lazaridou-Chatzigoga, D. (2019). Genericity. In C. Cummins & N. Katsos (Eds.), *The Oxford Handbook of Experimental Semantics and Pragmatics* (pp. 156–177). https://doi.org/10.1093/oxfordhb/9780198791768.013.12

leech, n.2. (n.d.). Retrieved July 9, 2020, from Oxford English Dictionary website: www.oed.com/view/Entry/106917

Leslie, S.-J. (2007). Generics and the Stucture of the Mind. *Philosophical Perspectives*, *21*, 375–403. https://doi.org/https://doi.org/10.1111/j.1520-8583.2007.00138.x

Leslie, S.-J. (2008). Generics: Cognition and Acquisition. *Philosophical Review*, *117*(1), 1–47. https://doi.org/https://doi.org/10.1215/00318108-2007-023

Leslie, S.-J. (2013). Generics Oversimplified. *Noûs*, *49*(1), 28–54. https://doi.org/https://doi.org/10.1111/nous.12039
Leslie, S.-J., & Gelman, S. A. (2012). Quantified Statements Are Recalled as Generics: Evidence From Preschool Children and Adults. *Cognitive Psychology*, *64*(3), 186–214. https://doi.org/10.1016/j.cogpsych.2011.12.001
Leslie, S.-J., Khemlani, S., & Glucksberg, S. (2011). Do All Ducks Lay Eggs? The Generic Overgeneralization Effect. *Journal of Memory and Language*, *65*(1), 15–31. https://doi.org/10.1016/j.jml.2010.12.005
Leslie, S.-J., & Lerner, A. (2016). Generic Generalizations. Retrieved June 12, 2020, from Stanford Encyclopedia of Philosophy website: https://plato.stanford.edu/entries/generics/
Lewis, D. (1975). Adverbs of Quantification. In E. L. Keenan (Ed.), *Formal Semantics of Natural Language* (pp. 3–15). https://doi.org/https://doi.org/10.1017/CBO9780511897696
Lieberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The Origins of Social Categorization. *Trends in Cognitive Sciences*, *21*(7), 556–568. https://doi.org/http://dx.doi.org/10.1016/j.tics.2017.04.004
Liebesman, D. (2011). Simple Generics. *Noûs*, *45*(3), 409–442. https://doi.org/10.1111/j.1468-0068.2010.00774.x
Lombrozo, T. (2006). The Structure and Function of Explanations. *Trends in Cognitive Sciences*, *10*(10), 464–470. https://doi.org/10.1016/j.tics.2006.08.004
Mahalingam, R. (2003). Essentialism, Culture, and Beliefs about Gender among the Aravanis of Tamil Nadu, India. *Sex Roles*, *49*(9–10), 489–496. https://doi.org/10.1023/A:1025828606094
Mandalaywala, T. M., Amodio, D. M., & Rhodes, M. (2018). Essentialism Promotes Racial Prejudice by Increasing Endorsement of Social Hierarchies. *Social Psychological and Personality Science*, *9*(4), 461–469. https://doi.org/10.1177/1948550617707020
Mandalaywala, T. M., Ranger-Murdock, G., Amodio, D. M., & Rhodes, M. (2018). The Nature and Consequences of Essentialist Beliefs About Race in Early Childhood. *Child Development*, *90*(4), e437–e453. https://doi.org/10.1111/cdev.13008
Mari, A., Beyssade, C., & Prete, F. Del. (2013). Introduction. In F. Del Prete, C. Beyssade, & A. Mari (Eds.), *Genericity* (pp. 1–92). https://doi.org/https://doi.org/10.1093/acprof:oso/9780199691807.003.0001
Markman, A. B., & Gentner, D. (2000). Structure Mapping in the Comparison Process. *The American Journal of Psychology*, *113*(4), 501–538. https://doi.org/https://doi.org/10.2307/1423470
Markman, A. B., & Stilwell, C. H. (2001). Role-Governed Categories. *Journal of Experimental and Theoretical Artificial Intelligence*, *13*(4), 329–358. https://doi.org/10.1080/09528130110100252
McArthur, L. A. (1972). The How and What of Why: Some Determinants and Consequenes of Causal Attribution. *Journal Ol Personality and Social Psychology*, *22*(2), 171–193. https://doi.org/https://doi.org/10.1037/h0032602
McGlone, M. S. (2007). What Is the Explanatory Value of a Conceptual Metaphor? *Language and Communication*, *27*(2), 109–126. https://doi.org/10.1016/j.langcom.2006.02.016
Medin, D. L., & Ortony, A. (1989). Psychological Essentialism. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 179–195). Cambridge University Press.
Meltzoff, A. N. (2005). Imitation and Other Minds: The "Like Me" Hypothesis. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 55–77). Cambridge, MA: MIT Press.
Meyer, M., Gelman, S. A., & Stilwell, S. M. (2011). Generics Are a Cognitive Default: Evidence From Sentence Processing. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *33*(33), 913–918.

Meyer, M., Leslie, S.-J., Gelman, S. A., & Stilwell, S. M. (2013). Essentialist Beliefs About Bodily Transplants in the United States and India. *Cognitive Science*, *37*(4), 668–710. https://doi.org/10.1111/cogs.12023

Miller, G. A. (1993). Images and Models, Similes and Metaphors. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 357–400). https://doi.org/https://doi.org/10.1017/cbo9781139173865.019

Milsark, G. L. (1974). *Existential Sentences in English*. Massachusetts Institute of Technology.

Murphy, G. (2004). *The Big Book of Concepts*. https://doi.org/10.7551/mitpress/1602.001.0001

Musolff, A. (2014). Metaphorical Parasites and "Parasitic" Metaphors: Semantic Exchanges Between Political and Scientific Vocabularies. *Journal of Language and Politics*, *13*(2), 218–233. https://doi.org/10.1075/jlp.13.2.02mus

Napier, J. L., Mandisodza, A. N., Andersen, S. M., & Jost, J. T. (2006). System Justification in Responding to the Poor and Displaced in the Aftermath of Hurricane Katrina. *Analyses of Social Issues and Public Policy*, *6*(1), 57–74. https://doi.org/https://doi.org/10.1111/j.1530-2415.2006.00102.x

Newman, G. E., Diesendruck, G., & Bloom, P. (2011). Celebrity Contagion and the Value of Objects. *Journal of Consumer Research*, *38*(2), 215–228. https://doi.org/10.1086/658999

Newman, G. E., Herrmann, P., Wynn, K., & Keil, F. C. (2008). Biases Towards Internal Features in Infants' Reasoning About Objects. *Cognition*, *107*(2), 420–432. https://doi.org/10.1016/j.cognition.2007.10.006

Newman, G. E., & Knobe, J. (2019). The Essence of Essentialism? *Mind and Language*, 1–21. https://doi.org/https://doi.org/10.1111/mila.12226

Oderberg, D. S. (2007). *Real Essentialism*. https://doi.org/10.4324/9780203356753

Ortony, A. (1975). Why Metaphors Are Necessary and Not Just Nice. *Educational Theory*, *25*(1), 45–53. https://doi.org/10.1111/j.1741-5446.1975.tb00666.x

Ortony, A. (1979). Beyond Literal Similarity. *Psychological Review*, *86*(3), 161–180. https://doi.org/10.1037/0033-295X.86.3.161

Ortony, A., & Fainsilber, L. (1987). The Role of Metaphors in Descriptions of Emotions. *Proceedings of the 1987 Workshop on Theoretical Issues in Natural Language Processing*, 181–184. https://doi.org/10.3115/980304.980346

Ortony, A., Schallert, D. L., Reynolds, R. E., & Antos, S. J. (1978). Interpreting Metaphors and Idioms: Some Effects of Context on Comprehension. *Journal of Verbal Learning and Verbal Behavior*, *17*(4), 465–477. https://doi.org/10.1016/S0022-5371(78)90283-9

Orvis, B. R., Cunningham, J. D., & Kelley, H. H. (1975). A Closer Examination of Causal Inference: The Roles of Consensus, Distinctiveness, and Consistency Information. *Journal of Personality and Social Psychology*, *32*(4), 605–616. https://doi.org/10.1037/0022-3514.32.4.605

Pappas, A., & Gelman, S. A. (1998). Generic Noun Phrases in Mother-Child Conversations. *Journal of Child Language*, *25*(1), 19–33. https://doi.org/https://doi.org/10.1017/s0305000997003292

Pelletier, F. J. (2010). Are All Generics Created Equal? In F. J. Pelletier (Ed.), *Kinds, Things, and Stuff: Mass Terms and Generics* (pp. 60–79). https://doi.org/10.1093/acprof:oso/9780195382891.003.0004

Pelletier, F. J., & Asher, N. (1997). Generics and Defaults. In J. van Benthem & A. ter Meulen (Eds.), *Handbook of Logic and Language* (pp. 1125–1177). https://doi.org/10.1016/b978-044481714-3/50025-4

Pesowski, M. L., & Friedman, O. (2019). Children Value Objects With Distinctive Histories. *Journal of Experimental Psychology: General*, *148*(12), 2120–2128. https://doi.org/https://doi.org/10.1037/xge0000606

Phelan, J. C. (2005). Geneticization of Deviant Behavior and Consequences for Stigma: The Case of Mental Illness. *Journal of Health and Social Behavior*, *46*(4), 307–322. https://doi.org/10.1177/002214650504600401
Pierce, R. S., & Chiappe, D. L. (2008). The Roles of Aptness, Conventionality, and Working Memory in the Production of Metaphors and Similes. *Metaphor and Symbol*, *24*(1), 1–19. https://doi.org/10.1080/10926480802568422
Prasada, S. (2010). Conceptual Representation and Some Forms of Genericity. In F. J. Pelletier (Ed.), *Kinds, Things, and Stuff: Mass Terms and Generics* (pp. 36–59). https://doi.org/10.1093/acprof:oso/9780195382891.003.0003
Prasada, S. (2012). Mechanisms for Characterizing Kinds and Classes. Retrieved May 23, 2020, from Recherches linguistiques de Vincennes website: https://journals.openedition.org/rlv/2058#text
Prasada, S. (2017). The Scope of Formal Explanation. *Psychonomic Bulletin and Review*, *24*(5), 1478–1487. https://doi.org/10.3758/s13423-017-1276-x
Prasada, S., & Dillingham, E. M. (2006). Principled and Statistical Connections in Common Sense Conception. *Cognition*, *99*(1), 73–112. https://doi.org/10.1016/j.cognition.2005.01.003
Prasada, S., & Dillingham, E. M. (2009). Representation of Principled Connections: A Window Onto the Formal Aspect of Common Sense Conception. *Cognitive Science*, *33*(3), 401–448. https://doi.org/10.1111/j.1551-6709.2009.01018.x
Prasada, S., Khemlani, S., Leslie, S.-J., & Glucksberg, S. (2013). Conceptual Distinctions Amongst Generics. *Cognition*, *126*(3), 405–422. https://doi.org/10.1016/j.cognition.2012.11.010
Prentice, D. A., & Miller, D. T. (2007). Psychological Essentialism of Human Categories. *Current Directions in Psychological Science*, *16*(4), 202–206. https://doi.org/https://doi.org/10.1111/j.1467-8721.2007.00504.x
Putnam, H. (1975). The Meaning of Meaning. In K. Gunderson (Ed.), *Minnesota Studies in the Philosophy of Science* (Vol. 7, pp. 131–139). Minneapolis, MN: University of Minnesota Press.
Rabb, N., Brownell, H., & Winner, E. (2018). Essentialist Beliefs in Aesthetic Judgments of Duplicate Artworks. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(3), 284–293. https://doi.org/10.1037/aca0000174
Rangel, U., & Keller, J. (2011). Essentialism Goes Social: Belief in Social Determinism as a Component of Psychological Essentialism. *Journal of Personality and Social Psychology*, *100*(6), 1056–1078. https://doi.org/10.1037/a0022401
Rehder, B. (2006). When Similarity and Causality Compete in Category-Based Property Generalization. *Memory and Cognition*, *34*(1), 3–16. https://doi.org/10.3758/BF03193382
Rhodes, M., & Gelman, S. A. (2009). Five-Year-Olds' Beliefs About the Discreteness of Category Boundaries for Animals and Artifacts. *Psychonomic Bulletin and Review*, *16*(5), 920–924. https://doi.org/10.3758/PBR.16.5.920
Rhodes, M., Leslie, S.-J., & Tworek, C. M. (2012). Cultural Transmission of Social Essentialism. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(34), 13526–13531. https://doi.org/10.1073/pnas.1208951109
Rhodes, M., & Mandalaywala, T. M. (2017). The Development and Developmental Consequences of Social Essentialism. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*, e1437. https://doi.org/10.1002/wcs.1437
Roeper, T., Strauss, U., & Pearson, B. Z. (2006). The Acquisition Path of the Determiner Quantifier Every: Two Kinds of Spreading. In T. Heizmann (Ed.), *Current Issues in First Language Acquisition* (pp. 97–128). Amherst, MA: Graduate Linguistic Student Association.

Rosch, E., & Mervis, C. B. (1975). Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology*, *7*, 573–605. https://doi.org/https://doi.org/10.1016/0010-0285(75)90024-9
Rosenblum, T., & Pinker, S. A. (1983). Word Magic Revisited: Monolingual and Bilingual Children's Understanding of the Word-Object Relationship. *Child Development*, *54*(3), 773–780. https://doi.org/https://doi.org/10.2307/1130064
Rosengren, K. S., Gelman, S. A., Kalish, C. W., & Mccormick, M. (1991). As Time Goes By: Children's Early Understanding of Growth in Animals. *Child Development*, *62*(6), 1302–1320. https://doi.org/https://doi.org/10.2307/1130808
Ross, L., & Nisbett, R. E. (2011). *The Person and the Situation: Perspectives of Social Psychology*. https://doi.org/DOI:10.5860/choice.29-3604
Roszkowski, M. (2016). *Generizität: Formale und Kognitive Aspekte*. Universitiy of Vienna.
Salomon, E., & Cimpian, A. (2014). The Inherence Heuristic as a Source of Essentialist Thought. *Personality and Social Psychology Bulletin*, *40*(10), 1297–1315. https://doi.org/10.1177/0146167214541659
Searle, J. (1993). Metaphor. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 83–111). https://doi.org/https://doi.org/10.1017/cbo9781139173865.008
Setoh, P., Wu, D., Baillargeon, R., & Gelman, R. (2013). Young Infants Have Biological Expectations About Animals. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(40), 15937–15942. https://doi.org/10.1073/pnas.1314075110
Shapiro, S. (1997). Philosophy of Mathematics: Structure and Ontology. In *The Philosophy of Science: A Companion*. https://doi.org/https://doi.org/10.1093/0195139305.001.0001
Shipley, E. F. (1983). Categories, Hierarchies, and Induction. *Psychology of Learning and Motivation*, *30*, 265–301. https://doi.org/https://doi.org/10.1016/s0079-7421(08)60299-6
Shtulman, A. (2006). Qualitative Differences Between Naïve and Scientific Theories of Evolution. *Cognitive Psychology*, *52*(2), 170–194. https://doi.org/10.1016/j.cogpsych.2005.10.001
Sousa, P., Atran, S., & Medin, D. L. (2002). Essentialism and Folkbiology: Evidence From Brazil. *Journal of Cognition and Culture*, *2*(3), 195–223. https://doi.org/10.1163/15685370260225099
Sperber, D., & Hirschfeld, L. A. (2004). The Cognitive Foundations of Cultural Stability and Diversity. *Trends in Cognitive Sciences*, *8*(1), 40–46. https://doi.org/10.1016/j.tics.2003.11.002
Suhay, E. (2017). Genetic Essentialism. In *The SAGE Encyclopedia of Political Behavior* (pp. 322–325). https://doi.org/doi: 10.4135/9781483391144.n149
Sutherland, S. L., & Cimpian, A. (2015). An Explanatory Heuristic Gives Rise to the Belief That Words Are Well Suited for Their Referents. *Cognition*, *143*, 228–240. https://doi.org/10.1016/j.cognition.2015.07.002
Taylor, M. G., Rhodes, M., & Gelman, S. A. (2009). Boys Will Be Boys, Cows Will Be Cows: Children's Essentialist Reasoning About Gender Categories and Animal Species. *Child Development*, *80*(2), 461–481. https://doi.org/10.1111/j.1467-8624.2009.01272.x
Teichman, M. (2016). The Sophisticated Kind Theory. *Inquiry*, 1–47. https://doi.org/10.1080/0020174X.2016.1267407
Thibodeau, P. H., Matlock, T., & Flusberg, S. J. (2019). The Role of Metaphor in Communication and Thought. *Language and Linguistics Compass*, *13*(5), 1–18. https://doi.org/10.1111/lnc3.12327
Toosi, N. R., & Ambady, N. (2011). Ratings of Essentialism for Eight Religious Identities. *International Journal for the Psychology of Religion*, *21*(1), 17–29. https://doi.org/10.1080/10508619.2011.532441

Tversky, A. (1977). Features of Similarity. *Psychological Review*, *84*(4), 327–352.
Tversky, A., & Kahneman, D. (1974). Judgment Under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/DOI: 10.1126/science.185.4157.1124
Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The Influences of Emotion on Learning and Memory. *Frontiers in Psychology*, *8*, 1–22. https://doi.org/10.3389/fpsyg.2017.01454
Vasilyeva, N., Gopnik, A., & Lombrozo, T. (2018). The Development of Structural Thinking About Social Categories. *Developmental Psychology*, *54*(9), 1735–1744. https://doi.org/10.1037/dev0000555
Waxman, S., Medin, D. L., & Ross, N. (2007). Folkbiological Reasoning From a Cross-Cultural Developmental Perspective: Early Essentialist Notions Are Shaped by Cultural Beliefs. *Developmental Psychology*, *43*(2), 294–308. https://doi.org/10.1037/0012-1649.43.2.294
Wilson, D., & Sperber, D. (2012). A Deflationary Account of Metaphors. In D. Wilson & D. Sperber (Eds.), *Meaning and Relevance* (pp. 97–122). https://doi.org/https://doi.org/10.1017/cbo9781139028370.007
Wisniewski, E. J. (1996). Construal and Similarity in Conceptual Combination. *Journal of Memory and Language*, *35*(3), 434–453. https://doi.org/10.1006/jmla.1996.0024
Wisniewski, E. J. (1997). When Concepts Combine. *Psychonomic Bulletin and Review*, *4*(2), 167–183. https://doi.org/10.3758/BF03209392
Wodak, D., Leslie, S.-J., & Rhodes, M. (2015). What a Loaded Generalization: Generics and Social Cognition. *Philosophy Compass*, *10*(9), 625–635. https://doi.org/10.1111/phc3.12250
Wolff, P., & Gentner, D. (2011). Structure-Mapping in Metaphor Comprehension. *Cognitive Science*, *35*(8), 1456–1488. https://doi.org/10.1111/j.1551-6709.2011.01194.x
Yamauchi, T., & Markman, A. B. (2000). Inference Using Categories. *Journal of Experimental Psychology: Learning Memory and Cognition*, *26*(3), 776–795. https://doi.org/10.1037/0278-7393.26.3.776

Abstract

Antisemitische Metaphern haben eine lange Tradition in öffentlich-diffamierenden Diskursen und treten in den unterschiedlichsten Variationen auf, jedoch allesamt mit dem Ziel, Juden und Jüdinnen als nutznießerisch, übermächtig oder gefährlich zu stigmatisieren, abzuwerten und zu entmenschlichen. Abseits solch inhaltlicher Assoziationen üben antisemitische Metaphern Einfluss auf die Struktur unserer mentalen Konzepte aus: So werden Juden und Jüdinnen als eine einheitliche soziale Kategorie wahrgenommen, die sich scharf von anderen Mitgliedern der Gesellschaft abgrenzen lässt und Rückschlüsse auf die Identität eines jeden Einzelnen ihrer Vertreter zulässt. In der vorliegenden Masterarbeit wird dieser Einfluss auf die konzeptuelle Repräsentation einer jüdisch-sozialen Kategorie aus einer kognitionswissenschaftlichen Perspektive beleuchtet und auf zwei unterschiedliche linguistische Elemente, die in antisemitischen Metaphern zum Tragen kommen, zurückgeführt: Generizität und Metaphorik. Basierend auf den Ergebnissen dreier Forschungsgebiete bezüglich sozialer Kategorisierung, generischer Sprache und Metaphern wird beleuchtet, wie sich die beiden Elemente auf unterschiedliche Aspekte sozialer Kognition auswirken und nachherige Denkprozesse begünstigen. Vor diesem Hintergrund wird argumentiert, dass antisemitische Metaphern nicht nur auf einer inhaltlichen Ebene operieren, sondern weitreichende Folgen für den Aufbau unserer mentalen Repräsentationen haben. Als solche schaffen sie den Nährboden für antisemitische Stereotypisierung, Vorurteilsbildung und anderweitige negative Konsequenzen sozialer Kategorisierung.