LEAPS AND BOUNDS:
Towards an Integrative Theory of Cultural Evolution

Sami Rissanen
Master's Thesis in Philosophy
Department of Social Sciences and Philosophy
University of Jyväskylä
Spring 2014

University of Jyväskylä
Faculty of Social Sciences

Abstract of
Master's Thesis

Author: Sami Olavi Rissanen
Title of Thesis: Leaps and Bounds: Towards an Integrative Theory of Cultural Evolution
Date: May 19th 2014
Pages: 135
Department: Department of Social Sciences and Philosophy
Chair: Philosophy
Supervisor: Professor Jussi Kotkavirta

Abstract:

This study attempts to combine the idea that culture gives rise to a form of natural election with Dan Sperber's positions concerning (1) the underlying dynamics of communication as inferring intended meanings according to their contextual relevance, and (2) the spread of distinct conceptions regarding different objects. Such a combinatory framework is intended as a viable answer to Sperber's criticism of applying Darwinism to cultural phenomana. Sperber's critiques include the claims that (1) the high mutation rate of culture prevents its cumulative evolution, (2) generalizations concerning representations are untenable because they oversimplify matters, and (3) instead of being replicated, transmitted cultural information is reconstrued. If successful, this project should benefit anyone who is inclined to utilize an evolutionary approach to research cultural phenomena.

The meme theory, the idea that there exist replicating cultural entities, is the primary representative of the idea that Darwinism applies to culture in this study. While alternative selectionist approaches such as the dual-inheritance theory, which argues that people inherit both genetic and cultural information, are also mentioned occasionally, they play a merely supporting role when detailing the integrative theory of cultural evolution is attempted. Instead, this study utilizes Dan Sperber's epidemiology of representations, according to which the transformations always occur when cultural information is handled, and these biases are caused by actors' dispositions. In particular, the relevance theory of communication constitutes an indispensable part of the newly introduced framework. Other resources include Pierre Bourdieu's theory of fields and Sir Karl Popper's theories about the evolution of science and the three worlds of (1) matter, (2) mind, and (3) culture.

The integrative theory this study introduces accepts the claim that representations often transform when they are inferred from the demonstrations which convey them. However, a framework which explains how their contents may be retained in practice despite this dynamic is provided to salvage the idea that culture contains a form of natural selection. According to this theory, people learn to associate certain traits with specific categories and minding another's position enables tailoring expressions to convey the intended meanings.

Keywords: association, communication, cultural evolution, epidemiology, meme theory, meme, metarepresentation, pragmatics, relevance theory, representations, selectionism

Jyväskylän yliopisto
Yhteiskuntatieteellinen tiedekunta

Pro gradu -tutkielman tiivistelmä

Tekijä: Sami Olavi Rissanen
Tutkielman nimi: Leaps and Bounds: Towards an Integrative Theory of Cultural Evolution
Päivämäärä: Toukokuun 19. 2014
Sivumäärä: 135
Laitos: Yhteiskuntatieteiden ja filosofian laitos
Oppiaine: Filosofia
Tutkielman ohjaaja: Professori Jussi Kotkavirta

Tiivistelmä:

Tutkimuksessa pyritään yhdistämään ajatus kulttuurista eräänlaisen luonnonvalinnan alaisena Dan Sperberin näkemyksien (1) kommunikaation taustadynamiikasta tarkoitettujen merkityksien päättelynä niiden kontekstisidonnaisen merkittävyyden perusteella ja (2) erilaisten käsityksien eri kohteista leviämisen suhteen. Tällaisen yhdentävän viitekehyksen turvin pyritään vastaamaan pätevästi Sperberin kritiikkiin koskien darwinilaisen luonnonvalinnan soveltamista kulttuurin ilmiöihin. Näihin kritiikkeihin lukeutuvat väitteet, että (1) kulttuurin muutostiheys on liian korkea evoluution tuloksien kasautumisen mahdollistamiseksi, (2) yleistykset koskien sisältöjä ovat kestämättömiä tehtyjen yliyksinkertaistuksien vuoksi ja (3) toisintumisen sijaan välittyvät kulttuurisisällöt ovat tulkinnanalaisia. Mikäli tämä projekti onnistuu, sen tulisi hyödyttää ketä vain, joka on taipuvainen hyödyntämään evoluutiolähtöistä lähestymistapaa kulttuurin tutkimiseksi.

Meemiteoria, eli ajatus toisintuvista kulttuuriolioista, on ensisijainen tutkimuksessa hyödynnetty lähde ajatukselle, että luonnonvalinta pätee kulttuuriin. Vaikka muihin vastaaviin lähestymistapoihin kuten rinnakkaisevoluutioteoriaan, joka väittää ihmisten perivän sekä geneettistä että kulturaalista informaatiota, viitataan ajoittain, ne lähinnä tukevat yritystä tarkentaa esitettyä yhtenäistävää teoriaa kulttuurin evoluutiosta. Sen sijaan tämä tutkielma hyödyntää Sperberin edustuksien epidemiologiaa, jonka mukaan kulttuuri-informaation käsittely muokkaa aina sen sisältöjä toimijoiden taipumuksien mukaisesti. Erityisesti kommunikaation relevanssiteoria on korvaamaton osa tätä esiteltävää kehikkoa. Muihin lähteisiin lukeutuvat Pierre Bourdieun kenttäteoria sekä Sir Karl Popperin teoriat tieteen kehittymisestä ja (1) aineen, (2) mielen ja (3) kulttuurin kolmesta maailmasta.

Tämän tutkielman esittelemä yhtenäistävä teoria hyväksyy väitteen edustuksien muutoksien yleisyydestä, kun ne päätellään välittävistä eleistä. Ajatuksen luonnonvalinnasta kulttuurin sisällä pelastamiseksi tarjotaan kuitenkin tarjotaan teoriakehystä, joka kykenee selittämään edustuksien sisältöjen säilymisen käytännössä näistä muutoksista huolimatta. Kyseisen teorian mukaan ihmiset oppivat yhdistämään tiettyjä piirteitä eri kategorioihin ja huomioimalla toistensa näkökulmat räätälöimään aiotut merkitykset välittäviä ilmaisuja.

Asiasanat: assosiointi, edustukset, epidemiologia, kommunikaatio, kulttuurievoluutio, meemi, meemiteoria, metarepresentaatiot, pragmatiikka, relevanssiteoria, selektionismi

FOREWORD

One of the most interesting questions any social scientist faces is the relationship between people and culture. Human subjects tend to produce relatively objective structures which in turn affect their worldviews and seemingly even the subconscious fashion in which they structure the reality they encounter. This reciprocal relationship complicates research which pertains to social and psychological phenomena but it is also endlessly fascinating and provides a neverending supply of meaningful problems.

This study contains one aspiring researcher's contribution, and the provided answer has undergone multiple revisions. Even by now, some new ideas which would clarify the account have arisen but they will be reserved for the future. Originally, even the explicitly stated research question was very different although all the stages during which the process of formulating this thesis has transpired ultimately pertain to the same phenomenon. What all these stages always addressed were the differences and the similarities between people's worldviews and the effects these often socially transmitted conceptions have on inter-personal affairs.

Finalizing this study would never have been possible without the gratuitous help of my family, my peers and the members of the faculty of philosophy. Thus, their contributions must also be acknowledged. Professor Kotkavirta is to be thanked for accepting the duty to supervise a project which hardly matches his areas of specialty, while Jarno Hietalahti deserves praise for supervising the initial phases and for providing helpful resources. Likewise, the occasional assistance provided by Professors Yrjönsuuri and Moisio has been most helpful. John Pajunen has bestowed particularly helpful insights regarding a plenitude of matters but the efforts of Tomi Kalimäki who provided feedback on an earlier versions of this study deserve recognition as well. Lassi Haapanen alongside Antti Hallamäki has also contributed to the mathematics this study occasionally utilizes.

Finally, the constant support of my parents, Päivi Rissanen and Arto Rissanen, has enabled finishing this project on numerous levels. The least of their contributions are hardly how they have endowed me with an interest in academic enterprises and supported the choices this preoccupation has led to despite having lacked such an ambition themselves.

TABLE OF CONTENTS

LEAPS AND BOUNDS:
Towards an Integrative Theory of Cultural Evolution

1. INTRODUCTION

"*Philosophers are never quite sure what they are talking about – about what the issues* really *are – and so often it takes them rather a long time to recognize that someone with a* somewhat *different approach (or destination, or starting point) is making a contribution.*"

– Daniel Dennett, *The Intentional Stance* (Dennett, 1989, p.350)

"*And always remember that it is* impossible to speak in such a way that you cannot be misunderstood*: there will always be someone who misunderstands you.*"

– Sir Karl Popper, *Unended Quest* (Popper, 1974/1992/2002c, p.29)

The phrase 'leaps and bounds' denotes dramatic improvements, and it indicates how progress may sometimes occur in leaps which skip expected intermediate phases of development. However, 'bounds' means both 'jumps' and 'boundaries', and the title toys with this homonymy. The leaps this study addresses pertain to the leaps of logic people execute when they jump to conclusions because they happen to associate certain traits with the classifications they use to structure experienced reality. On the other hand, bounds include (1) the skipping caused by the further associations which those traits engender and (2) the boundaries which delimit which associations are actually available.

The dramatic improvements the standard meaning of that phrase would imply should hopefully apply to the philosophical and scientific discussion on the evolution of culture. In the above quote, Dennett submits that philosophers oftentimes fail to comprehend the significance of each other's approaches (Dennett, 1989, p.350). This claim should ring particularly true for deviant perspectives, and while conceptualizing culture as an evolving entity is hardly a new idea among philosophers, the position that its evolution is parallel to natural selection remains quite peripheral. Thus, proving the fertility of researching the Darwinian evolution of culture in earnest would truly be a contribution worth attempting if this position's minor representation is not due to any true lack of merit.

The stance assumed towards the contributions of other theorists is quite ecletic in order to acknowledge the significance of their suggestions. This study attempts to integrate, not only two specific theories, but rather, the roles which individuals and their wider cultural environments play in the development of one another, into a relatively concise framework. Since that topic is unimaginably broad and pervasive, merging a multitude of theories to describe and explain the dynamics of either pole and their relationship expands the inclusiveness any theory about such a wide topic requires. Thus, the treatment of selec-

tionism focuses on the meme theory (e.g. Dawkins, 1976/1989/2006; Dennett, 1991; 1995; Blackmore, 1999; Distin, 2005) but is supported by some concepts extracted from the dual-inheritance theory (e.g. Boyd and Richerson, 1985; Henrich and Boyd, 1998). Meanwhile, Sperber's framework (e.g. Sperber and Wilson, 1986/1995; Sperber, 1996) is supported by some related considerations (e.g. Dennett, 1989; Baron-Cohen, 1995). Pierre Bourdieu's theory of fields (e.g. Bourdieu, 1972/1977; 1979/1984/2010; Grenfell, 2012) as well as Sir Karl Popper's evolutionary epistemology (e.g. Popper, 1959/2002b; 1972/1979) and three world ontology (e.g. Popper, 1972/1979; Popper and Eccles, 1977/1983) offer further assistance. Because the introduced framework is intended as a realistic and materialistic theory of communication, some attention is also paid to the neurophysiological foundation of representations (e.g. Edelman, 1987b; Cosmides and Tooby, 2000; Eichenbaum, 2004).

The integrative theory being advanced may be condended using Popper's words: '– – it is *impossible to speak in such a way that you cannot be misunderstood*: there will always be someone who misunderstands you.' (Popper, 1974/1992/2002c, p.29). The fundamental proposition this quote contains is the claim that any expression may be misunderstood by some individual. Sperber affirms the same thesis (Sperber and Wilson, 1986/1995; Sperber, 1996; Wilson and Sperber, 2004; 2012), and its admission here is due to his influence. This study diverges from the Sperberian framework by claiming that despite this insurmountable fallibility, at least associations may be propagated reliably. Because such conceptions proliferate differentially, a form of natural selection is claimed to emerge. Precursory theories which describe the dynamics of such cultural selection already exist (e.g. Boyd and Richerson, 1985; Heylighen, 1999; Distin, 2005; Heylighen and Chielens, 2009) but they must still be supplemented so that testable hypotheses about the evolution of culture may be formulated. Some of the chosen additions (e.g. Bourdieu, 1972/1977; Dennett, 1989; Baron-Cohen, 1995) should complement the picture thus.

Thus, this study inherits Sperber's framework for communication, according to which (1) meanings are never strictly encoded (Sperber and Wilson, 1986/1995, pp.24–28), (2) a communicator may only provide ambiguous ostensions to communicate some intended meaning (Ibid. pp.49–54), (3) these ostensions are attributed a meaning by the receivers who reconstrue a viable intention (Ibid. pp.50–64), and (4) this process of inferring a meaning is regulated by a modular system which assesses the relevance of available interpretations in light of available contextual evidence (Ibid. Chapters 2 and 3). Additio-

nally, Sperber provides important tools with which changes in people's positions may be dissected. These include (1) attractors (Sperber, 1996; Claidière and Sperber, 2007), (2) a typology of representations (Sperber, 1996; 2000c) and (3) cognitive causal chains (Sperber, 2001; 2006). Sperber's perspective emphasizes the significance of rational inferences and individual representations since he underscores that all communication consists of interpreting and reconstruing particular representations (Sperber, 1996, Chapter 2).

In contrast, selectionism introduces the radical claim that parallels may be drawn between natural selection and the evolution of culture (e.g. Dawkins, 1976/1989/2006b, Chapter 11; Dennett, 1995, Chapter 12). Behind this claim is the idea that natural selection is not determined by its objects but by the presence of certain substrate-neutral criteria such as (1) variation, (2) replication, and (3) differential fitness (Dennett, 1991, p.200; 1995, p.343). If culture indeed fulfills these requisite criteria, the fitness of different cultural objects may be evaluated in order to predict the rate of their propagation under set circumstances, or to explain the allure and widespreadness of certain conceptions and artifacts.

Chapter 2 elaborates this position by clarifying this idea of Universal Darwinism (Dawkins, 1983; Dennett, 1995; Blackmore, 1999, Chapter 2) and by providing a more detailed account of the meme theory. The latter relies heavily on the views of Francis Heylighen (e.g. Heylighen, 1999; Heylighen and Chielens, 2009) and Kate Distin (Distin, 2005) because Heylighen attempts a rigorous study of the criteria which affect the selection of culture (e.g. Heylighen, 1993; 1997; 1999) while Distin unravels themes older accounts ignore. She also attempts to answer Dan Sperber's criticism (Distin, 2005, pp.102–108), which is reviewed and rebutted tentatively at the end of chapter 2.

In turn, chapter 3 explores Sperber's position in more detail. It begins by detailing the relevance theory of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) and also expands on the function metarepresentations have in Sperber's framework (e.g. Sperber, 2000c). This section on metarepresentations also includes contributions from other theorists, particularly Dennett whose intentional systems theory (e.g. Dennett, 1989; 2009a) is present in many contemporary conceptions about social cognition (e.g. Baron-Cohen, 1995). Finally, this chapter introduces Sperber's alternative approach to the evolution of culture in detail. The attractors which define his epidemiology of culture (Sperber, 1996) are introduced but that section also criticizes Sperber's claim that selection and attraction are fundamentally different. In particular, Catherine

Driscoll's arguments which support the inclusion of selectionism despite the effects of attractors (Driscoll, 2011) are utilized to support this integration. Additionally, the possibility of actual leaps in the design of conceptions happening and the dynamics of such revolutions are contemplated because most theorists hardly address them.

Construing the actual outline of an integrative theory is finally attempted in chapter 4. That chapter begins by introducing some key conceptions present in Pierre Bourdieu's theory of fields: (1) doxa, (2) habitus, (3) (social) capitals, and (4) fields. These concepts are deemed necessary because Bourdieu utilizes them to link together social actors and the structures they produce – and by which they are molded themselves. The urge to describe this reciprocity motivates both his reflexive sociology (Bourdieu and Wacquant, 1992) and this study. Bourdieu also concentrates on the *differences* between distinct, inherited positions (e.g. Bourdieu, 1979/1984/2010). These distinctions must be addressed to answer Sperber's criticism because the claimed impossibility of cumulative selection relies partially on the effects of differences between individuals (Sperber, 2000a). Finally, the exposition of the doxa (Bourdieu, 1972/1977, pp.164–171) helps analyze the role conventions and unconscious biases may have.

Besides this introduction to the theory of fields, Chapter 4 contains clarifications regarding the particulars of different representations. Cultural representations – the generalizations Sperber eschews – are elucidated by structuring the relationship of the individual mental and public representations which comprise these species using Popper's three worlds (Popper, 1972/1979; Popper and Eccles, 1977/1983). That section also outlines the formal descriptions used to classify the contents of particular representations.

The section about mental representations exploits those descriptions and provides some neurophysiological arguments to support their viability in describing the contents of actual mental conceptions. This section also details how primary mental representations and mental metarepresentations relate to one another and how they might be formed in the first place. Both are claimed to be affected by the contextual relevance of certain traits as well as the stance being applied towards their functioning (Dennett, 1989; 2009a). The roles of propositional attitudes (Hintikka, 1962/2005; Distin, 2005, pp.169–170) and personally experienced certainty (Peirce, 1955, Chapters 1–4; Wittgenstein, 1969) are explored as well because they affect the probabilities of related conceptions becoming assimilated. Likewise, propositional attitudes are a part of the content metarepresentations involve.

Finally, public representations are tackled. The corresponding section ponders the relationship between mental and public representations and the concludes that Sperber is approximately correct: they do form causal chains, and each link of such a chain is an interpretation of its predecessor (Sperber, 1996, Chapter 2). However, after the elaboration of these processes, room for a form of selection is still left. Public expressions are claimed to function as the *interactors* (Hull, 1980; 2001) which transmit the contents of mental representations in order to facilitate their replication. Since this replication is claimed to emerge from how people infer and reconstruct the meanings of the expressions they encounter, the factors which enable it are specified. The section also provides a few conjectures about the forms in which different public products might become assimilated.

While many of the ideas this study utilizes have originated outside the institution of philosophy, hopefully their relevance will be demonstrated successfully on these pages. Both the theories this study involves and the original theory it promotes should prove serviceable to the philosophy of mind as well as the philosophy of language. Additionally, they help explicate how a materialist ontology may integrate different levels of explanation without deferring to familiar intuitions concerning human existence. Since most readers are likely inclined towards philosophy, both the inclusion of such a variety of non-philosophical theories and the made stylistic choices might feel alienating. Nevertheless, it is hoped that the contributions this study attempts are acknowledged according to their actual merits despite this eccentricity in relation to established philosophical practice.

2. SELECTIONISM

”*Most of what is unusual about man can be summed up in one word: 'culture'. I use the word not in its snobbish sense, but as a scientist uses it. Cultural transmission is analogous to genetic transmission in that, although basically conservative, it can give rise to a form of evolution.*”

– Richard Dawkins*, The Selfish Gene* (Dawkins 1976/1989/2006b, p.189)

”*Evolution is to analogy as statues are to birdshit.*”

– Steve Jones, cited in *The Darwin Wars* (Brown, 1999, p.v)

Dawkins's above claim summarizes the founding ideas of this chapter: (1) the existence of culture defines being human, (2) the propagation of culture approximately resembles natural selection, and (3) this enables a non-genetic, new form of evolution (Dawkins, 1976/1898/2006b, p.189). The phenomenon of culture is also claimed to be observable objectively. Such a scientific attitude is applied towards the phenomenon of culture itself, while (customary) evaluations apply to its manifestations. Mapping the latter may be an indispensable part of explaining cultural selection but this chapter instead attempts to demonstrate the laws under which the influence of such assessments operates.

The quote attributed to Steve Jones is a healthy reminder of the danger of soiling the theory of evolution by applying it carelessly. Despite its obvious dismissive connotations, the metaphor is actually quite befitting. The theory of evolution is indeed a sturdy and beautiful whole which careless analogs would merely tarnish. Avians, however, disseminate indigested seeds by defecating, and their faeces provides those seeds with a preheated parcel of nutrients. In other words, drawing parallels does not suffice. Despite the benefits existing theories confer, the conjectures those seeds contain must rely on their own assets to sprout and bloom into new theories able to stand beside such venerable statues.

Here, culture is understood approximately in the sense conferred to it by Dennett, according to whom culture is 'an *extra* medium of design preservation and design communication' (Dennett, 1995, p.338). Thus, the contents of culture are proposed to be subject to the same mechanisms of selection as the contents of genes, the standard medium of design preservation. Culture is thus posited to be reformed by the algorithm of natural selection (Dawkins, 1983; 1976/1989/2006b; Dennett, 1995). While other accounts of cultural selection exist (e.g. Boyd and Richerson 1985), this chapter focuses on the meme theory which grants that cultural entities are not completely subordinate to human interests.

The chapter will introduce (1) the principles of universal Darwinism, (2) the meme theory which applies them to explain the transmission of culture (e.g. Dawkins, 1976/1989/2006b; Dennett, 1991; 1995; Blackmore, 1999; Distin, 2005), and (3) relevant critivism of the meme theory (e.g. Sperber, 1996; 2000a). The reader is aspired to become acquainted with selectionism as well as the concept of the *meme* and its problems.

2.1 Universal Darwinism

This section explores natural selection's capacity to increase complexity of structures – their design – utterly mindlessly, like a sieve. Natural selection in this sense can be understood as a general process which applies in all situations in which a number of criteria, none of which is limited to organisms, is fulfilled. The claimed universality results from this substrate-neutrality of these requirements. Although not *everything* is posited to actually be subject to selection, *anything* at all may become subjected to it if the minimal requirements are fulfilled.

Richard Dawkins has argued that Darwinian natural selection is, even in principle, the only presented theory capable of explaining the appearance of complex structures (Dawkins, 1983; 2003, Chapter 2.2; 1986/2006a, pp.287–318). The reason is the algorithmic character of natural selection: it produces specified results according to the baseline circumstances (Dennett, 1995, pp.48–60). More precisely, natural selection is a *crane* (Ibid. pp.73–76; 2013, pp.227–231). Its origins are explainable without postulating pre-existing complex design, and the dynamics which erect new design are analyzable, unlike in proposed alternatives which are *skyhooks* (Dennett, 1995, p.74; 2013, pp.226–227). Unlike cranes, skyhooks postulate inexplicable, pre-existing designs, which often begs the question (Ibid.). Triggering the Darwinian algorithm only entails the fulfillment of three criteria, none of which is dependent on the complexity of existing structures.

These three criteria, which Dennett claims to necessarily produce a form of natural selection, are (1) variation, (2) replication and (3) differential fitness (Dennett, 1991, p.200; 1995, p.343). In defining natural selection as an algorithmic process, Dennett emphasizes how it is substrate neutral: any and all things which fulfill the above criteria are subjugated to selective processes (Dennett, 1995, pp.50–51). Similarly, Dawkins argues that all possible forms of life anywhere in the universe must have developed under

Darwinian selection regardless of whether structure-containing information is transmitted from one generation to another genetically or not (Dawkins, 1983). This process does not necessitate the emergence of (conventional) lifeforms. The original purpose of memes was illustrating how a familiar phenomenon could be perceived as a form of non-genetic natural selection (Dawkins, 1976/1989/2006b, Chapter 11).[1] David Deutsch also speculates that only the stability engendered by selection could in principle cause (1) the information which enables the formation of organisms and (2) the information they in turn produce to remain relatively constant across the multiverse (Deutsch, 1997, pp.170–193).[2]

This same basic dynamic of selection is also described by Dawkins under the name c*ore Darwinism*: '*Core Darwinism*, I shall suggest, is the minimal theory that evolution is guided in adaptively nonrandom directions by the nonrandom survival of small random hereditary changes.' (Dawkins, 2003, p.95) Dawkins emphasizes how this process is gradual and how selection only concerns adaptively significant changes (Ibid.). The views opposed to these premises, neutralism and saltationism, are inherently problematic as demonstrated by Dawkins (1986/2006a, pp.231–252, pp.303–312) and Dennett (1995, pp.287–288). While this definition actually describes biological evolution, the notion of core Darwinism is compatible with the replication dynamics Dawkins endorses elsewhere (e.g. Dawkins 1976/1989/2006b; 1982/1999). Hence, core Darwinism may be generalized because inheritance can be identified with the replication of any surviving information which contributes to its own preservation (Dawkins, 1976/1989/2006b, Chapter 2).

This definition of core Darwinism implicitly contains Dennett's three criteria. Variation signifies the existence of differences among replicators, and its continuation is dependent on the 'small random hereditary changes' mentioned by Dawkins (2003, p.95). Dennett defines variation as 'a continuing abundance of different elements' (Dennett, 1991,

1 Exactly this detaching of meme evolution from genetical advantage is the strength of Dawkins's original sketch according to, for example, Dennett (1991, pp.202–203; 1995, pp.361–365; 2009, pp.128–132) and Blackmore (1999, Chapters 3 and 9; 2009, pp.299–302).

2 Other examples of evolutionary speculations and actual theories are legion. Gerald Edelman has successfully modeled the immune system's production of antibodies as a form of selection (e.g. Edelman and Gally, 1964; Edelman, 1987a). Furthermore, he has proposed a form of neural Darwinism which allegedly explains the formation and strengthening of neural connections (e.g. Edelman, 1987b; 2006; Edelman and Tononi, 2000, Part III). Likewise, Karl Popper compared the logic of epistemic discovery to a natural selection of theories (Popper, 1959/2002b, pp.276–282; 1972/1979), and David Hull discusses the evolution of science (e.g. Hull, 2001, Parts II and III). While these are all respectable scientific conjectures, exceptions also exist. For example, Charles Peirce has claimed that laws of nature themselves may be subject to selection (Peirce, 1955, Chapter 26) and that an evolution of love will uncover the existence of the divine (Ibid. Chapter 27). Since processes of elimination and inheritance of traits among natural laws are nigh unimaginable, Peirce's hollow-seeming attempts echo the importance of Steve Jones's warning.

p.200; 1995, p.343). This continuation of variance is necessary because of the eliminative nature of natural selection which has even prompted Blackmore to dub this process *design by death* (Blackmore, 2009, p.297). The persistence of variety within an evolving set requires postulating the existence of mutations, or 'small random hereditary changes' (Dawkins, 2003, p.95). Without them, evolution would simply accomplish a local optimum in which all adaptively significant traits would be shared and the only variety would be among traits which are adaptively neutral in that environment.

However, the necessity of random changes implied above is misleading outside the biosphere. While mutations in the genome are indeed random and this randomness is often considered definitive of mutations, this feature should not be assumed to necessarily transfer to other substrates. The randomness of hereditary changes is not required for the algorithm of selection to function because all occasional divergences caused by outside influences suffice. A simple proof of this is the thought experiment that determinism is true. If everything was indeed determined, all mutations in the genome would also be nonrandom. Because the results these processes obtain in this deterministic example reality are *observationally equivalent* with random variation (Werndl, 2009, Chapter 5), selection can operate with nonrandom variation.

In turn, the survival of hereditary qualities in Dawkin's definition means replication because he stresses its importance for evolution (e.g. Dawkins, 1976/1989/2006b; 1982/1999).[3] Replication consists of objects or their elements having the capacity to multiply by producing replicas of themselves (Dennett 1991, p.200; 1995, p.343). While variation requires the existence of mutations, replication demands the relative scarcity of such changes. Otherwise, the qualities having already been selected would not reliably transfer to the next generation (Dawkins, 1982/1999, p.112; Sperber 1996, pp.102–103). In effect, their *survival* would become random, violating the conditions of core Darwinism. Every mutation is a case of the failed replication of existing design (Sperber 1996, p.102).

Finally, differential fitness denotes the relationship of differences between replicating units and selection pressures produced by their environment. Dennett formulates this idea as follows: 'the number of copies of an element that are created in a given time varies,

3 In this study, the term *replication* is preferred over *heredity* because that term reminds of the fact that the primary targets of selection are not the structures produced by replicators but the replicators themselves (Dawkins 1976/1989/2006b: 33–45; Dawkins 1982/1999: Chapter 6). Of course, the structures produced are those actually interacting with the environment to determine the replicators' fitness (Hull, 1980; 2001, Part I). The exclusion of the term *heredity* is merely an attempt to avoid the misunderstanding that conventional biological entities like organisms should be treated as replicators.

depending on the interactions between the features of that element (whatever it is that makes it different from other elements) and features of the environment in which it persists' (Dennett, 1991, p.200; 1995, p.343). In core Darwinism's case, nonrandomness is essential: when the survival of changes is nonrandom, evolution obtains a direction (Dawkins, 2003, p.95). The differential adaptive significance of these inherited changes ensures their non-random survival. Selection steers forms towards a local optimum under any prevailing circumstances this way (Dawkins, 1982/1999, pp.32–54).

In addition Dennett and Dawkins's criteria provided above, alternative requirements for selection have been suggested. Gerald Edelman lists the required existence of (1) a generator of diversity, (2) an environmental challenge which produces competition, and (3) differential reproduction of fitter variants (Edelman, 2006, p.27). A fourth, separate criterion is (4) *degeneracy*, or alternative inherited structures providing functionally similar outcomes (Ibid. p. 33–34; McAdams and Arkin, 1999; Edelman and Gally, 2001). While the latter is discussed primarily within the context of biological evolution, Edelman does admit its occurence at the level of culture (Edelman and Gally, 2001, Table 1).

Edelman's first criterion emphasizes the required persistence of variety ensured by a flow of new alternatives. However, localized selection processes could determine the fittest alternative from a set variety without further generation of diversity. Thus, this first criterion only applies to continuous selection. His second criterion is least clearly addressed by the previously provided accounts. However, Dennett does reference the relationship of features of elements and *features of the environment* (Dennett, 1991, p.200; 1995, p.343). While Edelman's use of the word *challenge* seemingly distinguishes a single instance, the features of an environment constantly provide a multitude of challenges. The third criterion is an amalgam of the earlier criteria of replication and differential fitness.

Yet another alternative set of criteria for selection is provided by John von Neumann. His criteria are (1) the sequestration of inherited information, (2) a clear-cut distinction between genotype and phenotype, and (3) no transmission of acquired traits (von Neumann and Burk, 1966 cited in Gabora, 2011, p.6). To a degree, all these criteria are variations of a common theme which concerns the separateness of the germ and somatic (or dead-end) lines (Dawkins, 1982/1999, pp. 83–85). Since cultural evolution is often posited to include the transmission of acquired traits (Sperber, 1996; 2000a; Gabora, 2004; 2011; Heylighen and Chielens, 2009), von Neumann's criteria could indeed prove

problematic to such accounts. However, this study conforms to Kate Distin's separation of the genotypic and phenotypic aspects of memes (Distin, 2005, pp.92–95).

The fitness of any replicator is constituted by its (1) longetivity, (2) fecundity and (3) copying-fidelity (Dawkins, 1976/1989/2006b, pp.17–18; Heylighen and Chielens, 2009, p.5). Longetivity signifies the average lifespan of individual tokens of the replicator. Longer lasting replicators have on average more opportunities to replicate themselves which increases the mean amount of replicas produced (Dawkins, 1976/1989/2006b, p.17). Fecundity, on the other hand, denotes the mean number of replicas produced during a designated timespan (Ibid.). Lastly, copying-fidelity means the ratio of successful replication to all instances of attempted replication (Ibid.). Since imperfect replication produces new types of replicators, its contribution to the propagation of the original is null (Ibid.). Once again, the alleged failure of cultural entities to propagate fidelitously motivates the criticism towards theories of cultural evolution (Sperber, 1996, pp.100–106; 2000a; Gabora, 2011, pp. 6–9;). Even Dawkins admits this as a potentially fatal flaw of the meme account (Dawkins, 1982/1999, pp. 112).

The general view provided above attempts to show how exactly the relatively simple process of selection has the capacity to supplant intentional creators as the producer of perceived diversity and complexity. Were this process to be personified, it would be 'heartless – even vicious – but boundlessly stupid'. (Dennett, 1995, pp.478–479) In other words, the blind watchmaker known as selection is but an algorithm with absolutely no foresight (Dawkins, 1986/2006a, p.5). In the case of nature, this is particularly true. However, it is claimed that intentionality is unnecessary elsewhere as well. Even when it is applied, it supplements rather than replaces selection.

The perspective being advocated here relates to Dawkins's description of the theory of the selfish gene as comparable to a Necker Cube. According to this comparison, both the replicator-centered and the organism-centered perspective are equally valid but depict different aspects of the same phenomenon (Dawkins, 1982/1999, pp. 1, 4–5). What is being asserted is that the focus on replication dynamics is not limited to biological evolution but is universal. For example, Blackmore utilizes the *meme's eye view* when speculating about the reasons for evolutionary developments such as the sudden growth of the human brain (Blackmore, 1999, Chapter 6) and the birth of altruistic tendencies (Ibid. Chapter 12). All that is required is the fulfillment of certain criteria.

Simply put, universal Darwinism describes how core Darwinism, the directedness of (random) changes in nonrandom ways due to selection pressures (Dawkins, 2003, p.95), manifests itself wherever there exist replicators which differ in their probabilities of being replicated. The driving idea of the meme theory is that this dynamic is applicable to explain cultural phenomena if cultural contents fit the three criteria of universal Darwinism. In other words, there ought to exist (1) variety, (2) replication, and (3) differential fitness among these contents (Dennett, 1991, p.200; 1995, p.343). While this chapter primarily utilizes these three criteria, the roles of (1) the imagination as a generator of diversity (Edelman, 2006, p.27) and (2) the degeneracy similar associations engender (Ibid. 33–34; Edelman and Gally, 2001) will be considered briefly in later chapters.

2.2 Meme theory

The foundations of the meme theory were laid in Richard Dawkins's *The Selfish Gene* (Dawkins, 1976/1989/2006b) where he attempted to illustrate the universality of replicator mechanics by using culture as an example (Ibid. Chapter 11). While calling this early conjecture a theory would be most inappropriate, framing cultural change as independent from actively intended results was an exciting new opening (Dennett, 1991, pp.202–203; Blackmore, 1999, Chapters 4 and 9). The efforts of Daniel Dennett to integrate memes in his the multiple drafts model of consciousness (Dennett 1991; 2005) and to popularize selectionism helped improve the theory (Dennett, 1995; 2006), although this version still contains notable problems (Distin, 2005, pp.77–91).

The meme theory claims that universal Darwinism applies to the evolution of culture because culture is claimed to be divisible into *memes*, which are replicators loosely analogous to genes. The term *meme* combines the terms *mimesis*, and gene (Dawkins, 1976/1989/2006b, p.192). Imitation is often conceived as the replicating mechanism of culture (Ibid. pp.189–193; Blackmore, 1999, pp.3–8). On the other hand, writers like Kate Distin would expand the list of possible mechanisms (Distin, 2005, Chapter 4). The connection with genes has two consequences: (1) memes are *selfish*, spreading mechanically whenever the conditions for their replication are fulfilled[4], and (2) attempting to etch

4 The intended definition of selfishness is Dawkins's who conceptualizes selfish entities as behaving in ways which increase their own welfare at the expense of others' (Dawkins 1976/1989/2006b, pp.4–5). It is the opposite of a behaviorist sense of altruism. The welfare of replicators means the probability of their continued survival (Ibid.).

the exact borders of individual memes is futile. This follows from Dawkins's conception of genes as 'any portion of chromosomal material that potentially lasts for enough generations to serve as an unit of natural selection' (Dawkins, 1976/1989/2006b, p.28). Serving as an unit of natural selection requires that a gene has phenotypic effects (Dawkins, 1986/2006a, p.304; Dennett, 1995, pp.125–126). Thus, Dawkins's definition also implies this demand.

Hence, the meme is *any amount* of cultural information which fulfills Dawkins's criterion of surviving a sufficient amount of generations to be exposed to (natural) selection. Similar definitions have been advocated by Blackmore (1999, pp.53–56) and Dennett (1995, pp.143, 344). Because such entities are only constituted by smaller pieces of information often capable of appearing independently or as parts of other memes, attempting to identify the smallest *significant* unit of culture seems trivial.

The problem of demarcating memes may be approcahed by applying comprehension about information in general. According to the standard definition of information, it is meaningful data (Floridi, 2005, p.353), or more precisely: 'Information is an objective (mind-independent) entity. It can be generated or carried by messages (words, sentences) or other products of cognizers (interpreters).' (Audi, 1999, p.435) The notion of information is often used in a loose sense which includes misinformation: meaningful and syntactically well-formed but untrue semantic contents (Floridi, 2005, p.366).[5] Since the fallibilism (Peirce, 1955, Chapters 2–4, 26; Popper, 1963/2002a, pp.154–160; 1972/1979, Chapter 2) to which this study adheres claims that all claims to ultimate truth are unjustifiable, the term 'information' is used here in this loose sense. Using these definitions, the smallest unit of culture would thus be the *infon* which means the smallest amount of meaningful and syntactically well-formed data (Ibid. 353).

The infon corresponds to the *cistron* which denotes the amount of DNA responsible for producing a single type of protein message (Dawkins, 1976/1989/2006b, p.28). If the analogy remains committed to Dawkins's definition of the gene, memes should not be identified with infons. In principle, an individual infon might become memetic but any given meme consists of one or more infons. A similar relationship on a higher level is illustrated by the concept of a *memeplex*: 'a collection of mutually supporting memes,

5 The Standard definition of information (SDI) does not include this requirement of truthfulness but Floridi has argued in its favor regarding semantic information (DOS) (Floridi, 2005). Semantic information denotes the field of information which is (1) declarative, (2) objective and (3) semantic (Ibid. p.352). Floridi's own definition (RSDI) includes the conditions of the standard definition but adds a new requirement of truthfulness. Thus, information (1) consists of data (2) when the data is syntactically well-formed and (3) meaningful, and (4) when these meanings are truth-abiding (Ibid. p.366.).

which tend to replicate together' (Heylighen and Chielens, 2009, pp.1, 9). This term is an abbreviation of the words 'co-adapted meme complex' (Dawkins, 1976/1989/2006b, pp.197–199; Blackmore, 1999, p.19;). The principle this notion embodies is simply that meanings become comprehensible and actions motivated only when they are supported by other information (Blackmore, 1999, pp.19–21; Heylighen and Chielens, 2009, pp.8–9).

Religious belief systems are often used as examples of memeplexes (e.g. Blackmore, 1999, Chapters 14 and 15; Dawkins, 1976/1989/2006b, pp.197–199; Dennett, 2006; Heylighen and Chielens, 2009, pp.8–9). For instance, Heylighen and Chielens list several candidates for the constituting memes of a memeplex for Catholicism. They include (1) belief in the existence of Hell, (2) belief in the creation of the world in seven days, and (3) belief in the virginity of Virgin Mary (Heylighen and Chielens, 2009, p.9). In principle, any of these beliefs could probably be held by any Christian – Catholic or not – although Catholics might be particularly prone to accept them. Another set of co-adapted memes listed by Heylighen and Chielens and likely held by a wider audience among Christians is (A) 'God is omnipotent', (B) 'God is good', (C) 'God punishes bad people', and (D) 'if you steal, you are bad' (Ibid. p.8).

Despite their stereotypicality, these examples may be used to illuminate the branching of forms which is characteristic of evolution. Let it be assumed that people who possess some combination of (A)–(D) are Christian. A subset of these people also pledge allegiance to one or more of conceptions (1)–(3) which are less strongly present outside this set. If a distinction becomes established between this subset and other Christians, two peculiar forms of Christian faith diverge. Those who only believe (A)–(D) but not (1)–(3) might be called *habitual Christians* who believe in the importance of believing in those beliefs (Dennett, 2006, Chapter 8) but do not actually commit themselves to doctrines. The other group may be nominated *fundamentalist Christians* because they assume the truth of Christian doctrines. The common ancestry of these two lineages might be traced by comparing common memes which characterize their constituting memeplexes' cores. If influences between the forms are severed, they should further diverge over time.

However, both sets of beliefs – (1)–(3) and (A)–(D) – are actually interconnected and may only become fully meaningful as a part of a common memeplex. For example, (2) becomes less comprehensible and convincing unless paired with (A). Likewise, comprehending the divine punishment (C) mentions would require believing (1) within the

Christian context.[6] In contrast, (B) might spread independently of (A) and (C) by infecting aboriginals with an existing conception of a similar enough god. To borrow an example of Dennett's, let these aboriginals revere *Feenoman*, a deity of the forest over whose eye color his followers disagree (Dennett, 1991, pp.82–85) as an overgod. While Feenoman may or may not actually denote a blue-eyed Tarzan look-alike with the power to heal the sick (Ibid. p.84), any conception of him hardly resembles the omnipotent, omniscient and agapeistic higher being Christians postulate. However, the closest parallel his worshippers would have for conceptualizing the claims of Christian missionaries would be Feenoman since both are the strongest spirits conceived by their respective followers. Thus, were missionaries to make claims (A)–(D), the aboriginals would relate this information to their existing conception of Feenoman.

Were those aboriginals under the conviction that Feenoman is fallible since a myth mentions him failing in a task, it would prevent them from assimilating (A). However, let the moral status of Feenoman have somehow remained a mystery until the missionaries arrived. Since there would now exist people seemingly claiming that Feenoman is good, and the aboriginals have always acknowledged Feenoman as a healer of the sick, (B) becomes assimilated as (B_F) 'Feenoman is good'. While details regarding the content of (B) have changed here, it retains most its components since both God and Feenoman designate a superior divine being to which benevolence is being attributed. The set to which the first component of such beliefs belongs remains the same, 'God-like beings', as does the second, 'good beings'. Such analysis is part of reverse engineering production rules inherent in a memeplex (Heylighen and Chielens, 2009, p.8). For the sake of argument, let (C) and (D) become blocked because (1) the Feenomanists believe Feenoman to also nurse the wicked and (2) because these imaginary people are required to pillage their neighbouring villages occasionally to survive.

No meme or memeplex transpires by being directly conveyed between people. Instead, they are always embodied (Dennett, 1995, pp.347–349). Basically, a meme is an information structure contained in its physical manifestation (Ibid. p.348). This physical aspect is a discrete unit integrated with its meme(s) and functions to preserve and propa-

6 How may habitual Christians maintain their integrity in this regard? They might have assimilated conceptions about alternative forms of karmic justice such as the eponymous belief in injustice entailing negative repercussions. Alternatively, such people might either not reflect on the matter, in which case the lack of coherence never becomes conscious, or they might pay lip service to (1) because they are aware of this conception and its inclusion in the Christian orthodoxy despite not having embraced it themselves.

gate those replicators (Dawkins, 1982/1999, p.114). Dennett utilizes the term *vehicle* (Dennett, 1991; 1995) inherited from Dawkins (1976/1989/2006b, pp.254–258; 1982/1999, pp.114–117) to describe this relationship. Other possible terms include *carrier* (Heylighen and Chielens, 2009), *interactor* (Hull, 1980; 2001; Distin, 2005;) and *host* (Blackmore, 2009). Within the confines of this study, Hull's terminology will be favoured. His account is more general – for example, Dawkins's use of vehicle has a genetical bias and contains unwelcome connotations (Hull, 2001, pp.32–33, Chapter 2). Unlike Dawkins's vehicles, interactors are also not limited to (phenotypic) products of replicators but also include the substrates embodying the replicators (Ibid. p.33). On the other hand, the people who possess specified memes will be referred to as their hosts (Blackmore, 2009).

While the role of replicators is emphasized in the account being discussed, they are not claimed to suffice by themselves. Hull's interactors highlight this duplicity of roles better than the vehicle- and carrier-terminology since vehicles are always subsidiary to their operators. Interaction with an environment and replication are equally necessary parts of the causal process of selection (Hull, 2001, p.23). This functional role in further reproduction bears stressing. For example, Distin criticizes Dennett for muddling the distinction between carriers and the capacity to propagate (Distin, 2005, pp.79–80).

Blackmore would dispose of any explicit differentation of a meme's genotype and phenotype (Blackmore, 1999, p.66). Heylighen and Chielens (2009, pp.7–8) as well as Dennett (1995, pp.347–348) suggest a similar abolishment since they all argue that a meme's replication is sufficiently mediated by its observable (*phenotypic*) effects. Distin opposes this view and proposes a distinction of a meme's genotype from its phenotype since she argues that exposure to phenotypic effects triggers reconstruction which only accidentally replicates the original information (Distin, 2005, pp.78–82, 92–95). This study advocates the latter view because it enables accommodating Sperber's critiques (Sperber, 1996, pp.100–106; 2000a). The role those critiques attribute to rational reconstruction during communication is also accepted as a realistic description of human interaction.

A simple example utilizing the religious illustrations adopted from Heylighen and Chielens (2009, pp.8–9) should clarify the assumed stance. Beliefs (C) 'God punishes bad people' and (D) 'if you steal, you are bad' form a production rule which allows inferring that (E) 'if you steal, God will punish you' (Ibid. p.8). When the relatively common disposition of not wishing harm upon oneself – or a more moral disposition of not wanting to

behave punishably – is accounted for in relation to (E), directive (α) 'Do not steal' is produced. Directive (α), generates an operational model (α') which consists of behavior devoid of stealing. This operational model (α') is the observable effect of (C) and (D) in suitably disposed people. Because (α') could be produced by multiple sets of varying beliefs like all the variants of (C) in which 'God' is replaced by another agent capable of effective punishment, successful transmission of (C) and (D) due to (α') is unreliable. The operational model (α') is phenotypic but due to degeneracy (Edelman and Gally, 2001; Edelman, 2006, pp.33–34), it is not univocally an expression of (C) and (D). Thus, while generated behavior like (α') may transmit memes, effectively functioning as their interactor (Hull, 2001, p.23), their ambiguity prevents them from sufficing. Replicatively significant forms of transmission other than such bare exposure are necessary (e.g. Blackmore, 1999, pp.61–62, 213–215; Dawkins, 2003, pp.143–147; Distin, 2005, pp.93–95).

The process of memes becoming spread may be divided into four phases which are (1) assimilation, (2) retention, (3) expression, and (4) transmission (Heylighen, 1999, pp.1–3; Heylighen and Chielens, 2009, pp.10–11). Assimilation signifies being recorded in a human mind due to its exposure to a meme's interactor. The meme must be (1) noticed, (2) understood, and (3) accepted (Heylighen, 1999, p.2). Retention requires that the thus created memory trace remains available for recall long enough to become expressed. Its value correlates with longetivity (Heylighen, 1999, p.2). These first two phases are required to conserve rather than proliferate a meme. Expression requires the hosting person to actively produce replicas of the meme's information content (Ibid.). Transmission determines the interactor, such as speech or text, a token possesses (Ibid. p.3). These two latter phases may produce multiples of the original replicator (Ibid.).

If the model above even approximately corresponds with the actual dynamics of the dissemination of culture, that process would contain replication and thus, potentially be subject to universal Darwinism. The rest of the criteria, variety and differential fitness, have also been analyzed by Heylighen (Heylighen, 1993; 1997; 1999). Table 2.1 has been reproduced according to one of Heylighen's articles (Heylighen, 1999, Table 1). The axes used to categorize selection criteria are (1) the phase of significance and (2) the system affecting a meme's success (Ibid. pp.3–4). The latter include different layers which specify the level at which the effects of the traits apply. Besides more traditional categories, Heylighen includes a level of meme-centered criteria.

stages/selectors	**Objective**	**Subjective**	**Intersubjective**	**Meme-centered**
Assimilation	distinctiveness	novelty simplicity coherence	authority formality	self-justification
Retention	invariance controllability	coherence utility	conformity	self-reinforcement intolerance
Expression			expressivity	proselytism
Transmission			publicity	proselytism
TABLE 2.1. Heylighen's selection criteria (Heylighen, 1999, p.5).				

All these criteria are primarily tied to their functions instead of certain structures a meme might contain (Heylighen, 1997). The values a meme attains in them are niche-dependent, and thus those figures are grouped accordingly (Ibid.). Studies supporting the inclusion of certain criteria, such as invariance (Van Overwalle and Heylighen, 1995) and formality (Heylighen and Dewaele, 1999), exist. Other factors, such as conformity, have been studied extensively (e.g. Boyd and Richerson, 1985, Chapter 7; Henrich and Boyd, 1998; 2002; Henrich and McElreath, 2003).

The layers of selection contain (1) objective, (2) subjective, (3) intersubjective, and (4) meme-centered factors. Objectiveness refers to factors independent of all individual memes and their hosts (Heylighen, 1999, p.4). For example, *invariance* may be further specified into (1) invariance over (sense) modalities, (2) invariance over time, and (3) invariance over persons (Heylighen, 1997). These three dimensions of invariance help specify an external entity. Likewise, *distinctiveness* is a trait shared by all such beings and *controllability* endorses it (Ibid.). These traits are universally required for distinguishing information structures, and higher values contribute to these structures' better preservation.

The subjective layer denotes the relationship of a (potential) host's and a meme's features (Heylighen, 1999, p.4). While *simplicity* could arguably be categorized as more of an objective criterion, it becomes relativized since an individual's existing information structures contribute to experienced simplicity. In effect, the term replaces Heylighen's earlier criterion of *learnability* (Heylighen, 1993, p.3–4), although *coherence* also retains some of that criterion's traits. *Novelty* helps capture attention but anything may only be novel compared to previous experiences. Similarly, *utility* only exists in relation to the contexts in which agents act and the experienced needs of these people.

Including a social dimension produces intersubjective criteria such as *authority*, *conformity* and *publicity*. Less obviously social are *formality* and *expressivity*. Both relate to the encoding of intended meaning: formality specifies unambiguity due to adherence to shared rules of expression (Heylighen, 1993, p.4; 1997) while expressivity only requires the capacity to become expressed in a comprehensible medium (Ibid.). While these are necessary criteria bound to intersubjective customs, the level of *salience* an expression has further contributes to its propagation (Heylighen, 1993, p.4). Another unlisted trait is *group utility* which affects every phase and specifies the adaptedness a group may derive by assuming a meme (Heylighen, 1999, p.4).

The last layer is meme-centeredness which specifies *selfish* features in the sense specified earlier (Dawkins, 1976/1989/2006b, pp.4–5). Such features contribute to overall fitness but only befit the memes' increased propagation (Heylighen, 1999, p.5). Because any external interests are eschewed, memes with high values in meme-centered criteria are called *parasitic* (Brodie, 1996; Dawkins, 2003, Chapter 3.2) or *selfish* (Heylighen, 1992; 1999) despite all memes being both parasitic (Dennett, 1991, p.221; 1995, p.341) and selfish (Distin, 2005, Chapter 8). All meme-centered criteria convey the sentiment that memes contain instructions or motivators which induce behavior which is only benefits the memes. For example, *self-reinforcement* might mean an instruction to rehearse the content regularly (Heylighen, 1999, p.5). This would increase retention but benefits derived by the hosts would remain dependent on the other features of the meme such as its utility to them.

According to Heylighen, calculating a meme's fitness value is possible by using the following formula: $F(m) = A(m) \cdot R(m) \cdot E(m) \cdot T(m)$; $0 \leq A \leq 1$, $0 \leq R \leq 1$, $E \geq 0$, $T \geq 0$. F(m) stands for a given meme's fitness while the rest depict ratios of successfully completing corresponding phases of selection: A(m) for assimilation, R(m) for retention, E(m) for expression, and T(m) for transmission (Heylighen, 1999, p.3; Heylighen and Chielens, 2009, p.12). A and R have values below 1 because the corresponding phases have a conservative function while E and T are not similarly limited and allow increases in the number of tokens. The limitation on the potential amount of meme hosts is also considered when calculating the rate of propagation: $dN/dt = (F-1) \cdot N(1-[N/K])$; $N \leq K$; $\{F, K, N\} \geq 0$. Here, dN/dt represents the amount of meme tokens at a given point of time. The value of N is a function of time, (t); meme fitness, F; and the totality of possible hosts, K; as well as the number of priorly existing tokens, N (Ibid.).

The meaning of the first formula may be clarified using the dissemination of fist bumping behavior as a hypothetical example. Suppose that M(x) denotes the set of all memes. Let fist bumps be a meme, M(b), which consists of the instructions (1) 'when you encounter a familiar person known to belong to the set of street credible people, S(p), greet them by raising a clenched fist horizontally towards them' and (2) 'if a familiar person known to belong to the set of street credible people, S(p), raises a cleched fist horizontally towards you, greet them by forming a fist and press your knuckles against theirs momentarily.' While the formality of these instructions may make them seem artificial, processing them consciously would not be required. M(b) is a memeplex since instructions (1) and (2) would likely not propagate individually since they only make sense together. Instructions (1) and (2) limit fist bumping to greeting but adjustments could easily broaden the applicability of such behavior, including moments of triumph in (loosely) co-ordinated operations.

Now, let it be assumed that witnessing a fist bump or being exposed to the idea of such behavior via a description induces learning instructions (1) and (2), which comprise M(b), half the time. Thus, A(b) = 0.5 since the assimilation rate would be this high. Similarly, half the people who assimilated this knowledge of an orthodox fist bump would retain the instructions in their memory until the associated behavior is expressed at least once. Many would never express it – possibly because they would be unmoticated to belong to set S(p) of street credible people and thus being unmotivated to bump fists. In any case, R(b) = 0.5 as well. On average, anyone deciding to express this behavior does so ten times in a set amount of time such as a week. This way, E(b) = 10. Each time M(b) is expressed, an average of five people are exposed to it. Thus, T(b) = 5. From these values, the fitness value of M(b) may be calculated as follows: $F(b) = 0{,}5 \times 0{,}5 \times 10 \times 5 = 12{,}5$.

This value, F(b) may be utilized to calculate the rate of dissemination for this meme which generates the appropriate behavior called 'fist bumping'. This rate is depicted by the function $N(b, t) = F(b) \cdot N(b, t-1)$, when F(b) = 12,5. Respectively, N(b, t) is the amount of tokens at time (t) in weeks after (b)'s formation. F(b) is used to multiply the existing amount of tokens at each interval of (t) since its value represents the average amount of prodigy produced by each existing token under the prevailing circumstances (Heylighen and Chielens, 2009, p.12).

This example also illustrates Blackmore's distinction between *copying-the-product* and *copying-the-instruction* (Blackmore, 1999, pp.61–62, 213–215). Copying the

product means reconstruing the instructions behind detected behavior and artifacts according to observations (Distin, 2005, pp.93–95). In contrast, copying the instruction signifies being directly exposed to and replicating the behavior-inducing information (Ibid. 93; Blackmore, 1999, p.61). Since there is no phase of reconstruction and the information is often much better encoded and often digitised (Blackmore, 1999, pp.213–214), successful replication becomes much more likely. In the example above, assimilating M(b) was mentioned to become possible due to exposure to the behavior of fist bumping itself or a description of it. Successfully interpreting the intent to greet as well as significant features of the behavior would require reconstruing the intentions of those participating in such behavior (Dennett, 1989; Baron-Cohen, 1995; Sperber, 2000a, pp.170–172). In contrast, reading a description which explains the actions' rationale and appropriate conduct would bypass this extra phase.

Distin opposes Blackmore's view that copying the product is replicating its originating memes because the information derived is a reconstruction dependent on prior knowledge (Distin, 2005, pp.94–95). On the other hand, she dismisses Dan Sperber's critique of the transmission of culture as reconstruction instead of replication (Sperber, 1996, pp.100–106; 2000a) by admitting such reconstruction as a possible mechanism of replication (Distin, 2005, pp.107–108). While this expansion beyond crude imitation as the mechanism of transmission seems appropriate, it would also seem to allow inferring from a product. After all, the selection environment of memes also includes other present memes, including the prior knowledge possessed by potential hosts (Dennett, 1995, p.349; Heylighen and Chielens, 2009, p.13). Thus, memes reconstruable with reliability comparable to replication under the influence of certain mutualist memes (Ibid.) may become reliably copied by exposure to their product when the mutualists enabling accurate reconstruction of their information content are widespread. However, since this situation requires quite the definite selection environment for which a meme must be fit, copying the instruction remains a more reliable form of transmission (Blackmore, 1999, pp.213–214).

While the criteria listed by Heylighen (Heylighen 1993; 1997; 1999) are tentative, his model regarding the four phases of meme selection (Heylighen, 1999; Heylighen and Chielens, 2009, pp.10–13) constructs a fine-seeming framework to conceptualize the transmission and propagation of memes. This is a welcome demystification of the process of replication which is admittedly too often presupposed instead of being explained. The

evaluation of the suggested criteria as well as the other possible factors affecting selection may be conducted by utilizing the results of established sciences such as sociology and psychology when such specification of the process are available. For example, features of memory affecting successful assimilation (absent-mindedness) and retention (e.g. transience and misattribution) are listed in *The Seven Sins of Memory* (Schachter 2001). Likewise, factors allowing forced assimilation may be analyzed by researching what circumstances allow affecting that phase (e.g. Taylor 2004).

This stresses the importance of Dan Sperber's claim that a deep understanding of cultural evolution is incompatible with shallow psychology (Sperber 2006). Also, comprehending at least the basics of sociology and anthropology would be desirable as well because cultural evolution is a process, the researching of which requires dissecting the relationship between (1) the capacities possessd by individuals and (2) the information they make publicly available. The importance of such comprehensive background knowledge about existing theories about culture is stressed by criticisms claiming that meme theorists oversimplify (e.g. Plotkin, 2000) and dismiss existing research on culture (e.g. Bloch, 2000; Kuper, 2000). However, the critiques being discussed within the confines of this study will be limited to Dan Sperber's counterarguments against selectionist approaches to culture (Sperber 1996; 2000a). The reasons for this are simply (1) the impossibility of answering all the critiques which have been voiced and (2) Sperber's framework being deemed plausible. Because his framework is accepted, what is attempted is refuting the claim that selectionism is incompatible with that framework.

2.3 Criticism of the meme theory

If it becomes established that universal Darwinism is an inappropriate approach to explaining cultural phenomena, the meme theory becomes falsified. Thus, demonstrating that culture in the sense used by meme theorists does not fulfill some of the requirements of universal Darwinism – (1) variety, (2) replication or (3) differential fitness (Dennett, 1991, p.200; 1995, p.343) – is a justified approach when attempting to refute the meme theory.

Dan Sperber's approach is questioning the replication of cultural contents. While he admits the existence of replication in culture, he also claims such cases to be exceptions to the general rule that culture does not replicate (Sperber, 1996, pp.101, 103–104).

Sperber's claim is that generally cultural contents transform when transmitted and that these transformations invalidate claims of replication (Ibid. p.101). Another prong of Sperber's critique is the claim that representations of observed phenomena are construed according to prior knowledge instead of being transmitted as such (Sperber 2000a).

Kate Distin has distilled Sperber's criticism into three counterarguments against memes (Distin, 2005, pp.102–104): (1) The high mutation rate during the transmission of culture prevents the effects of selection from cumulating (Sperber, 1996, pp.102–103). (2) Despite the similarity of content between representations, considering them tokens of a type is an oversimplification which underestimates their differences (Ibid. pp.29, 118). (3) Representations are not a code which becomes replicated during transmission; rather, they are reconstrued by inference and comprehension guided by considerations of relevance (Ibid. 106; Sperber, 2000a).

Both critiques (1) and (3) are motivated by Sperber's conception of communication as relevance-based reconstruction of intended meaning (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012). While the relevance theory of communication is elaborated in section 3.1, a brief introduction is provided here to clarify Sperber's position because Sperber does not contend to problematize but rather, asserts his own position simultaneously (Distin, 2005, p.104). Basically, Sperber claims that the meme theory does not correspond with the facts because his own theory does, and that the two are incompatible. Since the relevance theory is indeed quite well-corroborated (e.g. Jorgensen, Miller and Sperber, 1984; Happé, 1993; Wilson and Sperber, 2012, Chapters 13 and 14), Sperber's critiques (1) and (3) will be answered by asserting the relevance theory's compatibility with selectionism. Meanwhile, critique (2) is connected to Sperber's general critique of abstracting away causal connectedness (Sperber, 1996, Chapter 2, pp.70–76). Answering it requires elaborating the role and indispensability of these abstractions.

All three critiques are also inherent in Sperber's test for replication (Distin, 2005, pp.104–108; Sperber, 2000a). This test expands the criteria tentatively set by Dawkins's original test (Dawkins, 2003, Chapter 3.1) in which an intelligent observer is provided tokens of a lineage of meme products. If this agent fails to identify their order of descent, the tokens are functional replicas despite their differences (Ibid. 143–145; Sperber, 2000a, p.168). Since Sperber notes that some non-imitative and informationless behavior such as laughter fulfills this criterion (Ibid.), he proposes further specifications. To pass Sperber's

test and exemplify true replication, a descendant of (A) called (B) must (I) be caused by (A), (II) resemble (A) in relevant aspects, and (III) be generated by a process which conveys the information responsible for this resemblence with (A) from (A) itself (Ibid. p.169). Sperber's critique (1) attacks criterion (II) by claiming sufficient resemblance being the exception instead of the rule (Sperber, 1996, pp.100–106). Meanwhile, critique (2) questions the fulfillment of criterion (I). It affirms the necessity of causal descent and contrasts this view with the more generalizing conjectures provided by meme theorists among others (Ibid. Chapter 2). Finally, critique (3) engages criterion (III) by claiming (A)'s insufficiency in engendering replication and asserting the role of prior capacities responsible for its successful imitation (Sperber, 2000a, pp.169–173).

These prior capabilities include the propensity to attribute intentions to (human) agents (Sperber, 2000a, pp.170–171). An action's intended features may only be inferred thus (Ibid.), and even verbal descriptions require drawing such inferences (Ibid. p.171). This is the main tenet of the relevance theory of communication (Sperber and Wilson 1986/1995; Wilson and Sperber, 2004; 2012). All such inferences are reconstructions of intended meaning and while relatively reliable, they always transform the transmitted information since a communicator's and a receiver's cognitive environments are never strictly identical (Sperber 1996, pp.33–34).

This process of interpretation consists of three subtasks: (a) 'constructing an appropriate hypothesis of explicit content', (b) 'constructing an appropriate hypothesis about the intended contextual assumptions', and (c) 'constructing an appropriate hypothesis about the intended contextual implications' (Wilson and Sperber, 2004, p.615). These processes are parallel; not sequential (Ibid.). They are guided by two general principles: The Cognitive Principle of Relevance is that 'human cognition tends to be geared to the maximization of relevance' (Ibid. p.610; Wilson and Sperber, 2012, p.7). In turn, the Communicative Principle of Relevance complements it: 'every ostensive stimulus conveys a presumption of its own optimal relevance.' (Wilson and Sperber, 2004, p.612; 2012, p.6).

Both principles require an elaboration of Sperber and Wilson's concept of *relevance* to become intelligible. It means the degree of new information's usefulness to the receiving subject within the available context (Sperber and Wilson, 1986/1995, p.48). The disposition to optimize during exchanges ensures that the least energy-consuming possibilities should be favoured during both expression and interpretation (Ibid. p.125;

Wilson and Sperber, 2004, p.609; 2012, pp.62–63, 88). All these processes rely on prior understanding of the shared context and the other party's intentions (Sperber and Wilson, 1986/1995, pp. 38–46; Wilson and Sperber, 2004, pp.612, 614–623).

This preliminary introduction ought to suffice to explain the background of Sperber's critiques (1) and (3). Effectively, Sperber claims that the reconstruction of meaning is more unreliable than strict decoding of meaning would be due to the former's reliance on prior conceptions and the context of communication (Wilson and Sperber, 2012, pp.98–101). Despite being more expressive than actual codes, natural languages are simply incapable of encoding exact meanings sufficiently reliably, and this incapability always generates ambiguity (Ibid. pp.98–99).

If this framework for structuring communication is correct, the meme theory must adapt or be discarded. Countering critique (1) demands an explanation of how reliable preservation of defining information content is possible under the constraints of the relevance theory. Likewise, answering critique (3) necessitates explaining how criterion (III) shown above is fulfilled despite the role played by prior conceptions. The general outline of the provided responses resembles Distin's answer to Sperber's critique (Distin, 2005, pp.106–108). Basically, prior conceptions are accepted as elements of successful communication, and their role is elaborated to justify calling such transmission replication.

To a degree, critiques (2) and (3) are supporting arguments for critique (1). Thus, they will be dismantled first, beginning with the more central critique (3). As Distin notes, Sperber claims that the reconstruction process he describes is an *alternative* to replication (Distin, 2005, p.107). This, however, is far from obvious. Another perspective, assumed by Distin, is viewing such reconstruction as the *means* of replication (Ibid.). The justification of this approach demands explaining how the criteria of replication set by Sperber are fulfilled despite the role prior conceptions have (Sperber, 2000a, p.169).

Let H_1 denote the host carrying a meme (m), while H_2 designates the receiver of input acting as the interactor of (m) and realized by H_1. The minimum criteria of meme transmission are (1) that the designated information (m) is retained by a host such as H_1 during a moment in time (t_1), and (2) that the same information may be located in the memory of another individual such as H_2 at a later date (t_2) (3) due to interaction between H_1 and H_2. The actual process of transmission is effectively irrelevant although the necessity of H_2 not possessing (m) at time (t_1) must be considered.

Distin notes how the reconstruction process is initiated by an outside influence, be it another person or a public product (Distin, 2005, p.108). Effectively, an ostensive stimulus is provided, and such ostensive stimuli trigger a search for intentions and the corresponding meanings (Sperber and Wilson, 1986/1995, pp.48–54; Wilson and Sperber, 2004, pp.611–612). The inferences drawn by relating existing information to features of the ostensive stimuli provide new information by communicating contextually appropriate intentions reliably. In other words, acts of communication provide *positive cognitive effects* (Sperber and Wilson, 1986/1995, pp.265–266; Wilson and Sperber, 2012, pp.62–64), or increases in the quantity of the information available to consciousness. Thus, new information becomes available to the receiver, and this new information (1) resembles the intended meaning in relevant aspects, and (2) is caused by the interactor of previously existing information contained elsewhere. While prior conceptions do thus have a role in such successful communication, they will likely be considered when the interacting expressions are formulated. After all, communicators must also consider the relevance of the stimuli they provide from the viewpoint of their intended audience (Sperber and Wilson, 1986/1995, pp.50–64).

Thus, to convey (m) to H_2, H_1 would construe an interactor embodying (m) and specifically designed to cater to H_2's prior conceptions as they are conceived by H_1. The conceptions H_1 represents include H_2's conceptions about H_1 and thus, H_1 has a (limited) awareness of H_2 how perceives them. Interactors of (m) need not be identical to convey the relevant information which constitutes (m). This is probably where both some meme theorists and their critics stray because the replicating meme itself is identified with its publicly available interactors. Especially problematic cases include public products which are limited to a single appearance, such as the wording of a book, which may not be tailored for specific receivers' cognitive environments. In such cases Sperber's critique (1) indeed applies more forcefully because critique (3) supports it whenever incompatible prior conceptions exist among receivers. However, as long as compatible receivers also exist, these public products can reliably propagate the information they contain. Mutant misunderstandings will emerge but the original conception should also become more widespread. Since all these public products are also (for the time being) produced by human agents, their design incorporates the authors' suppositions of their intended audiences' existing conceptions.

Another problem with Sperber's critique (3) originates from the examples he chose to illustrate the issues of replication. These examples include (a) a random pattern drawn in one stroke, and (b) a five-branched star drawn without lifting the pen (Sperber, 2000a, pp.165–166). Sperber's analysis regarding (a) resembles Dawkins's because they regard the produced replicas as nothing but imitations of the perceived form and thus particularly susceptible to degeneration (Ibid. pp.166–167; Dawkins, 2003, pp.143–145). Copying (a) is merely copying the product in Blackmore's sense (Blackmore, 1999, pp.61–62). The analysis of (b), however, diverges from Dawkins's analysis of transmitting instructions regarding the crafting of Chinese junks (Dawkins, 2003, pp.144–146). Sperber does Dawkins no justice with his example (b) because the instructions in its case are implicit while Dawkins primarily discusses explicit, verbal instructions and only makes allowances to instructive behavior (Ibid.). While even verbal rules may become corrupted due to memory failures and misunderstandings, the above discussion regarding the relevance theory should suffice to explain such cases. Furthermore, particularly hard to remember information is simply unfit for retention unless there exist factors which enhance its preservation. Thus, the cases of people failing to remember such instructions can be easily accommodated by considering Heylighen's criteria (Heylighen, 1993; 1997; 1999) for the successful transmission of memes.

Even in the case of nonverbal, even implicit instructions, Sperber's example (b) remains problematic. A five-branched star is an archetypical shape, the drawing instructions of which are already widespread. While the prior capacities discussed by Sperber are not limited to realizing the star being copied being a token producible using familiar instructions, this would seem to be the default situation which biases evaluations of its plausibility. Especially Sperber's allusion to prior knowledge of five-branched stars makes this obvious: 'It [performing well] is an ability to recognize and re-produce, using, for this, knowledge of the five-branched star type that they already possessed before encountering the token.' (Sperber, 2000a, p.170) Less problematic forms of prior knowledge would include the capacity to realize the star being a product of human activity and simulating the easiest conceivable way to draw one.

Specifically, Sperber's example (b) supposes that the inferences are drawn according to perceptions of the finished product while Dawkins discusses observing a Chinese junk being folded (Dawkins, 2003, p.145). Even if the self-normalization Dawkins attri-

butes to instructions (Ibid.), was inherent in observed products, it should be understood as a form of reverse engineering. By backtracking the folding process of an origami based on the assumed intentionality of its original creator, uncovering instructions for producing necessary traits which are not visible in the finished product should be possible. While such reverse engineering would still require attributing intentionality and would thus exemplify the Sperberian framework, Sperber contends to remark how inferring such instructions from observing a finished Chinese junk would probably fail miserably (Sperber, 2000a, p.170). There exists a clear disparency. The actual discussion regarding the role of understanding another's intentionality (Ibid. pp.170–172) is legit, and its compatibility with the meme theory has been proven above. Sperber successfully dismisses only a naive replicationism founded on nothing but a limited conception of imitation (Distin, 2005, pp.107–108). However, other means of transmission, including processes which integrate rational reconstructruction, are available (Ibid.).

These problems may be illustrated using the *henohenomoheji* (see figure 2.1). It is a picture of a face which consists of certain Japanese *hiragana* characters. To those capable of decoding the fact that this form consists of writing, extra information is conveyed by the picture. Most non-Japanese will likely only discern the general shape of a face and not notice the fact that it consists of letters because they lack necessary information about the *hiragana* alphabet. Thus, these non-Japanese who attempt to copy the *henohenomoheji* may only rely on their implicit knowledge that the picture's design is intentional instead of being random since it has a clear structure and resembles a familiar object. However, they will probably fail to infer (1) the sequence of drawing one, as well as (2) the meaning of its parts. For example, the *he*-letters which represent the eyebrows and the mouth might become drawn as symmetrical reversed V-shapes. From the perspective of an abstract face pattern, this difference might be negligible. However, their function as letters would become disrupted by such changes – and only those who realize the symbols being letters could reliably avoid this mistake.

FIGURE 2.1. The *Henohenomoheji*. Picture acquired from Wikipedia (Wikipedia, 2005).

In summary, Sperber's examples (a) and especially (b), do not present all significant aspects of the problem of replication. Example (b) fails to reflect the relevant traits of Dawkins's original example (Dawkins, 2003, p.145), and it also depicts a biasingly familiar shape and process. Utilizing another example, the *henohenomoheji*, helps clarify this. This shape contains two aspects, only one of which is obvious and inferable using innate human capacities. The other is only available to those with a culturally conveyed prior understanding of Japanese alphabets. Sperber discusses both types of prior capability together when analyzing his example (b), a five-branched star (Sperber, 2000a, pp.167–171). The actual transmission of any new information in example (b) is thus questionable because a drawn star might simply motivate people to copy an already familiar shape. Hence, Sperber attempts to justify his conclusions (1) that prior conceptions matter decisively and (2) that this invalidates the true replication of cultural information (Ibid. pp.172–173) using an example in which those prior conceptions already include the content being transmitted.[7] Also, the cues to infer all the necessary instructions to draw a star are in plain sight in example (b), while uncovering them in a Chinese junk requires dismantling it systemically.

With this, it should become obvious that Sperber has overestimated the role of prior knowledge and underestimated and mispresented the possibilities of meme transmission. Thus, Distin's conclusions that Sperber's critique (3) should be seen as a demand to elaborate the processes of meme transmission and that reconstruction is viable as a means of replication (Distin, 2005, pp.107–108) seem valid. Also, since reconstruction seems capable of reliably transmitting the intended information, critique (3) does not support critique (1) which would then require a different justification.

Sperber seems to sway in his conception of replication. Originally, he demanded the perfect copying of all traits and focused on the articulated contents (Sperber, 1996, pp.103–105). This is made apparent by how he used chain-letters as an example of actual replication (Ibid. p.103). Later, Sperber would admit the sufficiency of non-identical tokens which share properties perveived as relevant (Sperber, 2000a, 167–169). This latter view is much more realistic since it allows the study of features instead of complete types, although successfully demarcating between the two will be demonstrated to be quite

7 In contrast, the culture-bound aspect of the *henohenomoheji* may be taught separately from its more general aspect of representing a face. Personally inferring a way to produce simulacra of that picture would convey no such instructions. This process should also increase the meme's fidelity since the exact characters constituting the figure receive meanings and since the image's name contains a mnemonic to those capable of connecting its syllables with the corresponding letters.

strenous. For example, Sperber's original view would deny replication when the character fulfilling a certain role in a story is switched. In contrast, the latter view should allow viewing the persistence of that role across versions of the story as the traits inherent in the role being selected. A possible role like this would consist of a character in the story being (a) an antagonistic (b) predator (c) who eats humans. This description would fit multiple possible characters and the distribution of such characters across stories would certify its fitness which could be further explained by researching the psychology and sociology behind the attractiveness of utilizing such conceptions.

Thus, critique (2) is actually also a supporting argument for critique (1) because relevant differences may only be identified by relating them to the type of meme being studied. According to critique (2), discussing memes as types is unjustified if a type accommodates multiple variants instead of only identical tokens (Sperber, 1996, pp.29, 118). This requirement is closely connected with Sperber's earlier conception of replication as the production of identical copies. Thus, being able to justify the position that 'studying memes as abstractions, which include multiple variations of a common trait, is reasonable' would abolish critique (2) and severely weaken critique (1).

The function such abstractions have is framing research. The other possibility, only studying single tokens and identifying their lineages, was Sperber's ideal during *Explaining Culture* (Sperber, 1996). However, even he has seemingly changed his view after introducing the terminology of *Cognitive Causal Chains* (Sperber, 2001; 2006). While the relevance of that concept will be detailed in section 3.2, its significance here is how Sperber has admitted more abstract generalizations in the form of *Cultural Cognitive Causal Chains* (Sperber, 2001, pp.309–312; 2006, pp.437–439). As the terminology implies, the focus of his project remains unchanged: only causally linked representations are to be analyzed together (Sperber, 1996, Chapters 2 and 3). The suggested approach which would also include studying traits which are common to diverse lineages dissents.

From the viewpoint of the analogy to biological evolution, Sperber's position is far more sustainable. However, the echoes of Steve Jones's warning regarding any analogy with evolution should prevent regarding this as a merit unless independent justification for such a claim is supplied. Biological evolution contains mechanisms which prevent reproduction between species. As Hull notes, 'That these populations remain potentially interbreeding means that they are potentially a single individual.' (Hull, 1980, p.324) By

individual, Hull means a single entity from the perspective of the dynamics of replication (Ibid. pp.322–324). Since species are nothing but sufficiently divergent lineages, using Hull's terminology (Hull, 1980, pp.327–329), the species of memes which may combine and exchange information are not really species in this sense. The lineages do not remain distinct but rather, become larger entities by interrelating because a lineage is *any* entity which 'changes indefinitely through time as a result of replication and interaction' (Ibid. p.327). However, new divergences among the descendants of such converged lineages may also branch into new lineages (even when they merely shed the shared influence again).

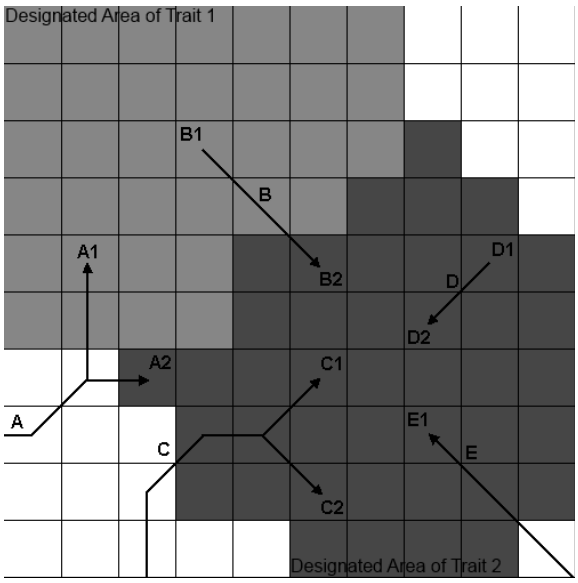
In the case of culture, such interrelating is relatively unhindered, which generates ambiguous ancestries (Gabora, 2011, pp.12–13). This clearly complicates Sperber's project of tracing causal ancestries. The alternative provided here consists of studying commonalities which affect the propagation of different information units. These generalizations are unbound by causal ancestries although they also play an important but not necessary role when one demarcates between species of representations. Their function is merely delineating the approximate probabilities of differentially large transformations.

In effect, what is being proposed is studying of (1) the fitness of traits shared by multiple lineages and (2) the interaction of those lineages. A crucial concept in understanding this project is *Design Space* (Dennett, 1995, Chapters 3 and 6; 2013, Chapter 38). The Design Space is the model of all possible structures, and more similar internal designs are depicted closer to each other (Dennett, 1995, pp.135–144; 2013, p.218). Basically, if something can be encoded using a limited system of symbols which may be further combined into new symbolic representations, its structure can be expressed using those symbols. The idea somewhat resembles early Wittgenstein's picture theory of language (Wittgenstein 1922). However, the function of the Design Space is not enabling the univocality of expressions but illustrating relations of similarity and disparity between distinct structures.

Any possible information structure may be located in the Design Space (Dennett, 1995, p.135; 2013, p.218–219). Within this model, more divergent structures are further away from each other than more similar ones, and minimally different options neighbour one another (Dennett, 1995, p.125). The amount of dimensions which is necessitated by such relations exceeds the everyday experience connected with the notion of space but imagining the Design Space in relevant detail using three dimensions suffices (Dennett, 2013, p.218). This study uses two-dimensional models for demonstration purposes.

Since structures which share traits are relatively close to each other in the Design Space on axes connected to the substructures which constitute those traits, inspecting the accrual of cultural contents in certain positions informs of the effects those traits have on fitness. Likewise, such traits have proven fit themselves and the reasons may be studied separately. However, these traits are not identified with levels in the earlier criteria for selection (Heylighen, 1993; 1997; 1999; Heylighen and Chielens, 2009, pp.17–20). Rather, they contribute to the latter according to the relationship between them and the selection environment. Examples of these traits in the case of supernatural beliefs could include anthropomorphism and animism. Both attributing archetypically human features and attributing pseudo-life to abstractions is relatively widespread traits among religions. For example, the earlier example (B) 'God is good' (Heylighen and Chielens, 2009, p.8) attributes an ideal of human morality to something often conceived as an invisible, omnipotent force. A possible explanation for this phenomenon could invoke evolutionary heuristics such as the intentional stance (Dennett, 1987; Baron-Cohen, 1995). In effect, conceptions which incolve mental features, such as motives, the audience recognizes appear simpler and are possibly more expressible and become thus understood more easily. This ease then increases the total representation of these conceptions to the detriment of more complicated alternatives, the thriving of which would require a widespread distribution of certain other assumptions such as more rigorous delimitations on potentially conscious entities.

Thus, it is argued that the generalizations criticized by Sperber allow studying more universal features which affect the direction of change in Design Space. While tracking the causal ancestries of representations allows constructing their genealogies, these genealogies depend on explanations regarding relations of the structure of the constituting representations and their environment. Sperber's own attractors (Sperber, 1996, pp.106–118) could be regarded as a similar generalization concerning the factors which affect movement in Design Space. A vital difference is how he excludes the possibility of nominating areas in Design Space in order to compose sets from the corresponding representations regardless of their causal connectedness (or lack thereof). However, such generalizations integrate the blending and exchange of information and the resulting ambiguous ancestries (Gabora, 2011, pp.12–13) better because what is being researched are not particular representations but the traits shared by them and capable of crossing over. Consequently, this approach may also allow formalizing those traits.



Explanations:

1. Emphasized areas:
a) Designated Areas of Traits 1 and 2
- Contain designs which share the specified traits

2. Drawings:
a) Arrows A, B, C, D and E
- Trajectories of change between designs
b) Cells A1, A2, B1, B2, C1, C2, D1, D2 and E1
- Example designs which contain the specified traits

FIGURE 2.2. Demonstration of the relevance of generalizations questioned by Sperber (Sperber 1996, pp.29, 118) using the grid he also favors (Ibid. pp.109, 112).

The above discussion regarding a taxonomy of traits and their significance will be illustrated using the grid Sperber utilizes to demonstrate the dynamics of attraction. His model consists of a hundred different possible representations which are divided between ten lines of ten possibilities each (Sperber, 1996, p.109). Some cells are designated as attractors while lines representing the trajectories of representation chains traverse across the grid (Ibid.). Here, the latter are denoted by arrows which show their direction of descent, while some areas are colored in shades of grey to designate designs which contain mutually incompatible traits. The visible ends of trajectories have been named to mitigate discussing their relations with one another.

Figure 2.2 contains a hypothetical representation of a minuscule portion of Design Space. It is divided between two incompatible traits but also contains neutral areas which contain neither. Seven alternative trajectories related to these traits are also depicted in the figure. While this demonstration is primarily an illustration of the answer to Sperber's critique (2), it also illuminates the connection between critiques (1) and (2), and thus provides a way to answer critique (1) as well.

Each of these trajectories depicts different possibilities: (1) The trajectories towards (A1) and (A2) diverge from a shared lineage but each reaches the area of a different trait. (2) Trajectory (B) illustrates a change from Trait 1 to Trait 2. (3) (C1) and

(C2) result from another diverging trajectory yet both endpoints remain within the same trait group. (4) Trajectory (D) only changes irrelevant contents within a single trait category. Finally, (5) trajectory (E) stems from a neutral area separated from Trait 1 by Trait 2 but heads towards the former.

The reason trajectory (A) branches would be of interest, were a similar trajectory located in reality because unlike (C), (A) involves two representations which contain incompatible traits. Why either trait is selected for requires understanding how non-objective criteria affect these transformations. Tokens (A1) and (A2) are differentially compatible with different selection environments, including people with deviating preconceptions. For example, the person who embraces (A2) may have prior beliefs which contradict all possible beliefs which contain Trait 1. This would effectively eliminate the possibility of this person embracing any of those beliefs. Likewise, (A1) might become selected for elsewhere due to its better compatibility which need not be limited to non-contradictoriness. It could maybe appear simpler in light of the host's previous conceptions. In other words, (A2) might be selected at individual level due to coherence and (A1) due to simplicity (Heylighen, 1999, p.4). The reasons provided are merely hypothetical examples but they should suffice to illustrate how the fitness of different traits is never universal.

In the case of trajectory (B), there exists a transformation from contents which include Trait 1 to contents which contain Trait 2 instead. In such a case, a token embodying Trait 1 would be eliminated in favor of another which contains Trait 2. In other words, Trait 2 would become selected over Trait 1 in this occasion. Were trajectories like (B) common, Trait 2 would flourish at the expense of Trait 1.

Heylighen and Chielens list three types of relationships between memes: (1) mutualists, (2) competitors, and (3) predator and prey (Heylighen and Chielens, 2009, p.13). While the reason Trait 2 becomes chosen may be due to existing mutualists among potential hosts, like in the example which explains (A1) and (A2)'s divergence, trajectory (B) emphasizes how the two traits are likely either competitors or predator and prey. Since the two are incompatible by definition, they cannot be mutually reinforcing mutualists (Ibid). Despite this incompatibility, their relationship is not necessarily one of the two other alternatives either. Competition requires that increases in a trait's representation reduce the share of others (Ibid.). Since neutral possibilities also exist, neither increases nor decreases in one example trait necessarily affect the other. Meanwhile, a predator-prey relation is

essentially a sub-type of competition: the predator flourishes due to the conversion of its prey (Heylighen and Chielens, 2009, p.13). When the prey representation is already held by a person, these hosts are particularly susceptible to replace that conception with its predator if subjected to the latter.

Such classifications of relations between designs would be impossible without construing general types like the traits being discussed. However, they allow drawing more general inferences than mere attractors, yet this account should be compatible with the latter. If (B) indeed depicts a predatory relationship between Traits 1 and 2, all forms which contain Trait 1 would be especially susceptible to convert to forms containing Trait 2 instead. Were this depicted using only attractors, an attractor which affects representations with contents which correspond to Trait 1 should be postulated somewhere in the territory occupied by the designs which contain Trait 2. The exact location of this attractor and the directedness of its pull would pose problems. Instead of attracting nearby designs equally like those Sperber discusses (Sperber, 1996, pp.106–118; Claidière and Sperber, 2007), such an attractor would only influence a particular set of representations not in its vicinity.

Trajectories (C) and (D) resemble each other, and their primary role is to exemplify how generalizations also answer Sperber's critique (1) which concerns the frequency of mutations (Sperber, 1996, pp.102–103). Both trajectories involve mutations but those transformations remain within the set of designs which contain the same trait variant. The branches of (C), (C1) and (C1), differ in some regard but both contain Trait 1. Likewise, (D2) is a mutation of (D2) but retains the information which constitutes Trait 1. If such retainment was clearly more frequent than changes in other features of design, the retained information could be described as being favored by selection despite the other changes.

Finally, trajectory (E) depicts another transformation from a trait-neutral form to another which embodies one of the example trait being utilized, just like (A) and (C). However, there is a difference: trajectory (E) stems from a location isolated from Area 1 by Area 2. Thus, the possibility of trajectory (E) reaching Area 1 without ever encroaching on Area 2 is neglible. This example demonstrates the role of causal trajectories: starting positions affect the probability of reaching differently positioned forms in Design Space (Dennett, 1995, p.125). However, selection enables even radical shifts. For example, the host of (E1) or any of its ancestors might have unproblematically replaced their held conception with an alternative which contained Trait 1 despite its remoteness. Were those

hosts to become exposed to distant alternatives, no principle prevents them from discarding their previous beliefs and adopting those conceptions. At most, the possibility that radically different representations contradict other held beliefs might be higher. This would lower the probability of the former being assimilated. Alternatives might also not be available because of an established doxa (Bourdieu, 1972/1979, pp.164–171; Deer, 2012).

Having discussed how this view of certain substructures of representations as the foundation for these representations' taxononomy answers Sperber's critiques (1) and (2), its other features are now examined. Firstly, this taxonomy diverges from the tracking of branching lineages in biology. Those views would only allow populations which resemble each other and share ancestors to be classified as a species (Dennett, 1995, pp.203–204). Thus, (C1) and (C2) could be seen as part of a common species but the likes of (B2), (D1), (D2) and (E1) could not be allowed into the same category. Dennett agrees that the lineage-crossing inherent to memes disallows the construction of species in the biological sense of common descent (Ibid. 355–356). His commentary on Hull's view that a common ancestry must be distinguished for similar tokens to be classifiable as members of the same meme (Hull, 1982) contains the claim that a sufficient amount of meme generations blurs the evidence necessary to confirm common descent (Dennett, 1995, pp.356–357). Only shared semantic structures might survive (Ibid.), yet they may also be produced by forced moves in Design Space (Ibid. pp.128–135) or free-floating rationales (Ibid. pp.132–133; Dennett, 2013, Chapter 39). Studying the differential propagation of selected structures instead of lineages seems plausible despite divorcing from the analogy with biology.

Secondly, this view enables formalizing the contents of memes in the lines of Russell's theory of descriptions (e.g. Russell, 1905). Basically, any representative token (m) may be described using its traits which are peculiar to the structures which correspond to them. For example, recycling Heylighen and Chielens's examples of the Christian concept of a god, (A) 'God is omnipotent', (B) 'God is good', and (C) 'God punishes bad people' (Heylighen and Chielens, 2009, p.8), a simple representation may be construed. This representation could be expressed as $(\exists!(x):G(x))\wedge\{\forall(x):G(x)\rightarrow[(A(x)\wedge B(x))\wedge C(x)]\}$. Here, $\exists!$ denotes unique existence, G is the attribute of 'being God', and A, B and C correspond with (A) omnipotence, (B) goodness and (C) 'being the punisher of bad people' respectively. While any actual conception probably includes more traits, only those being researched or seemingly contributing to these traits' fitness need be listed.

Of primary interest are not traits as such but their relations. Let B(m) denote a token which contains a structure which would implicate the goodness, B, of (m), the object being represented. All tokens contain multiple such descriptions, and object (m) is impossible to identify without referring to its other represented traits. Thus, a more polytheistic belief like $\exists(x):G(x)\wedge B(x)$, which depicts belief in one or more entities with the traits of G, 'being a god', and B, 'being good', is not prevented from being represented alongside the previous example. Both refer to the set of godly entities and attribute goodness to some of its elements. Such conceptions could then be contrasted with contradicting alternatives such as $\forall(x):G(x)\rightarrow E(x)$, where E, evilness, applies universally to godly entities. That alterbative contradicts any and all belief in benevolent deities since $\forall(x):[B(x)\rightarrow(\neg E(x))]\wedge[E(x)\rightarrow(\neg B(x))]$. Godliness, G, may in turn be defined as a family resemblance (Wittgenstein 1953/2001/2009b) which consists of the possession of certain traits such as (I) supernatural powers, (II) immortality and (III) worship-worthiness. The logic of such associations is detailed in chapter 4 of this study. The relative proportions in which these traits are present among the set of approved god-like entities is yet another problem which may be answered by utilizing theories of cultural selection.

In effect, what is being suggested is researching the distribution of specified traits inside a given set which is determined by its objects. Such traits are effectively alleles which share the same object (Distin, 2005, pp.169–170). Explaining their distribution may be based on theorizing about the differential fitness of specified traits within their cultural environments. Comparing existing beliefs with generalizations helps locate embodied traits of interest. Formal descriptions of the representations are also a helpful tool. Epistemic and doxastic logic (e.g. Grice, 2001; Hintikka, 1962/2005) and cognitive psychology (e.g. Cosmides and Tooby, 2000; Sperber and Hirschfeld, 2004) provide further assistance.

To summarize, it is possible to construct a selectionist account which accommodates and thus counters Sperber's critiques (1)–(3) regarding (1) the constant transformations of communicated content (Sperber, 1996, pp.102–103), (2) the problem of generalized accounts based on common content (Ibid. pp.29, 118), and (3) communication as the reconstruction of intended meaning founded on prior assumptions (Ibid. p.106; Sperber, 2000a). The theories which motivate Sperber's critiques (1) and (3) are accepted but these critiques are not perceived as valid rebuttals. In contrast, the truthfulness of critique (2) is denied despite individual tokens being accepted as also playing a crucial role.

Undoubtedly, the account provided here still retains some of the assumed problems of selectionism. It probably also raises new questions, such as the justification of abandoning lineage-centeredness in favor of generalizations which encompass multiple unrelated (in the sense of common descent) lineages (Hull, 1982). Criticisms which may apply to this account as forcefully as to orthodox meme theory include psychological shallowness (Plotkin, 2000; Sperber, 2006), replication-centeredness (Boyd and Richerson, 2000) and the need to further justify the conceptualization of cultural change as a selection of contents (Kuper, 2000). Despite this, the assumed position does diverge considerably from the introduced views. This will be elaborated in coming chapters, and maybe the introduced theory can be judged on its own merit in the end.

2.4 Chapter summary

Selectionism is defined by the concept of *universal Darwinism* – the idea that any process regardless of the entities involved follows the dynamics of natural selection if certain prerequisites are fulfilled (Dawkins, 1983; 1976/1989/2006b; Dennett, 1995; Blackmore, 1999, Chapter 2). While other variants have also been advanced (e.g. Von Neumann and Burk, 1966; Edelman, 2006, Chapter 3), the requirements discussed here are (1) variation, (2) replication, and (3) differential fitness (Dennett, 1991, p.200; 1995, p.343). This view focuses on the interaction and differential propagation of self-replicating entities.

An example theory of the development of culture which utilizes these dynamics of replication to explain the evolution of culture is the *meme theory*. That account proposes that researching culture should consist of examining which pieces of information are retained throughout the transmission of cultural contents. The claim which defines this approach is that cultural contents contain features which may assist their propagation regardless of the adaptive advantage conferred to their hosts (Dennett, 1991, pp.202–203; 1995, pp.361–365; 2009b, pp.299–302; 2013, pp.275–276; Blackmore, 1999, Chapters 3 and 9; 2009, pp.299–322;). While the concept of the meme and especially the way in which such entities are embodied are easily misunderstood, possibly the best definition is memes being replicable information packages which motivate producing interactors (Hull, 1980; 2001) capable of transmitting these replicators' contents. Different embodiments replicate this content with different reliability, and the fitness of a meme may be defined as

its average success during the necessary phases of the process of replication (Heylighen, 1999; Heylighen and Chielens, 2009, pp.10–13). A number of criteria affect the success rate during these stages of (1) assimilation, (2) retention, (3) expression, and (4) transmission (Ibid.). The particular factors which affect these processes include multiple levels of interaction, notably the relationship between a meme and its (potential) hosts (Heylighen, 1993; 1997; 1999; Heylighen and Chielens, 2009, pp.17–20).

The meme theory is far from unproblematic and has faced serious criticism since Dawkins invented the concept of a meme (Dawkins, 1976/1989/2006b, Chapter 11). While the above discussion regarding Dan Sperber's critiques (Sperber, 1996, pp.100–106; 2000a) is far from comprehensive, it is hopefully a step forward. Answering his claims that (1) culture almost never replicates (Sperber, 1996, pp.102–103), (2) discussing the general features of representations is mostly fruitless (Ibid. pp.29, 118), and (3) transmission of contents depends on prior capacities of the receiver (Ibid. p.106; Sperber 2000a) is of foremost importance because of the present aspiration to integrate his views and selectionism. The above discussion presents some details which should enable formulating such a combinatorial account. It was argued that (1) reconstruction of meaning is a means of replication and not an alternative to it (Distin, 2005, pp.107–108), (2) generalizations are indispensable because mere tracking of lineages is incompatible with the behavior of cultural information (Dennett, 1995, pp.355–356), and (3) these generalizations also allow the retainment of relevant information despite the transformation of other contents.

Thus, a preliminary framework has been established. While the next chapter will concentrate on introducing Sperber's theories such as the relevance theory of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) and the epidemiology of beliefs (Sperber, 1996; Claidière and Sperber, 2007), it will also further illuminate the emerging general picture of cultural evolution. Combining selectionism and the epidemiology of representations is far from impossible, and a number of arguments have been presented to support such an aspiration (e.g. Dennett, 1995, pp.357–360; 2006: pp.379–386; Driscoll, 2011; Henrich and Boyd, 2002; Sperber and Claidière, 2007).

3. EPIDEMIOLOGY OF REPRESENTATIONS

”*Just as one can say that a human population is inhabited by a much larger population of viruses, so one can say that it is inhabited by a much larger population of mental representations.*”

– Dan Sperber, *Explaining Culture* (Sperber, 1996, p.25)

”*The function of the linguistic meaning of an utterance is not to encode the speaker's meaning, but to provide* evidence *of her meaning.*”

– Preface of *Meaning and Relevance* (Wilson and Sperber, 2012, p.ix)

Because Dan Sperber perceives the relevance theory of communication pioneered by him and Deirdre Wilson (Sperber and Wilson, 1986/1995) as incompatible with the accurate replication selectionism is assumed to require (Sperber, 1996, pp.100–106), he is a long-standing critic of selectionist theories of cultural transmission, notably the meme theory (Ibid.; Sperber, 2000a, pp.163–173). According to the relevance theory, successful communication relies on the contextual relevance of available interpretations (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012). Sperber argues that this context-dependence and the subjectivity of interpretations are sufficient reasons to reject Darwinian models of cultural transmission (Sperber, 1996, pp.103–106). Since these critiques have already been answered in section 2.3, this chapter will instead focus on Sperber's positive contributions which help elaborate an integrative model for the evolution of culture.

Sperber is a vocal proponent of multidisciplinary approaches to culture (e.g. Sperber, 1996; 2000b; 2001; 2006). This commendable attitude is exemplified in his two rules of naturalistic causal explanation: (1) 'Don't recognise phenomena unless your grasp of their material mode of existence justifies your attributing them causal powers.' and (2) 'Don't make a causal claim unless you can back it with the description of a mechanism, a description fine-grained enough for it to be reasonable to ask neighbouring natural sciences to fill in the missing parts.' (Sperber, 2001, p.298) Both contain instructions required by the doctrine that 'Whatever has causal powers has them in virtue of its material properties.' (Ibid.) While sating Sperber's aspiration for multidisciplinary co-operation and materialist ideals is accepted as desirable, the aforementioned rules may at times be slightly rexaled.

This chapter will first introduce the basics of the relevance theory of communication to provide a foundation for understanding Sperber's alternative theory of cultural

evolution. Since his epidemiology of representations concerns the mental representations of encountered public representations (and vice versa) (Sperber, 1996; 2006), relevant forms of metarepresenting will also be explained. Finally, what is scrutinized last are the attractors (Sperber, 1996, pp.106–118) which determine Sperber's approach. The attractors were originally introduced to contrast Sperber's epidemiology with selectionist models concerning the dynamic of representations reaching ideal, stable forms. The incompatibility of the two approaches has been contested before (e.g. Henrich and Boyd, 2002; Dennett, 2006: pp.379–386; Driscoll, 2011) and a similar integrative motivation will guide the this chapter. Even Sperber himself has conceded that selectionism has an epidemiologic streak (Sperber, 2006, p.439) which is made explicit in the *virus of the mind* -metaphor (e.g. Brodie, 1996; Dawkins, 2003, Chapter 3.2).

3.1 Relevance theory of communication

The relevance theory of communication was originally conceived by Dan Sperber and Deirdre Wilson and introduced in their co-authored book *Relevance: Communication & Cognition* (Sperber and Wilson, 1986/1995). It has later been revised and supported with accrued empirical evidence in *Meaning and Relevance* (Wilson and Sperber, 2012). The relevance theory is an inferential account of communication and owes much to Gricean pragmatics (Grice, 1991)[1]. While Sperber and Wilson acknowledge that some human communication is encoded, they endorse the claims that (1) most acts of communication require more than is achievable by codes alone, and (2) codes are unnecessary for successful communication (Sperber and Wilson, 1986/1995, pp.24–28). Evidence supporting the latter has accrued because of research on constrained communicators (e.g. Goldin-Meadow, 2006; Goodwin, 2006; Liszkowski, 2006).

Grice's pragmatics are founded on his Cooperative Principle: 'Make your conversational contribution such as is required at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.' (Grice 1991, p.26) In addition to this overarching principle, Grice lists four maxims – the Maxims of (1) Quantity, (2) Quality, (3) Relation, and (4) Manner – which describe standard communication (Ibid. pp.26–27). Of these, the Maxim of Relation in particular seems to have

1 The referenced book, *Studies in the Way of Words* (Grice, 1991), contains Grice's William James lectures from 1967, in which he introduced his views on pragmatics.

inspired Sperber and Wilson, for it states: 'Be relevant.' (Grice, 1991, p.27) However, as Sperber and Wilson note, Grice actually leaves the concept of relevance unexplained (Sperber and Wilson, 1986/1995, p.36). Grice, on the other hand, claims that Sperber and Wilson rob that notion of its content (Grice, 1991, pp.371–372). However, their concept of relevance is hardly intended as a definition of the commonsensical meaning of that word. Rather, Sperber and Wilson's technical usage of the term *relevance* seems quite fertile. This usage includes evolutionary considerations: the efficiency acquired by maximizing the amount of inferred information using the least amount of resources, notably time and energy, would be a well-adapted strategy (Sperber and Wilson, 1986/1995, pp.261–263). Compared to Grice's model, the relevance theory also postulates less co-operation and reflection among communicators (Ibid. pp.161–162).

The core of the relevance theory of communication consists of these theses: (1) natural languages encode intended meaning in insufficient detail and expressions always contain ambiguity (Sperber and Wilson, 1986/1995, Chapter 1), (2) interpretations of expressions are always construed according to prior knowledge (Ibid. Chapter 2), and (3) the process of interpretation is reliant on the optimization of relevance, which is a variable of cognitive processing effort (Ibid. Chapter 3). The first two theses contain Sperber and Wilson's denial of code models while the third summarizes their alternative. In effect, their theory supports Grice's inferential account by applying cognitive (neuro)science. The relevance theory is supported empirically by controlled experiments (e.g. Jorgensen, Miller and Sperber, 1984; Happé, 1993; Wilson and Sperber, 2012, Chapters 13 and 14), which bolsters its plausibility as a realistic theory of human communication.

As was noted, the notion of relevance is re-defined by Sperber and Wilson in the context of the relevance theory. Their concept is bound to optimal efficiency in accruing new information according to inferences from the combination of new and old information available to a subject (Sperber and Wilson, 1986/1995, p.48). In other words, 'An assumption is only relevant in a context if and only if it has some contextual effect in that context.' (Ibid. p.122) By contextual effects, Sperber and Wilson refer to changes in the confirmation value (level of subjective certainty) of the assumptions being processed because of the context (Ibid. pp.108–111). Their definition is extended by two conditions which relate to optimality: the contextual effects achieved should be (maximally) large and the necessary processing effort should be (minimally) small (Ibid. p.125).

This dynamic is condensed in the form of the Cognitive Principle of Relevance[2]: 'Human cognition tends to be geared to the maximisation of relevance.' (Wilson and Sperber, 2004, p. 610; 2012, p.7) Furthermore, maximal relevance is accomplished only according to an ostension, such as remark's positive cognitive effects which are determined by its classificatory relevance as well as comparative relevance to the receiver (Sperber and Wilson, 1986/1995, pp.265–266). Classificatory relevance is bound to the interpretation's capacity to modify currently available, or *manifest* (see below), representations. In turn, comparative relevance is determined by the degree of optimality in the requisite cognitive processing of communicated information (Ibid.).

The Communicative Principle of Relevance contains the expectation of optimal relevance: 'Every act of overt communication conveys a presumption of its own optimal relevance.' (Wilson and Sperber, 2004, p.612; 2012, p.6) This presumption of optimal relevance can be further divided into two claims: (1) The ostensive stimulus acquired is sufficiently relevant for the receiver to process, and (2) it is maximally relevant among the possibilities available to the communicator due to his personal attributes (Sperber and Wilson, 1986/1995, p.270). Thus, when the Cognitive Principle and the Communicative Principle are combined, any act of overt communication is presumed to contain the most informative implications of significance (1) for the receiver (2) within that context. However, receivers must also consider the communicators' limitations.

A communicator's intended meaning is inferred according to such considerations of relevance. Of course, Sperber and Wilson do not claim that people calculate the relevance of communicative acts by consciously using the outlined principles. Rather, they claim that (successful) inferences are achieved by a central process (in the Fodorian sense) which receives its input from various input systems (Sperber and Wilson, 1986/1995, pp.71–72). The process of inferring transforms this raw input into conceptual representations with a logical form which may be propositional or non-propositional (Ibid. pp.72–73). Non-propositional conceptual representations archetypically contain free variables such as pronouns which leave the meaning open-ended (Ibid.). Thus, they are merely *incompletely* propositional. This degree of propositionality is closely tied to an acquired representation's degree of explicitness.

2 As it was called in *Meaning and Relevance* (Wilson and Sperber, 2012). It is also known as the *First Principle of Relevance* in the second edition's (Sperber and Wilson, 1986/1995) postface due to a previous misunderstanding regarding the two principles introduced in *Relevance* (Sperber and Wilson, 1986/1995). Likewise, the Communicative Principle of Relevance is also known as the *Second Principle of Relevance*.

The explicitness of a received assumption is defined as the measure to which its logical form corresponds with the logical, encoded form of the originating utterance (Sperber and Wilson, 1986/1995, p.182). It is a relationship between the amount of linguistically encoded information and the amount of conceptually inferred information. More precisely, the degree of explicitness is determined by the totality of encoded information which is inversely proportional to the totality of inferred information (Ibid.; Wilson and Sperber, 2012, pp.78–79). This establishes the relationship of the relevance theory and more traditional code theories. Sperber and Wilson invented the notion of an *explicature* (Sperber and Wilson, 1986/1995, p.182) to describe representations which contain any amount of explicitness in the sense defined above. The notion of explicature complements the notion of implicature (Grice, 1991, pp.24–40) which is standardly used in pragmatics.

According to Grice, an implicature indicates whatever contents a valid interpretation of a sentence has within a given context and independent of the literal, established meaning of the used words (Grice, 1991, pp.23–24). Intentional subversions of the maxims of communication are made because such deviations convey specific types of implicatures (Ibid. pp.32–37). Sperber and Wilson harness this definition according to which implicatures are context-bound and uncoded meanings inferrable from a message. As noted above, they also establish explicatures as the opposite of implicatures: an implicature is wholly uncoded while an explicature is coded to some degree (Wilson and Sperber, 2012, p.77).

The last concept from relevance theory to be introduced here is *manifestness* which yet again addresses cognitive science to help elaborate commonsensical views. In this case, the latter include problematic conjectures such as *mutual knowledge* and *shared information* (Sperber and Wilson, 1986/1995, p.38). Being manifest requires that a piece of information is (1) representable by an individual because it is perceptible or inferable and that (2) the individual would assume its truth or verisimilitude (Ibid. p.39). Notably, what is manifest is not necessarily yet embraced by anyone. Rather, manifest things are merely available (Ibid. pp.40–41). Since the whole of (contemporarily) manifest things forms what Sperber and Wilson call an individual's cognitive environment (Ibid. p.39), (1) overlap of (cognitive) environments and (2) consciousness of this overlap among participants (Ibid. pp.41–42) enable sharing potential inferences (*mutual manifestness*). Intersubjective common ground – jointly accepted presumptions (Enfield, 2006) – contributes greatly to this overlap between cognitive environments.

In order to summarize Sperber and Wilson's theory of communication, the introduced terminology is co-ordinated with the three theses presented earlier. The first thesis which relates to the necessary ambiguity of natural languages, is bound to the roles implicatures and explicatures play. The book *Relevance* is replete with examples with which Sperber and Wilson exemplify this idea (e.g. Sperber and Wilson, 1986/1995, pp.10–11, 188–190). In effect, any sentence's relationship with reality is indefinite because natural languages are incapable of sufficiently individualizing objects and circumstances. Some of the intended meaning must be inferred from the combination of the sentence and its context according to priorly acquired information. The concept of implicature designates pure inferences while the concept of explicature designates inferences which derive their logical form from some code. Together, they depict how natural languages achieve only partial ambiguity instead of collapsing into meaninglessness.[5]

The second thesis, about interpretations always being construed according to prior knowledge, is of foremost importance to Sperber's denial of selectionism (Sperber, 2000a). The concepts of manifestness and cognitive environment as well as the (second) presumption of optimal relevance the Communicative Principle of Relevance contains help understand it. Manifestness requires either being perceivable or inferable (Sperber and Wilson, 1986/1995, p.39), and being inferable implies that an agent possesses the required premises in the form of existing representations. No two people share a strictly identical perspective on their physical environment and no two people have identical cognitive abilities and memory traces (Ibid. p.41). Thus, no two cognitive environments, according to which inferences are drawn, are strictly identical. The second presumption of optimal relevance assumes maximal relevance within the boundaries of the communicators' attributes (Sperber and Wilson, 1986/1995, pp.270–272), including their cognitive environments. Due to this disparity of attributes, the relevance of what is communicated will never be equal between communicator and receiver. In effect, the intended interpretation and the actually made inferences may (and should) overlap but they are never strictly identical.[6]

5 Indeed, one may hypothesize how the transition from unambiguous codes perceived in animals (e.g. signal dances of bees) to natural languages in humans could have resulted from evolutionary advantage derived from the latter. After all, implicatures multiply the possible information content of expressions.

6 However, while the background for Sperber's critiques has been elaborated here, the answer that people may account for their differences during communication because they are capable of representing one another's conceptions should still apply. Likewise, Sperber establishes nothing but the fact that some details may change during communications because people interpret the same signals differently. If what is being researched are not the representations as such but certain features they contain, such transformations are tolerable because some features are still being replicated despite them.

According to Sperber and Wilson's final thesis, the process of interpretation consist of retrieving the most readily available and maximally consequental implications within some context (Sperber and Wilson, 1986/1995, pp.230–231, 266–278). Understanding this claim requires utilizing the eponymous concept of relevance, which is defined above as a variable of cognitive processing effort. Obviously, this claim also includes the assertion that every act of interpretation is guided by contextual considerations. Since Sperber and Wilson have already established the defunctionality of code models because expressions in natural languages are necessarily ambiguous, this is a natural consequence.

The process of inferring is best exemplified by an example adopted from *Meaning and Relevance* (Wilson and Sperber, 2012, pp.66–70). In this example, Lisa visits her neighbours, the Jones family, who invite her to join for supper. Lisa declines and explains how she has already eaten. While most people would readily understand her meaning, Wilson and Sperber observe that the used expression would literally mean that 'At some point in a time span whose endpoint is the time of utterance, Lisa has eaten something.' (Wilson and Sperber, 2012, p.66) Lisa's everyday words would actually not encode the intended meaning that she has eaten within a timespan which would make accepting the offer for supper inconvenient. Neither would they encode what she has eaten: a few peanuts, for example, would not suffice as a proper reason to decline joining the Joneses for supper (Ibid. p.67). However, Alan Jones, who voiced the offer, is supposedly well aware of the social conventions which govern supper as well as the average amount of food a normal human may intake within a limited span of time. As such, he would able to reconstruct the obvious meaning of Lisa's everyday response within the context of Lisa having declined an offer for supper. Sperber and Wilson provide a table elaborating this process of interpretation, and it also includes the weak implicature '(26g) Lisa might accept an invitation to supper another time.' (Ibid. 68) After all, she declined the offer due to temporary circumstances instead of communicating her commitment to a principle which would prohibit her from accepting the offer. If such a principle had existed, Lisa would in all likelyhood have preferred to communicate her commitment to it in order to save them both from the inconvenience of possibly having to repeat this discussion in the future.[7]

7 Unless, of course, she had a reason to conceal her commitment to the principle. For example, communicating the commitment to it might have also implicated her being involved in a socially questionable cult which requires such a strange vow from its members. If, however, Alan Jones was aware of her involvement in the cult, her attempt to conceal this would have informed him of Lisa's shame regarding her membership, despite her intent otherwise. Further imaginable circumstances always exist.

This example elaborates how Alan Jones's and Lisa's cognitive environments affect the implicatures they may infer from each other's everyday utterances. If an expectation derived from a previous representation is subverted somehow, the subversion will be interpreted according to the likeliest explanation available to one's cognitive resources. Since the present context is a common denominator for both the communicator and the receiver, it may and will be used as a reference for the search. Thus, relevance within a context is established as the principle which guides interpreting ambiguous utterances.

The relevance theory of communication describes the dynamics present during the transitions of representations. This introduction should enable understanding what motivates Sperber's problematization of selectionism (Sperber, 1996, pp.103–106; 2000a). However, the relevance theory will also be utilized in chapter 4 as an important part of the framework this study attempts to establish. Metarepresenting is discussed next since Sperber's epidemiology is inherently about how people tend to represent other representations instead of the objects in the world because the former have become much more numerous. After all, every mental infection consists of a public representation becoming mentally represented, and every act of transmission consists of publicly representing a mental representation. Objects in the world are only present when primary representations are formed, and the latter constitute a minuscule portion of people's mental contents.

3.2 Metarepresentations

In the introduction to *Metarepresentations: A multidisciplinary perspective* (Sperber, 2000b), Sperber lists four categories of secondary representation: (1) mental representations of mental representation, (2) mental representations of public representations, (3) public representations of mental representations, and (4) public representations of public representations (Ibid. 3). The distinction between these mental and public representations is addressed in *Explaining Culture* (Sperber, 1996): Mental representations include memories, beliefs and intentions, and their producer is the only person utilizing them. Public representations, on the other hand, exist in the environment of their users and their primary function is to enable communication between a producer and the interpreter (Ibid. p.32). Sperber's epidemiology of representations attempts to capture the transitions of such physically identifiable representations without invoking the existence of common essences

which he calls *cultural representations*[8] (Sperber, 1996, pp.61–66). Instead, his epidemiology focuses on the differing distributions of actual representations (Ibid. pp.75–76).

According to Sperber, metarepresentation gives rise to Cognitive Causal Chains (CCC), inclusing Social Cognitive Causal Chains (SCCC) and Cultural Cognitive Causal Chains (CCCC) (Sperber, 2001, pp.302–312; 2006, pp.434–443). Basic Cognitive Causal Chains also include links which are not strictly metarepresentative such as the changes in representations which processing sensory information causes (Sperber, 2006, p.434). Social Cognitive Causal Chains are strictly metarepresentative, however, since their links also include public representations and their interpretations (Ibid. pp.435–436). Such chains are formed when one individual's CCC-induced behavior becomes the sensory input of another, causing these Cognitive Causal Chains to connect (Ibid. p.435).

When a SCCC includes multiple instances of communication and its general content stabilizes, it becomes a Cultural Cognitive Causal Chain (Sperber, 2001, pp.309–312; 2006, pp.437–439). This boundary between the two types is blurry (Sperber, 2001, p.311) but the stability of the latter should ensure that participants correct each other (Schegloff, 2006) because the prevalent content is common ground (Enfield, 2006) they share. In effect, when participants have manifest expectations of its content because they have been exposed to its contents repeatedly, a SCCC becomes a CCCC. Similarly, if a CCCC is scarcely expressed, it may become eliminated from the pool of culturally shared representations (Sperber, 2001, p.312).

The metarepresentative nature of Sperber's epidemiologic model is made apparent since these chains exemplify his avoidance of invoking *abstract* collective representations. Strictly speaking, a public or mental secondary representation represents the representation from which it was derived instead of an object (Sperber, 1996, pp.33–34). Such a derived representation is an *interpretation* of its predecessor (Ibid. p.34). This view contrasts the position that new representations paraphrase the contents of a general type which corresponds to no actual representation in any physical medium outside its descriptions (Ibid. pp.35–41). Instead of common attributes only being shared because of a common object, Sperber argues that features become shared because of shared causal ancestries which tend to gravitate towards certain forms in Design Space (Ibid. pp.27–29, 49–55).

8 Indeed, Sperber makes it abundantly clear that he prefers *not* discussing representations on the abstract level of Platonic ideas (Sperber, 1996, pp.62–66). He accepts the possibility of studying the formal properties of abstract representations (cultural representations), but vastly prefers to focus on the causal chains of mental and public representations (Ibid. p.63).

Besides Sperber's taxonomy of representations, Dennett's intentional systems theory (Dennett, 1987; 2009a) will be presented to facilitate understanding the social role of metarepresentations. His theory includes the intentional stance which should allow elaborating the extent of manifestness. The psychology concerning the the *Theory-of-Mind* (ToM) will also be mentioned. David Premack and Guy Woodruff invented the concept to explain how chimpanzees may recognize the intentions of observed individuals (Premack and Woodfuff, 1978). It denotes the capacity of an agent to attribute mental states to itself and to others to explain their behavior patterns (Baron-Cohen, 1995, pp.51–53). Communication relies on the ToM, which is made apparent by autism being a form of *mindblindness* which results from a disruption in the Theory-of-Mind Mechanism (Baron-Cohen, 1995).

A further demarcation between forms of representation is Daniel Dennett's tentative division between florid and pastel representations (Dennett, 2000). This division relates to the question whether or not all representing and, consequently, metarepresenting is sufficiently uniform to be addressed as a singular phenomenon. As Dennett notes, within the volume in which his chapter lies (Sperber, 2000b), different people have differently endowed conceptions of metarepresenting despite claiming to be discussing the same phenomenon (Ibid. p.19). The introduced division between florid and pastel representations attempts to clarify the distinct functions of different metarepresentations.

Florid representing is deliberate and this purposiveness engenders self-consciousness (Dennett, 2000, p.18). The first impression might be that florid representations are necessarily metarepresentations because this self-consciousness involves representing yourself as representing (Ibid. p.19). After scrutinizing this impression and accepting it, Dennett contends that florid representing is *thinking about thinking*, as opposed to having *beliefs about beliefs*[9] (Ibid. pp.20–21). This prowess allegedly surfaces because by manipulating them, people internalize the symbols their environments contain (Ibid. pp.21–24).

In contrast, pastel representations arise from familiarizing oneself with the *objects* in one's environment (Dennett, 2000, p.23). They do not involve manipulating symbols (representations); only the manipulation of the represented objects. However, beliefs about beliefs are categorized as pastel, yet they involve a type of metarepresenting. The important difference is the fact that the pastel representations themselves are not being manipu-

9 To be distinguished from *belief in beliefs* introduced in *Breaking the Spell* (Dennett, 2006, pp.200–246). As Dennett notes, belief in beliefs is a result of reflection (Ibid. p.200), and the concept denotes a certain advocating as well as normative attitude towards *beliefs* as cultural, not mental, phenomena (Ibid. pp.200–203). In other words, belief in belief denotes voluntary fidelity to a meme.

lated (Ibid. pp.20–21). Dennett elaborates this point with a thought experiment which involves a mother dog and asks what it could possibly teach its puppies not to eat and how (Ibid. pp.25–26). Since dogs have no language capable of communicating concepts, the best the dam can do for its puppies is operantly conditioning them to avoid certain objects (orange toads, for example) which have become available in their environment (Ibid. p.25). Unlike linguistic communication, such conditioning of avoidance behavior may only establish distinctions by ostensively partitioning borderline cases as they arise.

Thus, manipulating symbols defines florid representing.[10] Communication involves such management of representations which makes public representations as well as the mental representations from which they are derived florid by definition. On the other hand, attributing intentionality, which is inherent in the Theory-of-Mind and the intentional stance, need not be done floridly. Pastel metarepresenting is involved when intentions are being inferred, and such inferences are necessary for successful communication according to the relevance theory. Hence, it may be tentatively suggested that pastel representations partially enable the appropriate manipulation of the conscious, florid metarepresentations.

The difference between florid representers and pastel representers may be illustrated using the way Dennett describes different manners of generating solutions by nominating four stages: (1) Darwinian creatures, (2) Skinnerian creatures, (3) Popperian creatures, and (4) Gregorian creatures (Dennett, 1995, pp.374–378; 1996, pp.81–93). These compose what Dennett nominates the *Tower of Generate-and-Test*, since they have a clear hierarchy (Dennett, 1995, p.373). Darwinian creatures utilize different hard-wired solutions and only those using the right solution multiply and flourish while the rest perish (Ibid. p.374). Skinnerian creatures react to new situation blindly but establish patterns by recycling effective strategies (Ibid. pp.374–375). Popperian creatures simulate the situation internally beforehand which enables selecting effective strategies more reliably without external conditioning (Ibid. pp.375–377). Finally, Gregorian creatures construct mind-tools such as words which improve their internal simulations cumulatively (Ibid. pp.377–381). Because most if not all vertebrates may be categorized as Popperian creatures (Ibid. p.376), the utilization of inner representations remains pastel while florid representing would also require those representations themselves to be available to conscious modification.

10 Dennett also explores this idea of *tools for thinking* in *Darwin's Dangerous Idea* (Dennett, 1995, pp.377–381) and *Kinds of Minds* (Dennett, 1996, pp.99–117) and showcases his own mental toolbox in *Intuition Pumps and Other Tools for Thinking* (Dennett, 2013).

The types of metarepresenting Sperber addresses resemble florid representing because both involve the content of the primary representations being re-represented (Sperber, 2000c, pp.117–120). The alternative is alluding to another representation without referencing its contents, as in Sperber's example (2a) 'Bill had a thought.' (Ibid. p.117) where Bill's thought is being represented but its contents are not. The two distinctions do not converge, however, since beliefs about beliefs do contain a rudimentary, non-propositional assessment of the contents of the primary beliefs. Such primal metarepresentations are illustrated by Sperber with his example (4a) 'He wants-to-mate.'[11] (Ibid. p.118) Sperber also echoes his omnipresent trifecta of representations by claiming that metarepresentational capabilities include the ability to re-represent using (1) mental representations, (2) public representations, and (3) abstract representations[12] (Ibid. pp.127–137).

The divisions suggested by Dennett and Sperber allow constructing a framework to classify metarepresentations. Firstly, there exists the division between representing the (1) existence or (2) content of other representations (Sperber, 2000c, pp.117–118). Since the difference between florid and pastel representations is manipulability (Dennett, 2000, pp.20–21), availability for manipulation constitutes another dimension. The third dimension would be mode of representation: whether a given metarepresentation is (1) mental, (2) public, or (3) abstract (Sperber, 2000c, pp.127–128). This framework will be used to help categorize different metarepresentational capabilities which are crucial for understanding the transmission of represented contents.

For example, communicating mental representations requires translating them into public representations, which is a form of manipulation. This qualifies the mental representations, the contents of which are being communicated, as florid. As argued, for instance by Luís Augusto (Augusto, 2010; 2013a; 2013b), even some traditionally higher-order mental representations may remain unconscious and, thus, beyond manipulation.[13] A fine example

11 This is Sperber's formulation, and *wants-to-mate* signifies a single, actually unarticulated attitude being represented (Sperber, 2000c, p.118).

12 Sperber alludes to anthropological generalizations in his discussion of cultural representations (Sperber, 1996, pp.70–74) while these abstract representations involve folk psychology. He suggests that the latter are processed by a *logical* module which attends to the logical content of encountered mental and public representations (Sperber, 2000c, pp.133–136). The cultural representations produced by anthropologists are such results but similar generalizations are made by everyone. Such folk essentialism (Medin, 1989; Griffiths, 2002; Mahalingam, 2007) will be addressed in section 4.3.

13 Augusto criticizes what he calls the *traditional model of human cognition*, or TMHC (Augusto 2013a). It includes the divide between conscious, symbolic representations and unconscious, non-symbolic (signal- or sign-based) representations (Ibid. 646–650). Pastel and florid representations might appear to follow the TMHC but Augusto's criticism is upheld because while florid representations are available for conscious recall and symbolic, it is not being argued that they exhaust symbolic forms of representing.

is the *halo effect* in which appearances unconsciously activate corresponding, stereotypic associations beyond what is visible (Augusto, 2013a, pp. 658–659; 2013b, p.22). Similarly, information about the receiver is manifest – available – but not necessarily being processed consciously during interaction. Generating public representations without deriving their information content from handled mental representations is also possible: one's posture might convey information publicly to those able to infer a person's mood from its details, and such communication need not be intended. In effect, the interpreter would explain why such behavior occurs by postulating a mood which generally causes such changes.

The introduced forms of pastel representing may be explained using Dennett's intentional systems theory, the central claims of which are that (1) people treat each other as intentional systems (Dennett, 2009a, pp.5–7), (2) there are multiple levels of intentional systems, the behavior of which may be predicted by attributing them layers of metarepresentations (Ibid. p.10), and (3) there exists no real demarcation between the intentionality of people and other intentional systems (Ibid. p.8). To treat something as an intentional system is to assume the intentional stance in order to explain its behavior (Ibid. p.1). Thus, a posture signals a mood because we attribute mental states such as moods to agent-like objects such as other humans to explain their outward actions.

The intentional stance is one of three perspectives concerning the predictability of one's surroundings (e.g. Dennett, 1987, pp.16–22; 2009a, pp.2–5). These three stances are (1) the physical stance, (2) the design stance, and (3) the intentional stance. The physical stance consists of deriving expectations from the constitution of some thing according to perceived physical regularities (Dennett, 1987, p.16). Ultimately it may reach applying the perspective of Laplace's demon[14] locally, like in Dennett's example of the Martian interpreter (Ibid. pp.25–28). The design stance, on the other hand, ignores this level of explanation in favor of predicting from the functional design of an object or organism (Ibid. pp.16–17). Instead of this design or appearance thereof needing to be intentionally produced, it may as well result from what Dennett calls *free-floating rationales*: locally optimal structures towards which selection will steer the designs which may realistically achieve them (Dennett, 1995, pp.132–133; 2013, Chapter 39). They closely resemble Sperber's

14 However distorted such a vision might be, as pointed out by Stephen Hawking (Hawking, 2011). Like Hawking paraphrases the original thought experiment, Laplace's demon is a theoretical entity which could calculate the position of all particles at any given time from information regarding their position and speed during one moment. He debunks Laplace's legacy which he calls the *hidden variable theory*, the view that there is a determined truth regarding the positions and speeds of particles despite our inability to determine them (Ibid.). The physical stance does not require this in-principle possibility of determinism.

attractors (Sperber, 1996, p.110), albeit Dennett's notion is more generally applicable. Finally, adopting the intentional stance to generate predictions means eschewing the two previous layers of explanation and instead attributing the thing rationality and mentally simulating the actions this would entail (Dennett, 1987, pp.17–22). Most naturally, the simulated level of rationality may be adjusted if the derived predictions fail to accommodate a system's observed behavior (Ibid. 21; Dennett, 2009a, p.6).

This listing of the stances generally sets from greater accuracy and expenditure towards increased generality. The physical stance, for example, is standardly only cost-effective when applied to extremely simple actions such as the falling of objects. The more general stances may overlook many details available to others but the processing costs associated with explaining more complex phenomena are much smaller. They simply describe most observed phenomena more cost-effectively. Thus, the intentional stance is generally the default stance human beings assume towards complex phenomena since natural selection favors such cost-efficiency (Dennett, 1987, p.18; 1995, pp.132, 237). Dennett calls this natural disposition to assume the intentional stance when predicting the actions of other people *folk psychology* (Dennett, 1987, pp.50–57).

Notably, the representations derived from the intentional stance, which represent the intentional states attributed to objects and organisms in one's environment, re-represent actual representations only contingently. Simulating the rationality of another requires attributing these representations to them.. In the case of vertebrates, personal experiences of being conscious provide a good reason for this attribution (Dennett, 1987, pp.53–54) but the intentional stance may even be applied to objects such as toasters. This demand for consistency also motivate Dennett's project of heterophenomenology (studying the phenomenology of others): there exist no good reasons to surmise that similar structures would not support the emergence of similar experiences (Dennett, 1991, pp.404–406). This idea resembles Sperber's materialist doctrine of causal power always relying on material structures (Sperber, 2001, p.298). Similar claims about the necessary consistency and continuity of nature are also made by Popper (1959/2002b, pp.249–251) and Deutsch (1997, Chapter 7): natural laws allow no inexplicable exceptions.

According to the relevance theory, successful communication relies on mutually manifest intentions (Sperber, 1996, pp.43–44), and this requires them being accessible to all involved parties (Ibid. pp.41–42). Despite their fallibility in representing anything real,

the intentions derived from the intentional stance are with what humans operate (Baron-Cohen, 1995, Chapters 3 and 4). Unless brought under scrutiny forcefully because expectations become subverted for example (Grice, 1991, pp.30–37), such representations should remain pastel. Neither of the associated theories requires people to consciously reflect on their contents. Like mutual manifestness is a requirement for successfully inferring intended meanings (Sperber and Wilson, 1986/1995, pp.163–170), the representations of others' intentions the intentional stance enables may simply constitute a pastel background required for deploying more vivid metarepresentations.

This necessary, relatively reliable prediction of the actions of higher order intentional systems requires multiple layers of metarepresenting – the re-representation of metarepresentations (Dennett, 1987, pp.242–250; 2009a, p.10). Such predictability is a prerequisite for deriving appropriate expectations regarding the systems' intentions. Effectively, the user of the intentional stance must be a higher order intentional system than is required to generate the output being utilized in order to derive another's intentions. The reason is how the (meta)representations one encounters need to actually be re-represented.

This dynamic is explored by Sperber in his example of Mary and Peter (Sperber, 2000c, pp.124–125). He aspires to demonstrate that metarepresentational capacities precede language, or at least that communication is enabled primarily by metarepresenting rather than the use of language (Sperber, 2000c, pp.121–127). In Sperber's example, Peter examines Mary picking berries in five different scenarios. As the scenarios advance, the amount of layers of metarepresenting increase as Mary attempts to communicate with Peter through her actions and Peter attempts to derive Mary's intentions from her actions. Finally, in the last scenario, '[1] Mary intends / [2] that Peter should believe / [3] that she intends / [4] that he should believe / [5] that these berries are edible', (Ibid. p.125) and (6) Peter must represent this intention of hers. In this scenario, Mary's communicative intent requires the capacities of a fifth-order intentional system and Peter must be a sixth-order intentional system to comprehend her intent reliably.

According to Sperber, upon reaching this level of deriving sixth-order representations from fifth-order communicative intents, people shift from being required to use ostensive behavior to being able to use symbolic behavior (Sperber, 2000c. pp.125–127). In effect, they become capable of florid rather than merely pastel representing. Sperber argues that symbolic behavior only becomes meaningful when interpreted as an intention

to influence an audience's representations (Ibid. p.126). Since symbolic behavior operates on symbols, it establishes the possibility of a common code between participants (Ibid.). Thus, utilizing languages in communication actually requires this order of metarepresenting because they are such systems of codes. Nevertheless, the real genealogy of languages might actually predate using them for communication (Ibid. p.127).

In summary, understanding Sperber's theoretical framework for the transmission of representations is facilitated by three factors: (1) the representation of content or mere existence (Sperber, 2000c, pp.117–120), (2) the division to pastel and florid representations according to manipulatibility and accessibility (Dennett, 2000), and (3) the division to (I) mental, (II) public and (III) abstract representations according to platform and the module responsible (Sperber, 2000c, pp.127–136). The first dimension separates uninteresting cases of metarepresenting from those Sperber discusses (Ibid. pp.117–118). Pastel and florid representations coexist in human cognition, and their relationship may be described as pastel representations forming a foundation for the florid representations people may actually express. Such important pastel representations include the metarepresentations about the representations attributed to others because of the intentional stance. Finally, humans can represent the same information content using different media. Indeed, the epidemiology of representations is based on trailing the causal chains of one type of representation being re-interpreted as another type (Sperber, 1996, pp.56–76; 2001; 2006).

3.3 Attractors

Sperber proposes attraction as a viable alternative to the selectionist, Darwinian models of cultural inheritance such as the meme theory and dual-inheritance theories, or DIT (Sperber, 1996, pp.98–118). Attraction denotes probability-based differences between the vectors which affect the chains of inheritance connecting representations when these biases are skewed towards *attractors* by psychological and ecological factors (Ibid. pp.106–118). These attractors are ideal forms of representations, fit for the totality of a person's cognitive and socially acquired traits (Ibid. pp.111–114). Sperber emphasizes both psychological and ecological factors: the first include individual psychology, as mentioned above, and species-specific capacities (Ibid. pp.113–115), while the latter consist of interaction between individuals and their *cultural environments* (Ibid. p.115).

Four time-scales at which psychological and ecological factors interact are distinguished by Sperber: (1) biological evolution, (2) socio-cultural history, (3) individual cognitive development, and (4) processes of transmission (Sperber, 1996, p.113). The levels of psychological factors are thus bound to the time required to develop them. For example, on the time-scale of biological evolution, specialized mental modules are produced (Ibid. pp.113–114, 123–129). Socio-cultural history, on the other hand, includes contents which individuals encounter during their cognitive development. Such ecological factors constitute the cultural environment of an individual. The aforementioned cognitive development consists of encountering some existing cultural contents and selectively assimilating some of those representations (Ibid. p.115). This process of assimilation is the final level, but it is affected by what has been made available by the other processes.

For example, biological evolution generates modular processes which are cognitively efficient as well as adaptively specialized (Sperber, 1996, p.114). The module(s) responsible for successful communication the relevance theory postulates are included. Sperber's view on the basis of metarepresentative capacities (Sperber, 2000c) and the spilling of these modules' functions to domains beyond their evolutionarily intended targets (Sperber and Hirschfeld, 2004) exemplify the importance of such modularity.

Besides the socio-historical formation of cultural environments, ecological factors include the long-term changes in those environments. Sperber lists the following examples of factors which affect the formation and annihilation of attractors: 'over-exploited ecological niches lose their economic attraction; rarely walked paths become overgrown; some practices tend to increase, and others to decrease, the size of the populations that might be attracted to them, and so on.' (Sperber, 1996, p.115) Their relevant similarities with the processes of social distinction (Bourdieu, 1979/1984/2010) will be explicated later. The interaction of psychological and ecological factors is the primary catalyst of change in both elements. The environment affects which representations are available for assimilation. However, individuals host these contents, and whether or not they become expressed publicly affects these representations' relative density in the cultural environment to which others are exposed (Sperber, 1996, pp.115–116).

The differences between these factors of attraction and the forces of selection may be clarified using Sperber and Claidière's example of developing smoking habits (Claidière and Sperber, 2007) which was used to rebut Henrich and Boyd's claim about the effects of

selection cumulating despite gravitation towards attractors (Henrich and Boyd, 2002). The demonstration contains 31 alternative smoking habits which represent the average amount of cigarettes smoked daily, varying from 0 to 30 (Claidière and Sperber, 2007, pp.92–93). The authors illustrate selection forces by arbitrarily assigning people who smoke 10 cigarettes a day as the likeliest models for the people adopting this habit (Ibid. p.93). The behavior of these people is thus being copied most. However, the attractors affecting the habit's evolution are situated at 0 and 25, due to (1) tobacco's addictiveness and (2) people's reactions (abandonment or dependence) on the former (Ibid. pp.93–94). Thus, while people might replicate the habit of smoking by mimicking certain individuals' behavior because those individuals have qualities which encourage copying them, latter practices will be further affected by the factors of attraction.

A common misconception is that movement in Design Space must necessarily be directed towards the closest attractor (Claidière and Sperber, 2007, pp.99–100). However, attractors do not exert such a deterministic influence. Rather, their influence is probability-based (Ibid. pp.99, 105–107; Sperber, 1996, pp.110–112). Consequently, after any number of generations of reconstruction, attractors will have biased the proportions and distributions of related representations but all possible forms will be varyingly present after an arbitrarily high number of iterations (generations) (Claidière and Sperber, 2007, pp.95, 97). Also, tobacco's addictiveness in their example would exert different influences if (1) tobacco were never smoked or (2) people were unaware of its addictiveness. In scenario (1), addiction would not ensure the public displays of tobacco-smoking behavior, nor increase the average amount of cigarettes a day. Likewise, in scenario (2), the negative connotation of being addicted to a substance would not cause people to abstain from smoking tobacco. Also, (3) were there no such negative connotations concerning the loss of health, free will and similarly valued traits, a result similar to the one caused by the lack of knowledge in scenario (2) occur, since abstinence would not be an attractor. These example scenarios demonstrate how attractors depend on existing cultural models.

Misunderstandings of the sort depicted above between differently oriented theorists of cultural evolution might be explained by differences in the phenomena selectionists and epidemiologists emphasize (Driscoll, 2011). Selectionists, such as DIT-theorists,[15] are

15 Since the meme theory argues for the selection of selfish content, the adaptive advantage derived by human beings (Driscoll 2011: 302–304) seems secondary. However, such advantage includes how conforming to a society's conventions might bolster one's status, which increases the desirability of retaining and expressing the corresponding memes. Thus, Driscoll's analysis (Driscoll, 2011) should apply.

generally interested in cumulative changes and the cultural complexity they generate (Ibid. pp.302–304). Their models also often depict population-wide trends of transmission (Ibid. p.302). In contrast, epidemiologists, including Sperber, stress relative stability despite the unreliability of communication. They also focus on instances of transmission between individuals and the way in which they adopt habits and representations (Ibid. pp.302, 305–306). While the inclusion of both cumulative change and relative stability is necessary to explain the evolution of culture, the representatives of these two parties have tended to clash. For example, the inclusion of attractors is often understood as stunting the possibility of any further change after the attractors have been reached (Ibid. pp.306–307).

These two perspectives' focus on different examples is noteworthy because selectionist (especially DIT) theorists utilize adaptively meaningful cultural phenomena such as tools and techniques while epidemiologists prefer adaptively (mostly) neutral examples such as stories (Driscoll, 2011, p.308). Driscoll argues that this difference is reflected on the theorists' accounts of what constitutes a meaningful breach of fidelity (Ibid. pp.307–311). Meaningful differences between transmitted adaptively significant items are those which affect their contribution to adaptiveness under any, and especially under the prevailing, circumstances (Ibid. p.308). In contrast, any and all changes in the content of a story, song or similar cultural item will somewhat shift its identity. As long as the functionality of cultural items of the first sort is preserved, the changes occurring during their transmission are practically meaningless despite being bound to occur (Ibid.).

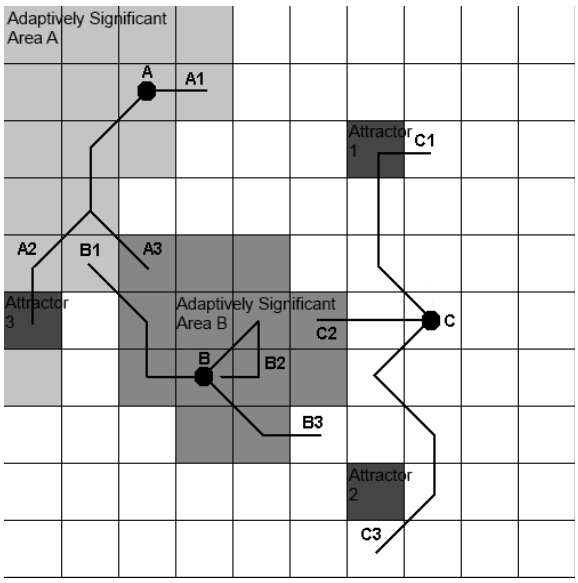
In all, Driscoll lists five reasons why attractors do not prevent cumulative cultural evolution (Driscoll, 2011, pp.307–320). Reason (1) is the aforementioned retainment of adaptively significant traits (Ibid. pp.307–311). Reason (2) maintains that differential adaptive significance suffices to explain why the effects of attractors might become inhibited or suppressed (Ibid. pp.311–313). Reason (3) implies the effects of a multitude of attractors. The more densely packed Design Space is with attractors, the less movement will be directed away from adaptively significant forms (Ibid. pp.313–314). According to reason (4), attractors will likely influence certain components of transferred representations but leave the functionality of their products more or less intact (Ibid. pp.314–316).[16] Finally,

16 New species of representations may also result from the amalgamating of representations brought together by attraction (Driscoll, 2011, p.315; Gabora, 2011, pp.72–73). Each part of such an amalgam representation would be easily retained since it fits psychological demands. Also, such representations could overlap only partially since no story, for example, consists of adding more of the exact same content. Any given attractor could only affect a limited number of components of the constituting representations as well if Driscoll's hypothesis is correct.

reason (5) is attractors not being universal even within local populations (Ibid. pp.317–320). Driscoll notes how Sperber includes the features of individuals among factors of attraction and affirms how they change during a person's life cycle (Sperber, 1996, pp.115–117). Additionally, specialized sub-populations may retain even cultural traits which are utterly redundant outside these populations' fields of specialization (Driscoll, 2011, pp.317–318). Parallels to Bourdieu are invited since the value of (social) capitals is field-dependent (Bourdieu and Wacquant, 1992, pp.126–128; Maton, 2012, pp.50–59; Thomson, 2012, pp.67–73). Driscoll's arguments also incorporate the continual interferences and selective emphases of originating hosts, the so-called *teachers* (Driscoll, 2011, pp.310–312, 314–315, 319–320). This mentor dynamic is exemplified by how communicators correct their interpreters during everyday communication (Schegloff, 2006).

Driscoll's insights may be demonstrated by utilizing Sperber's grid (Sperber, 1996, pp.109, 112). While Sperber leaves all the dimensions a representation's placement within the grid affects unexplicated, it should probably be assumed that cells which are closer to each other represent more similar iterations since understanding the effects of attractors requires this. In this regard, this illustration resembles figure 2.2 although figure 3.1 does not denote traits in any way. Although one axis could represent the adaptive advantage which would be made available to a representation's hosts, this demonstration will instead contain designated areas of equal adaptive advantage. These areas are independent of attractors because such a model will illustrate Driscoll's claims (1)–(3) regarding movement in Design Space within an unchanging range of adaptive advantage (Driscoll, 2011, pp.307–316) better than the alternatives. Contribution to adaptedness is a derived trait and a two-dimensional grid could never represent all the relevant factors.

In figure 3.1, there are nine example trajectories which are affected by three attractors but limited by relative adaptive advantage between forms. They illustrate different causal pathways, each traversable but not equiprobable within the present theoretical framework. While the actual probabilities of different trajectories are excluded, they are assigned relative probabilities according to predictions derived from Sperber's (1996; Claidière and Sperber, 2007) and Driscoll's (2011) accounts. The shown trajectories are also extremely simplified. They should actually branch in almost every generation, every transition between cells, since outside repeated communication between only two people the interpretations made would likely not remain this uniform.



Explanations:

1. Emphasized areas:
a) Adaptively Significant Areas A and B
- Contain all equally adaptively beneficial iterations
b) Attractors 1–3

2. Drawings:
a) Dots A, B, and C
- Teachers in Driscoll's sense (Driscoll 2011)
b) Trajectories A1, A2, A3, B1, B2, B3, C1, C2 and C3
- Examples of causal chains of reprenting

FIGURE 3.1. Demonstration of Driscoll's arguments (Driscoll, 2011) by applying it to Sperber's grid (Sperber, 1996, pp.109, 112).

The teachers symbolized by dots A, B, and C are individuals who host the representations which correspond to their coordinates. Teachers A and B hold adaptively significant representations while Teacher C's representation is approximately neutral in this regard. These individuals' cultural traits have been chosen for imitation, and for the sake of argument, Teacher B's representation is a selection peak. The reason might be prestige-based bias (Henrich and Boyd, 2002, pp.102–108; Henrich and McElreath, 2003, pp.129–130) or exposure to that representation being particularly likely (Boyd and Richerson, 1998, Chapter 7). Thus, if Teacher B is among the models which are available for imitation, the corresponding representation is particularly likely to be replicated.

The adaptively significant areas dubbed A and B denote different representations which have equal effects on the adaptive advantage of the individuals who host them. The adaptive significance of representations in Area B is greater than of those in Area A. Driscoll's claims (1) and (2) are directly related to the properties of these locations: according to her, adaptively significant representations are unlikely to transform into adaptively less significant forms during transition (Driscoll, 2011, pp.307–311), even under the effects of factors of attraction (Ibid. pp.311–313).

The provided nine trajectories may be analyzed in the light of these circumstances. (A1) represents movement within an adaptively significant area towards Attractor

1 which lies outside it but this trajectory does not cross the boundary which separates adaptively significant representations from the adaptively neutral. It illustrates how an attractor should be unlikely to force a representation to abandon what confers it adaptive significance since the latter also weighs in the scales (Driscoll, 2011, pp.307–313). On the other hand, the second trajectory which originates from Teacher A is attracted by Attractor 3 which lies within the boundaries of Area A. It branches to (A2) and (A3) which illustrate two distinct alternatives. (A2) approaches Attractor 3 uninhibited since its adaptive advantage remains unchanged during the required transitions. Those changes between generations are adaptively neutral between the forms. If the ratio of attractors is suitably high, such uncontroversial mode of movement would become the norm since each adaptively significant area would likely contain attractors. Exactly this is asserted by Driscoll's reason (3) (Ibid. pp.313–314). The branch of (A3), on the other hand, diverges by being transformed into a form with greater adaptive advantage than representations within Area A offer. If Driscoll is right, its future generations would likely resemble (A1) by remaining inside the range of possibilities with the greatest adaptive advantage.[17]

Of the trajectories which originate from Teacher B, whose example the forces of selection favor, (B1) crosses to adaptively less significant territory, (B2) drifts around Area B before returning, and (B3) is attracted beyond adaptively advantageous regions by Attractor 2. Thus, (B1) mirrors (A3), but by becoming less advantageous instead. The existence of Attractor 3 in that general direction might provide an excuse. Unlike (B3), (B1) does however retain some adaptive significance. The shortest path towards Attractor 3 would have strayed to adaptively neutral territory. As long as adaptive significance's effects exceed those of attraction, (B1)'s trajectory would be more probable than (B3)'s since it retains more adaptive significance. However, since (B2) retains all its adaptive edge throughout its transformations, it is the likeliest of the three unless not tracking any attractor affects its probability more than retaining adaptive significance does.

17 How much would a trajectory's history affect the probability of different transformations? For instance, someone already acquainted with the possibilities of Area A might be able to compare them with the advantages of Area B more effectively, and thus reduce the chance of backtracing. However, someone acquainted with the advantages of Area A but unable to compare them with Area B's could prefer returning to Area A. These musings only apply to trajectories of representations contained within one individual's development, and to cases of the learner being able to consciously evaluate the encountered representations and relate them to alternatives. In both cases, the learner also retains a token of the less adaptively significant representation but as a metarepresentation of a representation attributed to another – such as their bygone self. Distin's discussions of active, passive, and recessive memes (Distin, 2005, pp.44–45, 58–59, 89–91) and propositional attitudes (Ibid. pp.169–170) clarify this dilemma of possessing many alternative or even contradictory representations.

The fictional trajectory of (B2) which returns to the selective peak represented by Teacher 2 introduces yet another question. Since the representation held by Teacher 2 is the most commonly adopted form within the imaginary community *regardless* of the types of biases which ensure this, the conformist bias should apply to it and any other selection peaks (Henrich and Boyd, 2002, pp.101–102). Favoring the most prevalent contents often causes stagnation and inertia (Ibid. p.101) – in other words, stability. Like Driscoll observes, the explanation of stability despite high rates of transformation is the primary motivation behind attractor-based accounts (Driscoll, 2011, pp.302, 305–306). Since the conformist bias is constituted by a psychological disposition to favor the status quo (Henrich and Boyd, 1998; Henrich and McElreath, 2003, pp.130–133) and the prevalence of a given form is an ecological factor par excellence (Sperber, 1996, pp.115–116), conformism closely resembles a factor of attraction. Were this the case, the cell which contains Teacher B should actually be marked as Attractor 4, and its influence would require being considered in the cases of trajectories (A3), (B1), (B3), and (C2). The wider implication of this conjecture is, of course, that differentiating between forces of selection and factors of attraction may be unfounded.

Finally, the three trajectories which originate from Teacher C begin outside adaptively significant areas but relatively close to two of the conjectured attractors. Their primary purpose is to illustrate the indeterminate nature of attraction in Sperber's model (Sperber, 1996, pp.110–112; Claidière and Sperber, 2007, p.99). (C1) is initially pulled towards Attractor 1 but after reaching this attractor, this trajectory deviates from it. It closely resembles trajectory (B) in Sperber's original illustration (Sperber, 1996, p.112). The next transformation undertaken by (C1) would be greatly biased towards reverting to Attractor 1 or maintaining the same distance. However, further deviations remain *possible*. (C3) is similar to (C1) since both illustrate the probabilistic, indeterministic nature of attractors. Unlike (C1), (C3) never reaches the attractor. Instead, it gravitates around it, and retains a fixed distance to Attractor 2. Such trajectories may be less probable than the likes of (C1) but they remain more likely than any trajectories which escape an attractor's pull.

The last trajectory, (C2) is exceptional but within the realm of possibility. Instead of being pulled towards either of the attractors in the vicinity of Teacher C, it crosses the boundary and reaches adaptively significant forms, much like (A3). The earlier problem returns: would the nearby attractors' influence suffice to cause (C2) to return to a less

adaptively significant form? If trajectories like (B1) and (B3) are possible, so is the recession of (C2). However, if Driscoll is correct in her proposal (Driscoll, 2011, pp.307–313), this further development should remain highly improbable. Also, if the form represented by Teacher B is indeed an attractor besides being a selection peak, (C2)'s path should continue in its general direction across Design Space.

Thus far, all transformations undertaken by the cognitive causal chains have been supposed to represent minimal steps in Design Space, because this reflects the original examples in *Explaining Culture* (Sperber, 1996). However, as long as the axes of the grid represent some dimensions, along which changes happen, the supposition that only minimal changes at a time are possible may be unjustified. Postulating a race across the separating cells is unjustified because people rarely change their mind by processing all differences between two alternatives.

More substantially, the evolution of science as described by Sir Karl Popper (e.g. 1959/2002b; 1963/2002a; 1972/1979) and David Deutsch (1997; 2011) would allow bold conjectures to leap across Design Space. Once a held theory is falsified by critical tests and thus eliminated, conjecturing a new theory is an exercise in creative imagination (Popper, 1959/2002b, pp.8–9, Chapter 4; Deutsch, 2011, p.26). While elimination does occur in this scenario, it differs from the competition with which selectionism is concerned because no alternative representation supersedes the original as a part of this process. While science remains Popper's primary subject, he does attempt to generalize this account by arguing that solving problems via conjecturing and refuting tentative theories is how cognition works (Popper, 1972/1979, pp.242–244). Deutsch's elaboration of the *theory-ladedness*, or *theory-impregnatedness*, of experiences (Deutsch, 2011, p.10) which is also shared by Popper (1972/1979, pp.30–31), reveals similar aspirations.

However, a new theory should remain able to assimilate the same phenomena as the previous theory and it should not be falsified by the same critical test which prompted its creation (Popper, 1959/2002b, pp.276–281). The jumps across Design Space being investigated are thus not random but limited by this requirement to at least apply to known phenomena. As such, Sperber's model of causality is not abolished but must simply be expanded. This change is easily elaborated by contrasting a version of Sperber's illustration of the direction of change with a model which would permit the jumps proposed by Popper and Deutsch. Of course, in any particular case, the made modifications may also be

minimal and only address the encountered difficulties by adding singular auxiliary hypotheses (Popper, 1959/2002b, pp.62–63). Nevertheless, even such small changes which would also be compatible with Sperber's model are easily accommodated by the more general model as well. As the illustration in figure 3.2 should make abundantly clear, the transformations the latter may describe include all those Sperber would allow as well.

In the two models pictured in figure 3.2, (A), which is located in cell (3.3), denotes the mutating representation. All cells being shown have been numbered according to their row and column. According to Sperber, the descendants of (A) would be limited to either retaining its content or assuming one of the neighboring forms which are represented by cells (2.2), (2.3), (2.4), (3.2), (3.4), (4.2), (4.3) and (4.4) (Sperber 1996: 111). In turn, any representations which correspond to those cells may only beget descendants which correspond to their adjacent cells. Thus, all the hypothetical reaches of Design Space are potentially traversable though most parts will never become actualized. Not only are many parts incompatible with human psychology, the amount of these possibilities is Vast (Dennett, 1995, pp.118–120). Nevertheless, according to Sperber, any given point in Design Space may only be reached by an undisturbed trajectory of minimal movement. All the potential intra-individual leaps would result from some superseding alternative replacing the held representation. However, such alternatives would in turn have been generated by gradual transformations during mental and public transitions.

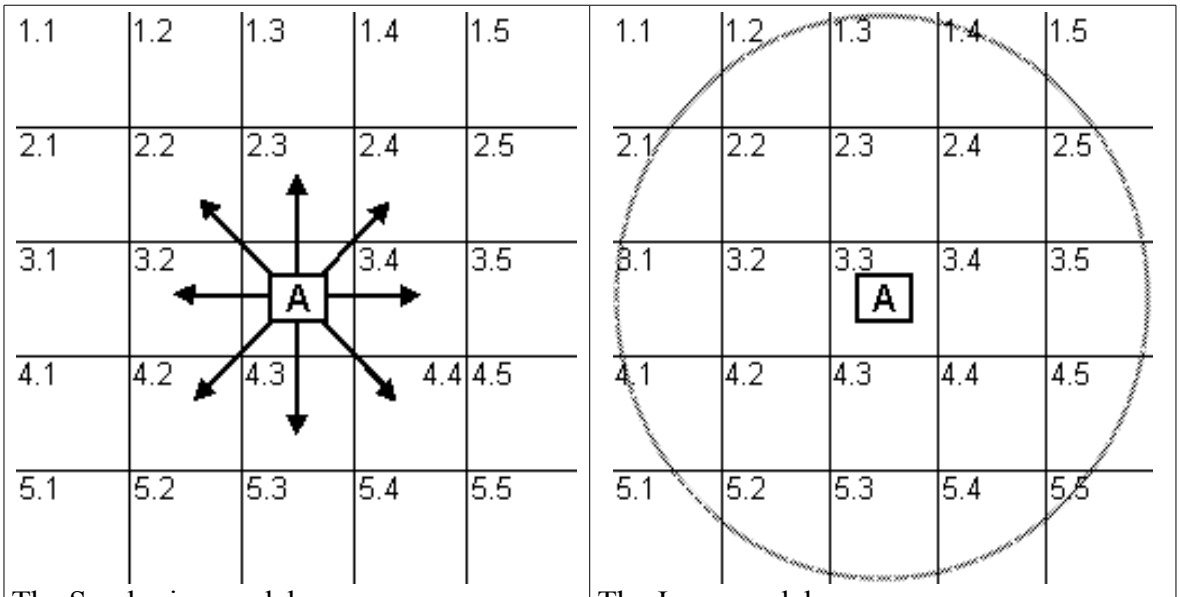


FIGURE 3.2. A comparison of Sperber's model of representational descent (Sperber, 1996, p.111) and a model integrating the possibility of bigger transformations, or *jumps*.

In contrast, the other model, nominated here *the Jump model*, would allow discontinuous causal chains of representing. In the illustration figure 3.2 contains, this is represented by an indefinite circle around (A) instead of arrows denoting the possible forms. Since the descendants of (A) are still limited by the merits of (A), their possible forms are likewise bounded. In this model, any cell even partially within the circle contains a representation of a potential descendant's form. For example, a causal trajectory might jump to (3.1) and to (1.3) from there instead of spanning across intermediate cells. While a circle was chosen for illustration purposes, successfully modeling the range of possible descendants a representation may beget is not limited to any particular shape.

Due to factors of attraction, the probabilities to reconstrue the particular forms different cells represent differ. In Sperber's model, there are always exactly ten possible moves in a two-dimensional model of the Design Space: into one of the nine adjacent cells or stasis. Probability may be apportioned between these possibilities according to psychological and ecological factors, which total the certainty of one of the ten occurring. Thus, $P=1=P_1+P_2+P_3+...+P_{10}$; $0\leq P_x\leq 1$, where P denotes the total probability value of all possible moves and its subindexed proportions match the distribution of the ten available forms' probabilities. P_x denotes any given possibility. Utilizing a similar formula to capture the moves available during any given generation in the Jump model is possible despite the higher and indefinite amount of possible forms. It would have the more general form $P=1=P_1+P_2+P_3+...+P_n$; $0<P_x\leq 1$, where (n) denotes the number of available forms while the rest of the notations are preserved. Unlike in Sperber's model, no possibility value may reach zero since the corresponding form would be eliminated from the set because no other principle than having a positive possibility value delimits its elements.

Also, while the average possibility of any given form fluctuates according to the amount of available possibilities, the distribution of probability may contain comparable chances between models. For example, in Sperber's model, the possibility of (2.2) becoming actualized during the generation following (A) may have any value between one and zero. Because the amount of in-principle possible forms is fixed, (2.2)'s probability may have any value, including certainty and impossibility. In contrast, any of the possible forms shown in the Jump model's presentation, such as (3.5), may have any probability value other than zero. Were it extremely high such as $P_{(3.5)}=(0,999...)$, the remaining probability value could still be distributed between the remaining possible forms. Thus, an

equally big value may occur in either model at the expense of the rest of the possibilities. Should this value reach one for any given move in either model, the rest become eliminated from the set of possible moves.[18]

The Jump model can assimilate Sperber's model, and since Sperber offers no arguments to support the limitation to minimum movement, this assimiliation becomes desirable because it allows larger, actually possible changes. An example using Sperber's own example of smoking habits (Claidière and Sperber, 2007) should suffice to justify this transition. Let Kim be a smoker who habitually smokes five cigarettes a day. Under normal circumstances, the addictive qualities of tobacco would tend to increase this amount gradually. Were Kim to recognize this developing addiction, the habit might be dropped completely.[19] Such a change might remain within the ten allowed possibilities but circumstances such as depression caused by a sudden trauma might instantly tranform this average tobacco consumption into any of the amounts specified by Claidière and Sperber (Ibid.). Since this amount of possible moves surpasses ten, it would be indigestible within Sperber's framework. It might be argued that the actual amount of possible moves under any circumstances does not exceed ten but such a claim would require some justification. The example used only has one dimension of change, the amount of cigarettes smoked daily. Most examples would include more dimensions, and a two-dimensional model may only depict two. Change along other axes is still possible, including three or more axes at a time, and the possibilities differring along a third axis must be distributed according to the two criteria being modeled. Such incontinuity of similar forms is simply accommodated better by the Jump model.

Assimilation is made possible by the fact that a given representation's reach may coincide with Sperber's model. The malleability of the zone of possible descendants in the Jump model yields this result. In effect, any given element, P_y, of the whole set of all related designs has a probability value between one and zero, $0 \leq P_y \leq 1$, for becoming represented when a token of the representation which determines that set transforms. The

18 Such situations may be called *forced moves*: the transformation necessarily occurs in a dictated way (Dennett, 1995, pp.128–135). While such dictation might be limited to one trait (representations differing in this respect having a probability value of zero), the situation discussed here is absolute because there would only exist one possible descendant and it has a probability value of one. Of course, such situations seem scarce but they are in principle possible. Even then, any change in the factors affecting the probability of different changes could overturn the situation.

19 Changing the habit may embody either a transformation of existing habits or applying a previously held but unexpressed alternative. This is yet another proof of the importance of Distin's recessive memes (Distin, 2005, pp.44–45, 58–59): Kim would surely already be aware of the possibility of not smoking.

subset of possible descendants contains all elements with probability value greater than zero, $0<P_x\leq 1$. This possibility is determined by actual psychological and ecological factors (Sperber, 1996, pp.41–55, 113–118) which may be studied independently, which renders the definition non-tautological. All situations in which this set of possible descendants consist of ten or fewer elements, one of which is may be identical with the original and the rest within minimal distance of it on two different axes, correspond with Sperber's model.

While the difference between Sperber's model and the Jump model may seem almost irrelevant, it has dire consequences. First of all, the Jump model allows punctuated trajectories, and such causal chains would likely be even harder to track and predict than continuous changes in form. At the very least, reconstructing possible missing links would become impossible, which would cause hardships to researchers who are interested in such missing portions. Secondly, the set of possible descendants becomes indefinite which complicates assigning probability values and deriving predictions. Thirdly, biological evolution would hardly allow such major mutations. While this could merely emphasize the unsufficiency of a naive analogy between biological and cultural evolution, the possible inclusion of saltation and hopeful monsters (Dennett, 1995, pp.287–288) endangers any analogy between these two processes. After all, the mentioned two conjectures rank among those which Darwinian evolution is supposed to dismiss by being capable of explaining speciation without conjuring them. Both saltation (Ibid. pp.285–289) and hopeful monsters signify macromutations which, against all odds, produce increased fitness (Ibid. p.288). Saltation in particular often includes outside interference in the form of ready-made design. These conjectures counteract gradualism, one of the main tenets of Darwinism, which is defended by Richard Dawkins in *The Blind Watchmaker* (Dawkins, 1986/2006a).

Fortunately, there exists a way to exorcise such implications without invoking *ad hoc* conjectures. The improbability of hopeful monsters stems from the fact that a major leap in Design Space produces *terata,* or monsters, whose fitness is devastated by the macromutations' effects on their phenotypes (Dennett, 1995, p.287). The vastness of the whole of Design Space compared to the set of adapted or even viable designs ensures that any major differentiation from accumulated adaptations almost necessarily fails (Ibid. p.288). However, in the case of ideas, those ideas may die in our stead.[20] Since the Jump model is founded on Sir Karl Popper's theory of scientific discovery, it is only befitting

20 This is, of course, a reference to Sir Karl Popper's words: ”Let our conjectures, our theories, die in our stead!” (Popper, 1978, p.354)

how his trail of thought coincides in this regard. When Popper describes the thought processes of Bertrand Russell and Alfred Einstein based on personal acquaintance, he notes how Einstein admitted having considered multiple possible models and eliminating them before realizing the theory of general relativity (Popper, 1978, p.347). In Russell's case, such thought processes were conjectured to explain the different amounts of editors' corrections in Russell's and Popper's manuscripts (Ibid.).

In effect, a jump in Design Space may be described as the result of an internal process of selection. This process need not follow the Sperberian model's rules of advancement but it avoids the problem of hopeful monsters since the conceived monsters are culled before ever being birthed. The process may be facilitated by a variant of the influence model (Sperber, 1996, pp.104–106) where, instead of tokens of similar representations being averaged, the creative search for new possibilities is influenced by normally unconnected representations.

The problem of saltation is also connected to the unexplained appearance of new features of design. Unlike the proposed alternatives regarding presumed saltation in biological design, human creativity is a real, researchable phenomenon. Saltationism is problematic because it shrouds the question of where and how the new design appeared against all odds. Since human creativity is available for research unlike the historical circumstances postulated by accounts which advocate such improbable evolutionary leaps, a similar mystification of the process does not occur. The punctuated causal chains of the Jump model could in principle be predicted and explained by applying sufficient knowledge of the processes of imagination which produce them. Thus, the variant influence model suggested above becomes unnecessary if human creativity and imagination may be modeled without invoking such synthesization.[21]

Once human creativity is sanctioned, its role in the true generation of new cultural material must be explained. Liane Gabora argues that Darwinist models, including Sperber's epidemiology of representations, cannot even in principle integrate forms of creativity other than the synthesization hypothesis presented above (Gabora, 2011, pp.69–71). The main reason seems to be these Darwinist models' aspiration to identify particular

21 Liane Gabora has also noticed the significance of human creativity regarding cultural evolution models (Gabora 1997; 2005; 2011; Gabora and Aerts 2009). She is, however, a persistent opponent of Darwinist models of cultural evolution (Gabora 1997; 2005; 2008; 2011) and her arguments concerning the dynamics of cultural transmission and the cognitive retrieval of representations would deserve being considered seriously.

lineages (Gabora, 2011, pp.69–71). Creativity, however, often involves distinct influences becoming mixed, and the descendants cannot be justifiedly categorized as part of any of the parent lineages (Ibid. pp.72–73). In biological evolution, there no longer exists generation of new information besides random mutations. However, in the primeval soup, the self-replicating chemicals which would later become the origin of life, were likely[22] to be spontaneously generated by chemical reactions which were only possible under those circumstances (Dawkins, 1976/1989/2006b, pp.14–16). Gabora herself is among those who promote the view that cultural evolution should probably be compared with this period instead of the modern state of the biosphere (Gabora, 2008; 2011, pp.75–76). While she might disagree, it would however seem possible to apply a form of Darwinism under these conditions: new cognitive causal chains may emerge due to true creative processes but their spread is approximately Darwinian as long as they persist. Thus, while new material may join the struggle, as soon as it does, the new content is subdued by the same rules to which the veterans of the game of selection adhere. Creativity is thus effectively a generator of (cultural) diversity (Edelman, 2006, p.27). Also, while lineages are currently being discussed, the unnecessity of *distinct* lineages was already established in section 2.3.

The last remaining question is how the Jump model relates to differential adaptive significance. The above discussion regarding Driscoll's insights (Driscoll, 2011) utilizes Sperber's model. As an alternative to accommodate the Jump model is construed, certain new factors need to also be considered. Firstly, the eliminative search connected with the new model is more likely to involve multiple differentially adaptively significant possible moves and contrast them with each other. This would likely affect the probability of adopting one of those representations. Secondly, the trigger for such a search in Popper's model sets minimum criteria for future generations of representations: (1) they must accommodate all previous observations compatible with the falsified notion (Popper, 1959/2002b, pp.249–251, 276), and (2) they must accommodate the incompatible observation(s) which falsified it (Popper, 1972/1979, pp.357–358). While the latter criteria were originally conceived to differentiate two theories according to their *verisimilitude*, or truth-likeness, (Ibid. pp.47–60), they may be applied to adaptive significance due to the greater reliability of action guided by representations which correspond better with the facts. As

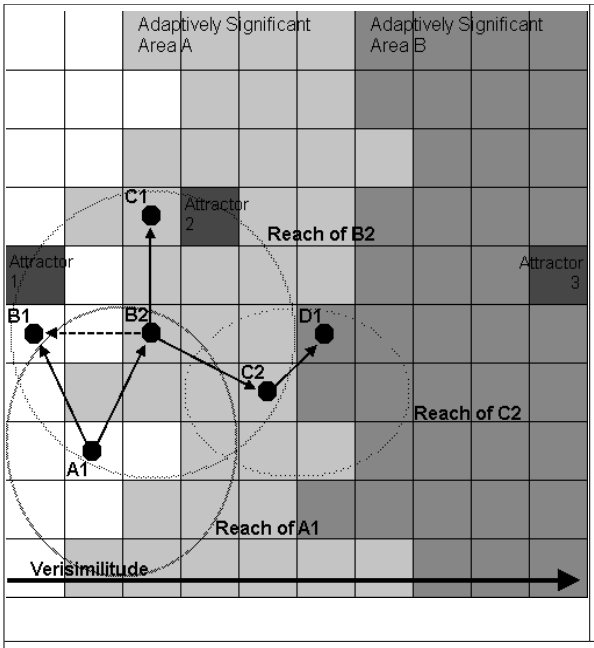
22 In all honesty, the probability of any given instance of this happening was infinitesimal. It was, however, possible and, as such, considerably more likely than in any other conditions. It could only have happened multiple times due to the amount of time involved during this period (Dawkins, 1976/1989/2006b, p.15).

David Deutsch notes, at least scientists tend to adopt well-corroborated theories exactly because they seemingly correspond to some degree with how reality functions (Deutsch, 2011, pp.256–257).

In the illustration in figure 3.3, the features of a more elaborate Jump model, which includes the notation for adaptive significance and attractors used in figure 3.1, are considered. This is combined with the elliptical notation of the hypothetical reach of transformations during a representation's reconstruction used in figure 3.2. New features of figure 3.3 include (1) denoting each example representation with a named dot as well, (2) illustrating the transformations with arrows, and (3) designating one axis to correspond with a dimension of design, verisimilitude in this instance. Thus, to have such a relation to truth, the contents of these example representations must be propositional.[23] The names of the dots signify their generation: (A1) is a first generation representation, both (B1) and (B2) belong to the second generation, and so forth.

The starting point, (A1), is adaptively insignificant and far removed from all attractors. During the next generation, (B1) is much closer to an attractor while its adaptive significance is equal to (A1)'s. (B2), in turn, is adaptively more significant than either. The fact that (B2) lies relatively close to Attractors 1 and 2 is irrelevant for demonstrative purposes. Either of these two second generation representations could be considered more probable depending on the strategy being utilized for reconstruction. If multiple moves available to (A1)'s holders under the circumstances are systemically considered, the adaptively more significant forms such as (B2) ought to prevail in comparison. However, the pull of Attractor 1 might bias the initial impression. Unless there is a systematic search beyond such impressions, (B1)'s probability of becoming adopted should exceeds (B2)'s. The dashed arrow from (B2) to (B1) designates the already discussed hypothetical possibility of transformation towards an adaptively less significant form. Since adaptive significance is conjectured to approximately correlate with verisimilitude, and since survival of crucial experiments is an indicator of comparative verisimilitude (Popper, 1972/1979, p.14), anyone aware of (B1)'s relation to (B2) should not make such a transition.

23 Because what has been discussed for the most part has been representational content in any case, this background assumption has also been haunting earlier examples to a degree. While forms of representing which have a weaker connection to truthfulness such as instructions are also considered elsewhere, they are explicitly distinguished as special cases. One such case, discussed in section 4.4, are representations which retain exact features such as the lyrics of a heard song. However, despite hardly being propositional in the sense being discussed, such cases do have a connection to verisimilitude because correspondence is of utmost importance in their case. While they may not have truth values as such, a mismatch between what is represented and what the object is resembles falsehood.



Explanations:

1. Emphasized areas:
a) Adaptively Significant Areas A and B
   - Contain all equally adaptively beneficial iterations
b) Attractors 1–3

2. Drawings:
a) Dots A1, B1, B2, C1, C2 and D1
   - Examples of representations
   - Alphabet denotes generation
b) Dotted perimeters
   - Reaches of designated representations
c) Arrows between dots
   - Transitions between forms

FIGURE 3.3. Combination of Driscoll's hypothesis (Driscoll, 2011) and the Jump model.

During the next generation, (C1) and (C2) relate to each other similarly. (C1) is drawn towards a nearby attractor and retains an adaptive significance equal to its predecessor (B2)'s, reflecting the transition from (A1) to (B1). However, there is also no improvement in adaptive significance between (B2) and (C2). (C2) has slightly higher verisimilitude than (B2) and it lies farther away from attractors but no other significant traits distinguish the two. In the pictured scenario, reaching (C2) and similar forms capable of further transformations into adaptively more significant forms which require such a transitive phase would be a gamble.[24] Because all possible transitions are only differently probable, such moves would nevertheless remain actualizable as long as the circumstances in reality indeed permitted them. For example, the different verisimilitude between (C1) and (C2) would imply the existence of at least one crucial experiment capable of falsifying (C1) but not (C2). Discovering such a crucial experiment in practice or during mental simulation would lower or even eliminate the possibility of ever adopting (C1).

Finally, the transition from (C2) to (D1) would increase adaptive significance again. Unless (C1) had a longer reach than any of the shown examples, Area B could only be reached by a trajectory involving (C2) or similar, unnamed forms. This transition is also

24 This only applies to the illustration since the is no reason why attractors and other significant factors should steer away from such transformations instead of guiding them. For example, the cell below (D1) could have housed an extra attractor if the demonstration were to apply to a real situation. Driscoll's claim (3) regarding the potential quantity of attractors applies (Driscoll, 2011, pp.313–314).

similar to Sperber's examples (Sperber, 1996, pp.111–112). If the relevance theory is indeed correct, interpretations should gravitate towards the correct interpretation since longer leaps would involve a greater misunderstanding. As long as communicators' understanding of what is manifest to their receivers is approximately correct, they should be able to provide sufficient cues to avoid such failures of interpretation (Sperber and Wilson, 1986/1995, pp.163–170). If communication is continued, wide misunderstandings should become corrected because they would disrupt any further exchanges (Schegloff, 2006).

Thus far, the role of Attractor 3 has been left unexplained but it actually represents the hypothetical possibility of reaching a truth. This truthfulness is to be understood in the sense of correspondence and perfect truth content from which only true consequences may be derived (Popper, 1972/1979, pp.54–55). In principle, such a conception could never be falsified although its truth could also never be verified if it was non-tautological. As such, no further falsification would force outgoing jumps by inducing aporia concerning the representation's content. Repeatedly misunderstood communication or badly construed public representations could however cause trajectories to deviate. Truth as an attractor thus seems justified although it may only activate under specific circumstances. For example, the above speculation would imply that were the corresponding representation ever to become represented by someone, this could avert future deviations in their case.

To summarize, let us contend that movement in Design Space is likely more diverse than Sperber's illustrations (Sperber, 1996, pp.111–112; Claidière and Sperber, 2007) would lead to understand. Firstly, Driscoll seems to correctly demand for adaptive significance to be considered (Driscoll, 2011). Secondly, there exist cases of discontinuous trajectories in Design Space. It bears repeating that while the reaches of such jumps have been depicted as relatively regular ellipses, they might even contain gaps, depending on the dimensions of design contained in the used graphic representations. Despite their importance having been downplayed in section 2.3, also considering the role causal chains of transmission (Sperber, 1996; 2006) play is noteworthy because they still remain a part of the wider picture. Firstly, representations which contain similar contents are likely to share an ancestry although alternative explanations are also available. Secondly, representations would indeed seem to form causal chains although the links of these chains may sometimes become tangled or untraceable. Thus, while Sperber's epidemiology is of secondary interest compared to the relevance theory, it still contains plausible claims worth consideration.

3.4 Chapter summary

This presentation of the theoretical background and main features of Dan Sperber's epidemiology of representations (Sperber 1996; 2001; 2006) first introduced the relevance theory of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) and the role of the human capacity to metarepresent in Sperber's account. However, the attractors Sperber conjectures (Sperber, 1996; Claidière and Sperber, 2006) received the most attention. The latters' influence was admitted but delimited by the adaptive advantage people may derive from different representations (Driscoll, 2011). The presumption of only minimal movement in Design Space which is inherent in Sperber's model (Sperber, 1996; Claidière and Sperber, 2006) was also criticized and replaced with the *Jump model* which integrates Popperian influences regarding the evolution of knowledge (Popper, 1959/2002b; 1963/2002a; 1972/1979).

While the the relevance theory and metarepresenting which lie behind Sperber's epidemiology are scarcely referred to in later exposition of his theory, they are inherently present in the whole discussion of the attraction model. The intentional stance (Dennett, 1987; 1996; 2009a), for example, is crucial to inferential communication (Baron-Cohen, 1995). The whole project of the epidemiology of representations, on the other hand, is motivated by the relative unreliability of such inferences (Sperber, 1996).

The discussed views will be attempted to be combined into an unified framework compatible with the view of selectionism discussed earlier. Important contributions introduced in this chapter include (1) Sperber's inferential view of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012), (2) the three stances: physical, design, and intentional stance (Dennett, 1987; 1996; 2009a), (3) features of metarepresenting (Dennett, 2000; Sperber, 1996; 2000c), and (4) Cognitive Causal Chains (Sperber, 2001; 2006). Further elaboration of the emerging model also requires attending to (5) factors of attraction (Sperber, 1996, pp.106–118; Claidière and Sperber, 2006) and (6) adaptive significance (e.g. Boyd and Richerson, 1985; Henrich and Boyd, 1998; Driscoll, 2011). Depicting the change of conceptions realistically may also require inspecting the possibility of non-minimal movement in Design Space. An outline integrating these and earlier themes will finally be provided in the coming chapter.

4. OUTLINE OF A THEORY OF ASSOCIATING

"*Every established order tends to produce (to very different degrees and with different means) the naturalization of its own arbitrariness.*"

– Pierre Bourdieu, *Outline of a Theory of Practice* (Bourdieu, 1972/1977, p.164)

"*Yet the theory that truth is manifest not only breeds fanatics – men possessed by the conviction that all those who do not see the manifest truth must be possessed by the devil – but it may also lead, though perhaps less directly than does a pessimistic epistemology, to authorianism.*"

– Sir Karl Popper, *Conjectures and Refutations* (Popper, 1963/2002a, p.11)

This chapter complements previous accounts in order to help formulate an unified account about the evolution of culture. An important theme is experienced certainty: despite the existence of multiple alternative beliefs regarding the same objects, many seem to indeed regard their own rationales as justified above those of others. Bourdieu offers one explanation for this phenomenon: when conventions become established, they produce an experience of being justified (Bourdieu, 1972/1977, p.164). Most naturally, such a claim coincides with the claimed selfishness of memes (Heylighen, 1992; 1999; Distin, 2005, Chapter 8). Both the effects of established conventions and the connections between key theories are among the themes which will now be handled.

This chapter is divided into four broad subjects. First, Pierre Bourdieu's theory of fields (e.g. Bourdieu, 1972/1977; 1979/1984/2010) and its connections to the already provided outline will be introduced briefly. His theory is an important thread which helps knit together the earlier positions. The rest of the sections are named according to Sperber's trifecta: (1) cultural representations, (2) mental representations, and (3) public representations (Sperber, 1996; 2000c). Each section will outline a general view which utilizes previous themes and binds them together using some further considerations.

The introduction to the theory of fields consist primarily of Bourdieu's reflexive sociology's key concepts such as (1) doxa, (2) habitus, (3) capitals and (4) fields (Bourdieu 1972/1977; 1979/1984/2010; Bourdieu and Wacquant, 1992; Grenfell 2012). These concepts will be utilized when the rest of the account is parsed together, and their connectedness with previously discussed themes is thus established.

The section about cultural representations attempts to clarify the exact relationship of the mental and public representations which constitute cultural representations. This relationship will be structured by using Sir Karl Popper's theory of three worlds

(Popper, 1972/1979; Popper and Eccles, 1977/1983). Despite Popper's interactionism (Ibid.), this treatment will originate from a materialist viewpoint which accepts the inclusion of explanations which integrate facts regarding mental states and cultural factors (Sperber, 1996, Chapter 1; 2001, p.298; Deutsch, 1997, pp.20–23).

Mental representations are claimed to be the true progenitors and progeny of culture. The corresponding section will address the exact nature of mental representations, including the effects which metarepresenting and experienced certainty (e.g. Peirce, 1955, Chapters 1–4; Wittgenstein, 1969) have on those mental tokens. That section will also explore the formation of primary representations using Dennett's three stances (Dennett, 1987; 1996; 2009a) and contextual relevance (Sperber and Wilson, 1985/1996; Wilson and Sperber, 2004; 2012). That account will also be expanded to include metarepresentations.

Public productions are argued to be better understood as the interactors (Hull, 1980; 2001) of existing mental representations because they are the *means* of the latter's propagation. They are formulated by the hosts of mental representations and mostly manifested to transmit these conceptions' relevant features. Were the public representations conceived as replicators themselves, the reliability of information transmission made possible by the expressions' relationship with their audience would be abolished. This would justify Dan Sperber's critique regarding the role of transformations in communication (Sperber, 1996, Chapter 5; 2000a). Thus, only mental representations are attributed this role. Both the relatively reliable, audience-catering expressions and artifacts derived from mentally stored instructions are addressed as such public products. The pragmatic relation between deliberately communicative formulations and the intended audiences will be analyzed using the relevance theory (Sperber and Wilson, 1985/1996; Wilson and Sperber, 2004; 2012). The problem how the instructions to produce copies of artifacts may be derived from observing them although the original information becomes reconstrued only accidentally (Distin, 2005, pp.81–83, 92–95) will also be addressed.

The account being presented is a highly abstract framework and the claims it contains should be tested to verify their plausibility. Were the theory instead to become falsified, its contributioin would not be abolished. Such an eliminative process is a requirement for the evolution of theories (Popper, 1959/2002b; 1963/2002a; 1972/1979). This framework should still convey a new way in which the reader may relate to culture. It may even contains kernels of truth which will allow the construction of yet better accounts.

4.1 Theory of fields

Bourdieu's sociology provides further tools to analyze the dynamics of cultural transmission as well as the differential preservation of contents, and it complements the views introduced earlier. In particular, an explanation regarding the role actors have in the perpetuation of prevailing conceptions is provided. Bourdieu links together established practices and people by examining the effects of the objective on the subjective and *vice versa* (Bourdieu, 1987/1994, p.65 cited in Maton, 2012, p.49). Stabilized contents of different cultural environments are an important ecological factor which affects further entry into the collective consciousness (Sperber, 1996, pp.115–116; 2006, pp.444). The terminology Bourdieu provides helps analyze how the stabilization of contents occurs and elaborate the dynamics of different cultural selection environments.

The main concepts being introduced are (1) doxa, (2) habitus, (3) (social) capitals, and (4) fields. Bourdieu's project is quite holistic, and these concepts only become intelligible in relation to one another (Bourdieu and Wacquant, 1992, pp.4–7; Maton, 2012, pp.48–54). However, this applies to most of the introduced notions. Similarly to Bourdieu's theory, the universal theory of selection contains ultimately relational elements: the fitness of replicators may only be evaluated in relation to the selection environments they inhabit. Thus, the criteria for selection may only be understood relationally as well.

The meaning of *doxa* in Bourdieu's sense differs slightly from its standard usage in philosophy. Instead of calling all opinions which do not fulfill the formal criteria of knowledge doxastic, Bourdieu discusses the mentalities engraved in different cultures (Bourdieu, 1972/1977, pp.164–171; Deer, 2012, pp.114–115). Such doxa is explicitly contrasted with consciously accessible opinions which may be either orthodox or heterodox in relation to it (Bourdieu, 1972/1977, pp.167–170). Similarly to the traditional sense of doxa, the widespread conceptions which constitute the doxa in Bourdieu's sense are ultimately arbitrary (Deer, 2012, pp.114–115). This arbitrariness resembles the role cultural drift has in the formation of distinct societies, or social populations, ascribed to it by Boyd and Richerson in their criticism of adaptionism in memetics (Boyd and Richerson, 2000). Thus, while ecological factors affect which assumptions can become common sense within a population, the original pools of available conceptions limits what representations any given society may be expected to host (Ibid.). Likewise, populations will become differen-

tiated into distinct ethnicities by the beliefs their members share because these common assumptions facilitate co-ordination within these groups (Nettle and Dunbar, 1997; Henrich and Boyd, 1998; Henrich and McElreath, 2003, p.133; McElreath et al, 2003).

The doxa of a culture is thus not wholly arbitrary because cultural selection shapes it by preserving the adaptively significant aspects of the conceptions it contains (Boyd and Richerson, 1985; Driscoll, 2011) for example. Instead of being random, the arbitrariness Bourdieu discusses is used to criticize the claimed objective justification, the *misrecognition*, of different beliefs (Bourdieu, 1979/1984/2010, pp.440–434; Moore, 2012, pp.100–101). Effectively, all (doxastic) beliefs could have been any of multiple alternatives despite the possible benefits they confer. No universal justification for any given impression is possible. Popper also supports this impossibility of justifying held conceptions (e.g. Popper, 1959/2002b; 1963/2002b; 1972/1979).

For the objectives of this study, though, the effects of the prevailing doxa of any given cultural selection environment are more important than its origins. If the dynamics of cultural evolution are universal, this process of dominant doxa being formed simply reminds how new ideas may become the common sense of tomorrow despite the influence of their predecessors. In effect, emerging awareness about a doxa produces a conscious orthodoxy, the adherents of which attempt to restore the doxa, and a heterodoxy which challenges those established practices (Bourdieu, 1972/1977, pp.167–169; Deer, 2012, pp.118–119). This heterodoxy remains subjugated by the doxa because communicating it would require utilizing the prevalent structures (Ibid.). However, if a heterodoxy supersedes the previous doxa, the latter is eliminated and the former instituted as the new doxa.

What defines the doxa is the unableness of its hosts to recognize the existence of its constituents (Bourdieu, 1972/1977, p.164). To those people, there exist no conceivable alternatives. If they realize the existence of different opinions regarding the contents of a doxa, these contents lose their doxastic nature. Instead, they become mere opinions, towards which people and societies may have different attitudes. These beliefs may now be classified as either orthodox or unorthodox (Ibid. pp.168–170) and manipulating them consciously is enabled. However, doxas are not particularly vulnerable: the doxas of large societies might be shielded by the sheer amount of people because emancipated individuals may only affect so many. Besides, social factors such as the conformist bias (Boyd and Richerson, 1985, Chapter 7; Henrich and Boyd, 1998) would still hinder the spread of

deviating viewpoints. Even Bourdieu accepts the role such frequency-dependent biases have in the maintenance of doxa by noting how the actions of a group's other members affirm the certainty of their unquestioned views (Bourdieu, 1972/1977, p.167).

Thus far, doxas have only been attributed to homogenous populations where everyone shares the same doxa. While Bourdieu indeed originally conceived ethnic groups as the hosts of doxas, the concept may be expanded to cover the participants in specific fields (Deer, 2012, pp.116–125). There exist field-specific doxas, and individuals may host a multitude of them because people partake in multiple fields (Ibid.). Bourdieu himself has researched the assumptions which regulate different spheres of life such as aesthetics and everyday life (Bourdieu 1979/1984/2010), and higher education (Bourdieu, 1984/1988). Deer argues that the latter dissection of *symbolic power* (e.g. Bourdieu, 1991) also pertain to doxa because the capitals social agents (including groups) accumulate help subjugate fields to certain patterns of thinking (Deer, 2012, pp.116–117). Connections with the dual-inheritance theory are present because such situations resemble cases of prestige-based bias (Boyd and Richerson, 1985, Chapter 8; Henrich and McElreath, 2003, pp.129–130).

In effect, the doxa upheld by the participants on a field is a powerful ecological factor which affects the fitness of all conceptions which share niches with doxastic beliefs. Within a field, it is the ultimate common ground (Enfield, 2006) shared by all participants. Thus, even attempting to communicate any alternatives may fail because the premises required to decipher the intended meanings are prevented from becoming *manifest* to the receiver. Whatever the doxa, it would seem to consists of particularly successful cultural representations, the continued success of which may be attributed to its mental tokens having become almost pastel (Dennett, 2000). In one sense, a doxa is public – people share seemingly the same beliefs but its inconspicuousness complicates matters. This indiscernability is generated by no alternatives being observable, which makes a doxa's contents seem obviously true (Bourdieu, 1972/1977, p.164). Because a doxa determines how experiences are structured, its adherents may integrate even contradictory observations as confirmations of its truth (Ibid. 164–170; Deer, 2012, p.118).

The (1) habitus, (2) capitals and (3) fields are a too tightly interwoven to be truly discussed separately. While the doxa concerns their contents having become absolutized, these concepts themselves explain how dispositions are formed (Maton, 2012, pp.49–52). Bourdieu expresses their relationship thus: '[(habitus)(capital)]+field=practice' (Bourdieu,

1979/1984/2010, p.95; Maton, 2012, p.50). In other words, a habitus and its owner's set of different capitals comprise an assortment of possible practices which are realized according to their experienced appropriateness on different fields of life.

The habitus is effectively a social actor's accumulated set of dispositions which may be expressed as the habitus being a 'structured and structuring structure' (Bourdieu, 1987/1994 cited in Maton, 2012, p.50). Maton interprets this as the habitus (1) being structured by all the actor's experiences so far, (2) structuring present and future actions, and (3) having a systematic structure or patterning (Maton, 2012, p.50). What the term denotes is a system of dispositions which yield 'a *way of being*, a *habitual state*' (Bourdieu, 1972/1977, p.214). A person's bearings is evaluated socially based on two dimensions: (1) accomplishment and (2) transposability (Moore, 2012, p.111). The former depicts the field-dependent evaluations of the forms of capital embodied in a habitus being expressed and the latter the range of fields in which positive attitudes towards such inclinations prevail.

What is accumulated during a habitus being formed are multiple types of social capital. Economic capital depicts assets with monetary value and the relative economic security of agents (Bourdieu, 1979/1984/2010, pp.375–381). It affects how and what a person habitually consumes and thus both shapes and enables the expressing of their habitus (Ibid.; Moore, 2012, pp.103–106). However, the habitus is primarily concerned with forms of symbolic capital, the value of which is field-dependent (Bourdieu, 1979/1984/2010, Chapter 4; Moore, 2012, pp.100–103). These include the likes of (1) cultural capital, (2) linguistic capital, (3) scientific capital and (4) literary capital (Ibid.) – in other words, levels of mastery pertaining to definite types of public information. Possible examples of the differential evaluation of such capacities between fields include how literary knowledge is a necessary condition to amass status amongst members of the humanities but affects the appreciation an actor receives less in the natural sciences. In less academic contexts, expressing being well-read may even alienate peers. While at least minimal levels of key capitals are required for participationship on certain fields, causing differences on the mean levels of those capitals between them, the amounts of possessed capitals also differentiate actors within fields (Moore, 2012, p.99).

As has been stressed above, the perceived value of possessing and expressing symbolic assets is always proportional to the fields of life one inhabits. In the formulation '[(habitus)(capital)]+field=practice' (Bourdieu, 1979/1984/2010, p.95; Maton, 2012, p.50),

the habitus and capitals of a person constitute one operator while the fields are another. The fields are social spaces which contain actors who constantly interact with their constitution: both being affected by the established rules and attempting to secure desirable standards. Their actions are delimited by the existing norms and expectations because all possible moves available to an actor are determined by their dispositions and resources while the subset of available maneuvers is defined by the present contexts (Bourdieu, 1979/1984/2010, Part II; Thomson, 2012, pp.66–72).

The theories of cultural evolution and Bourdieu's project are both attempts to answer the question how intersubjective patterns of conduct arise from the behavior of individuals even without explicit commands (Bourdieu, 1987/1994, p.65 cited in Maton, 2012, p.49). According to Bourdieu norms are reproduced to incorporate new subjects into the proper sectors which causes these people to construct their identities accordingly (e.g. Bourdieu and Passeron, 1970/1977/1990; Bourdieu, 1979/1984/2010; 1984/1988). Like all the theoretical entities invoked in this study, the entities Bourdieu discusses actually only exist as embodiments of information in (physical) mental and public structures.

The interrelatedness of the theories being investigated begins with the recognition that Bourdieu explicitly discusses single actors which host a distinct set of representations developed during interaction with one's environment. While the contents and value of these social resources are not fixed, the same set characteristic of any given individual is retained across contexts. The earlier criteria affecting a habitus – accomplishment and transposability (Moore, 2012, p.111) – help understand the significance of this simple fact and its relationship to Sperber's attractors (Sperber, 1996, pp.106–118). Specifically, these effects of a field may be understood as specific ecological factors which are characteristic of that cultural environment (Ibid. pp.115–116). These factors change between contexts but the objects of their attraction remain the same.

Thus, the influence of factors such as the biases discussed by Dual-Inheritance Theorists (e.g. Boyd and Richerson, 1985; Henrich and Boyd, 1998; Henrich and McElreath, 2003) may each attract the same representations in multiple directions in Design Space because expressing those same representations is evaluated differently on separate fields. This resembles Driscoll's claim that sub-cultures contain their own attractors, due to which no attractors within a population may be universal (Driscoll, 2011, pp.317–320). The value in accomplishment representations have denotes the content of

such evaluations and transposability describes the ratio of favorable contexts to unfavorable and neutral ones (Moore, 2012, p.111). Driscoll describes how a high amount of attractors tends to dampen their average effect (Driscoll, 2011, pp.313–314). But if the attractors on multiple fields were layered according to the design features they entail, some designs could be promoted by multiple attractors from different fields. Each attractor could theoretically be assigned an attraction value according to the strength of its pull it as well as an actor's level of conformity towards the demands of the corresponding field. This value would represent the force of a vector towards the attracting design, and the summing of such influences would be greatly affected by the multiplicatory effects of overlap.[1]

Bourdieu's analysis also applies to Heylighen's system of criteria for selection (Heylighen, 1993; 1997; 1999). First, subjective criteria such as novelty, simplicity, coherence and utility (Heylighen, 1999, p.4) receive their value in relation to the habitus a person embodies. While all these criteria relate to the contents of a person's mental resources, the evaluation of different traits is also affected by espoused habitus which are characteristic of the fields to which a person belongs. For example, Bourdieu researches the different logics of aesthetic evaluation particular to different social classes in *Distinction* (Bourdieu, 1979/1984/2010). Notably, the bourgeois appreciate novelty because these people always attempt to distinguish themselves from the commonplace (Ibid. Chapter 6). In contrast, members of lower classes are raised to prefer simple and useful beliefs and items (Ibid. Chapter 7).

Secondly, the intersubjective criteria of authority, formality, conformity, expressivity, and publicity (Heylighen, 1999, pp.4–5) are subject to the prevailing misrecognitions.[2] Particularly the doxa determines what is deemed appropriate (Bourdieu, 1972/1977, pp.164–171; Deer, 2012). It often determines the conventions to which conceptions should conform as well as which sentiments remain intelligible when expressed. Also, the emergence of heterodoxy and orthodoxy when people become aware of the doxa (Bourdieu, 1972/1979, pp.168–171) should affect the contributions these criteria offer. At that point, people's allegiances become divided and questions are raised: to whose doxy and its rules does a given opinion's formulation conform?

1 The adaptive significance Driscoll emphasizes (Driscoll, 2011) also includes how highly a given conception is evaluated between fields because such evaluations affect whether the hosting individual will thrive.

2 This means the claims of objectivity and intrinsic value of one's beliefs, when establishing them as the appropriate rules of a field is attempted (Moore, 2012, pp.100–101). It is also called misattribution (Bourdieu, 1979/1984/2010, pp.440–434).

Finally, the meme-centered criteria like self-justification, self-reinforcement, intolerance, and proselytism (Heylighen, 1999, p.5) fit Bourdieu's framework perfectly. Any doxa is deemed particularly self-reinforcing because all actions confirm it in the absence of alternatives (Bourdieu, 1972/1979, p.167). On the other hand, the different discourses fields contain are depicted as constantly struggling, and standard strategies include attempting to establish a conception's objectivity, or otherwise causing misrecognitions which obfuscate the viewpoint's arbitrariness (Moore, 2012, pp.100–103; Schubert, 2012). Such self-justification also involves intolerance towards deviating perspectives.

Further connections this union between the theories engenders will be discussed in the appropriate sections in this chapter. The current deployment of Bourdieu's terminology and its theoretical commitments will only be partial because his project is too vast to be fully described here. It should also be noted that an eclectic approach like the one being utilized may easily mispresent Bourdieu's theory (Bourdieu and Wacquant, 1992, pp.4–7). The present account may thus contain some necessary recontextualization of the introduced concepts despite their intended meaning being attempted to be retained.

Bourdieu's already overexploited formula '[(habitus)(capital)]+field=practice' (Bourdieu, 1979/1984/2010, p.95; Maton, 2012, p.50) continues to encapsulate the interrelatedness of his key concepts. A social actor's habitus and possessed capitals constitute how they may act. The habitus itself is also produced by the possession of different capitals because such resources make certain patterns of action available (Moore, 2012, pp.107–108). For example, an inclination to buy books may reinforce literary capital but expressing it at any given point in time also requires certain amounts of economic capital. Likewise, inheriting literary capital may be necessary for such habits to ever even emerge (Bourdieu, 1979/1984/2010, pp.73–78).

These available actions are turned into a practice by being selected according to the criteria inherent in the fields being attended. Those criteria are founded by the actors themselves and grounded by attempts to justify them (Moore, 2012, pp.100–103; Thomson, 2012, pp.69–72). What playing this game secures are further capitals, particularly on the almost all-encompassing *field of power* (Bourdieu, 1979/1984/2010, Part II and Chapter 8; Thomson, 2012, pp.68–69). A central position in the hierarchy of power will allow determining the rules which affect relatively less powerful actors who share the same social space (Ibid.). Thus, the fields could also be likened to playing fields where the

players are allowed to construct obstacles and each occupies a position and a role pertaining to it (Thomson, 2012, pp.66–67).

The doxa is present in all these levels. The unuttered, subliminal rules which affect each field and are shared by all those who attend it are infused into these actors' ways of being. Ultimately, the doxa determines the rules to which all must adhere (Bourdieu, 1972/1979, pp.164–171). These dynamics of power stem equally from agents and the conceptions they share, and both those sources are equally subjugated by the rules thus established. This duality of the actors and their shared beliefs affecting one another and vying for dominance is what connects Bourdieu and cultural evolution.

4.2 Cultural representations

Thus far, it has been established that the claimed problems of generalizing accounts (Sperber, 1996, Chapters 2 and 3) are exaggerated. Sperber uses the notion *cultural representations* in at least two senses – (1) the unjustified abstractions made by (most) anthropologists (Ibid.) and (2) 'anything that is both cultural and a representation' (Ibid. p.77) – and objects to the first usage. Indeed, the latter view which is highlighted by Sperber's later writings on Cultural Cognitive Causal Chains (Sperber, 2001; 2006) and considers a lineage of mental and public representations a cultural entity (Sperber, 1996, Chapter 4; 2001; 2006) is quite compatible with the views being presented here. Arguably, that account does not suffice and generalizations beyond individual lineages would thus be justified. The study of the formal properties of representations which is deemed a fancy by Sperber (Sperber, 1996, pp.62–63) is not merely possible – it seems necessary.

Cultural representations in the sense discussed here are information structures which may be shared by representations regardless of their descent. The argument which supports this view could be construed as follows: (1) New representations are constantly generated by human creativity which acts as a generator of diversity (Edelman, 2006, p.27). (2) An account which only tracks individual lineages could not explain the processes which direct this generation of new representations (Gabora, 2011, pp.9–14). (3) These processes are likely affected by factors which bias the distribution of different information structures being generated. The structures frequently being produced may be described either as *attractors* (Sperber, 1996, pp.106–118) or *free-floating rationales* (Dennett, 1995,

pp.132–133; 2013, Chapter 39). (4) The process of generating new representations may be described as a form of selection because these biases produce differential fitness. (5) Because the same structures are being selected for among the distinct lineages produced by this process, researching the distribution and fitness of these general features is viable. (6) Thus, the lineage-dependent accounts of Sperber (Sperber, 1996; 2001; 2006) and Hull (Hull, 1982) are not the only viable option for studying the evolution of culture.

In effect, what is being suggested is an account which disengages from the analogy with biological evolution because unlike genetical information, new cultural information is seemingly generated constantly, and because modeling these new representations as mere recycling of old information is untenable (Gabora, 2011, pp.9–11). Despite the impossibility of biological lineages which do not branch from the tree of life but may intermingle with it,[3] this implausibility of multiple beginnings is not a necessary criterion for evolutive processes. Because the processes of handling cultural information feature crossing over (Dennett, 1995, pp.355–357) which causes ambiguous ancestries (Ibid.; Gabora, 2005; 2011, pp.12–13), tracking lineages accurately becomes impossible. Instead, the convergent evolution of multiple lineages may also be described as the selection for the shared traits among these lineages. What is being studied are not particular lines of descent but existing tokens, and the latter are categorizable regardless of their muddled origins.[4]

The position established thus far is summarized above, and elaborating it further may now begin. The most important remaining question concerns the exact relationship of this view and cultural lineages. While the importance of lineages is de-emphasized, it is far from abolished. The generation of new lineages is admitted but this does not imply that all representations are the result of personal creativity. Indeed, the main level of cultural selection is probably between existing lineages despite new contenders appearing occasionally. Possibly the greatest problem is posed by how these lineages intermingle constantly (Dennett, 1995, pp.355–357; Gabora, 2005; 2011, pp.9–13). Thus, what must be

3 Actually, while this would be a near-impossibility in nature as such, synthesizing genes should not be impossible in principle. Since such synthetic genes would not originate from the natural replication processes of existing genetic material, they would effectively reflect the view on the generation of cultural material being presented here. Also, the possibility of the genetic code having emerged multiple times in the conditions of the primeval soup and these trees of life having converged is not impossible in principle – it is merely extremely improbable (Dawkins, 1976/1989/2006b, p.15).

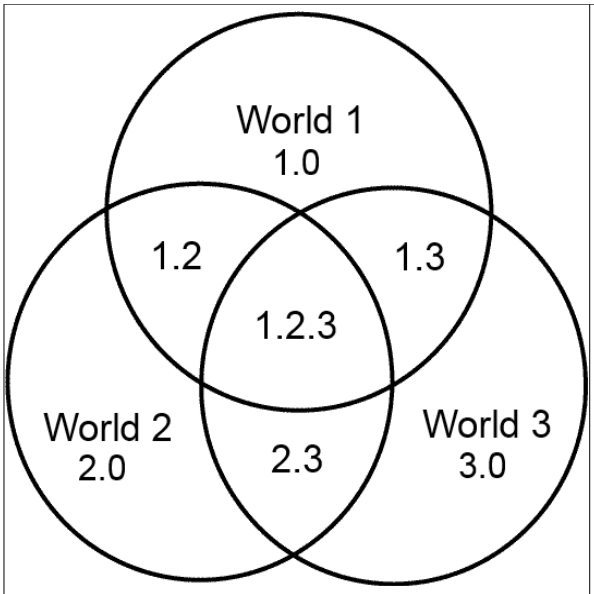
4 Although another possibility is that representations are actually not distinct at all. In this case, what would be studied are the structures present in the holistic worldviews held by people. However, the analysis presented here should accommodate this possibility since both the traits of interest and their objects are designated by the researcher. Supposing the naturalness of these distinctions is unnecessary. The holistic replication of entire worldviews has been explored by Gabora (2004; 2011; Gabora and Aerts 2009).

explained is how such interrelating does not prevent a form of selection from occurring. The more detailed answer to this question will be structured using Sir Karl Popper's three worlds (Popper, 1972/1979; Popper and Eccles, 1977/1983).

Popper's three worlds are essentially an attempt to further develop classic interactionism (Popper and Eccles, 1977/1983). Besides (1) the material layer and (2) the mind, being posited, a third dimension, (3) culture, is also included (Popper, 1972/1979, Chapters 3 and 4; Popper and Eccles, 1977/1983, Chapter P2)[5]. However, although Popper identifies himself as an interactionist, his position contains some quite unorthodox features. Firstly, Worlds 2 and 3 are claimed to emerge from World 1, the material, instead of existing independently (Ibid. pp.15–16, 554). What makes them possible are the properties of the human brain which has been generated by natural selection (Ibid.). Secondly, Popper questions the utility of the notion of a *substance* (Ibid. p.105). Thus, he is clearly no substance dualist, unlike more traditional interactionists like Eccles (Ibid. Chapter E7). This might suffice to justify using his theory to describe a materialist position.

Thus, despite the original author's aspiration to defend a form of dualism, his concepts may be utilized to illuminate a wholly materialist framework. Hull defies the postulation of Popperian worlds besides the material when memes are being researched (Hull, 2001, p.33). Likewise, Dan Sperber argues that every subject of psychological and anthropological research must be materially embodied and that discussing these embodiments should not be mere lip service to superficial materialism (Sperber, 1996, Chapter 1, pp.79–82; 2001, p.298). However, Sperber does not deny the possibility of utilizing non-materialist vocabulary – on the contrary. Like Dennett, he allows such language as long as the serious research pertaining to its objects can identify the corresponding material structures (Dennett, 1987, Chapter 6). Of course, Dennett is better known for questioning the justification of the folk psychological concepts which still direct much of psychological and neurophysiological research (e.g. Dennett, 1991) but these two positions overlap. In effect, both Dennett and Sperber deny approaches which unquestioningly postulate the existence of any mental or cultural entities assumed by commonsensical conceptions. Paricularly Dennett's project is an heir to his teacher's, Gilbert Ryle's, attack on dualism (Ryle, 1949/2000) in this regard.

5 *The Self and Its Brain* (Popper and Eccles, 1977/1983) is divided into three parts: (1) Popper's philosophical arguments, (2) Eccles's neurophysiological arguments, and (3) dialogues between the authors. Thus, the notation 'P2' which is also used in the book, denotes chapter two written by Popper.



Explanations:

1. Worlds 1, 2, and 3
   - The three worlds conjectured by Sir Karl Popper (Popper 1972/1979; Popper and Eccles, 1977/1983)
   - (1.0), (2.0) and (3.0) denote the sets of entities existing only in single worlds.

2. Intersections (1.2), (1.3), (2.3), and (1.2.3)
   - Denote the sets of entities which exist in all the worlds listed by the numbers in these intersections' names.

FIGURE 4.1. A Venn diagram demonstrating the possible relationships between Sir Karl Popper's three worlds (Popper 1972/1979; Popper and Eccles, 1977/1983).

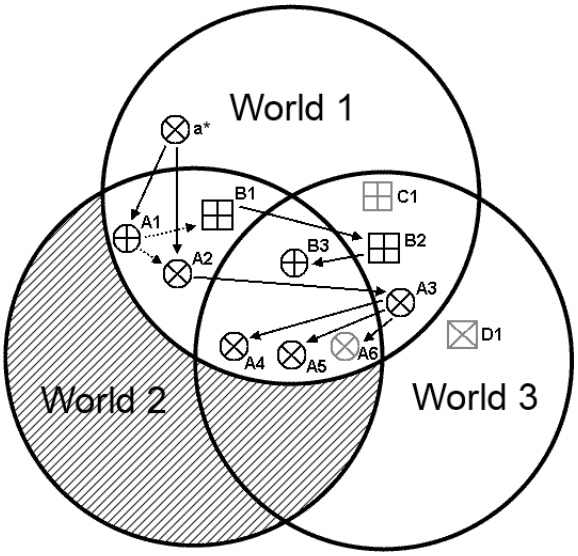
A simple Venn diagram of the three worlds should clarify the positions assumed by different theorists as well as the viewpoint this study represents. In figure 4.1, Popper's three worlds are represented by the corresponding numbers from one to three. Since Popper never claimed the full independence of these worlds – he accepts the physical embodiment of the denizens of Worlds 2 and 3 (Popper and Eccles, 1977/1983, pp.36–41) – there exist intersections. While he has argued that Worlds 1 and 3 do not interact directly and that World 2 interacts with both by conveying their relations (Popper, 1972/1979, pp.155–156), this does not deny the existence of physically embodied cultural objects. Nevertheless, this diagram which includes all possible relations between the worlds may also depict perspectives which deny some of these possibilities.

Figure 4.1 depicts Worlds 1, 2 and 3 as intersecting sets which are denoted by identical circles. Their intersections are named according to their constituting worlds, and the numbers of these worlds are deparated by a period. Here, the preceiding worlds are always mentioned first because of the assumed primality of World 1 in relation to the rest and of World 2 with respect to World 3. However, an ontological idealist might rearrange this order. For example, (1.2) would become (2.1), were World 2 perceived as the primal layer instead of World 1. Subsets (1.0), (2.0), and (3.0) denote entities which are particular to their respective worlds and not contained in any of the intersections.

This figure helps clarify the relationships of different perspectives on ontology. Basically, any subsets, the existence of which is denied, are designated as empty sets. For example, a materialist viewpoint like Sperber's (Sperber, 1996, Chapter 1; 2001) where everything is materially embodied may be described by affirming the emptiness of sets outside World 1. Thus, a materialist would conjecture that $[(2.0)\cap(3.0)]\cap(2.3)=\{\varnothing\}$. This example allows the existence of mental and cultural entities which are also contained in World 1. Set (1.2) would contain mental contents while (1.3) holds culture-embodying artifacts and (1.2.3) consists of the mental representations of cultural entities.

Sperber's musings regarding the existence of unrepresented beliefs (Sperber, 1996, pp.86–87) complicate this picture a little. It is argued that conjecturing that propositions which are inferable from a person's existing representations may also be called their beliefs is reasonable (Ibid.). This resembles Popper's claim that World 3 also contains the propositions which may be derived from existing theories (Popper, 1972/1979, pp.115–119; Popper and Eccles, 1977/1983, pp.41–43). Both views superficially contain a sort of Meinongian subsistence. This is particularly obvious in Popper's works where all the theories which may ever become conceived are also assumed to somehow exist in World 3 (Ibid.). However, one interpretation of Sperber would actually deny the peculiar existence of these unrepresented beliefs. He highlights how these beliefs are inferable and how that capacity is innate to humans (Sperber, 1996, pp.87). What triggers this process of inference is an outside stimulation such as attempted communication (Ibid.). Thus, such beliefs arguably *do not exist* until generated because of an intervention. Once again, this seems in line with Ryle's critiques. Deriving an exhaustive list of the postulated entities called personal beliefs introspectively seems simply futile (Ryle, 1949/2000, pp.156–160).

This diagram is also useful when one dissects cultural representations. However, minor adjustments will be necessary for that task. An example illustration which integrates them is provided in figure 4.2. It depicts the position outlined earlier in this study. In this illustration, the tokens being studied may be located in the diagram according to the set(s) in which they belong. Each token represents one of two object types according to its color. In turn, each drawing which corresponds to a token consists of a shape and an interior design, both of which have two alternatives. These two features of their form represent two distinct attributes, each of which has two alternative traits corresponding to it. Arrows denote the trajectories the lineages of these representations produce.



Explanations:

1. Worlds 1, 2, and 3
   - The three worlds conjectured by Sir Karl Popper (Popper 1972/1979; Popper and Eccles, 1977/1983)
   - Diagonal stripes: empty sets

2. Drawings:

a) Tokens (A1), (A2),... (A6), (B1), (B2), (B3), (C1), and (D1); and Object (a*)
   - Color (black/grey): type of object
   - Shape (circle/square): Allele 1
   - Interior (cross/check): Allele 2

b) Arrows
   - Trajectories of descent between tokens
   - Dashed when merely influencing

FIGURE 4.2. An illustration of the view about cultural representations being promoted.

Sets (2.0) and (2.3) are conjectured to be empty because the assumed position is materialistic: there exists no mind beyond its physical embodiments. The only allowance beyond World 1 is how (3.0) denotes the Design Space because although the latter is but a theorists' fiction, its postulated Vastness would be impossible to be contained physically (Dennett, 1995, pp.107–113, 143–144). This is approximately the role Popper envisions for World 3 (Popper, 1972/1979, pp.115–119; Popper and Eccles, 1977/1983, pp.41–43). Set (1.0) only contains a single example token but one should remember that this set contains all the non-representative physical states of affairs. Thus, while its instantiations are of little interest within the confines of this study, this set is far from insignifact.

The tokens figure 4.2 depicts are all singular but gerenalizations about them are produced by classifying those representations according to their different colors and designs. The trajectory which originates from object (a*), for example, only contains round tokens, most of which are black. In this example, black tokens denote human features and grey tokens the characteristics of vaguely humanoid extraterrestial life forms. Circles represent the trait of rotundness and squares the trait of robustness. Finally, crosses imply gullibility while the diagonal crosses called *checks* represent cynicism.[6] Each pair consists

6 While the chosen traits are essentially arbitrary, each does contain a simple mnemonic. A certain popular type of humanoid alien often encountered in fiction is called a *grey*. Round and square shapes resemble simplified visualizations of corpulent and muscular body types. Finally, cynicism may be expressed as a sort of generalized denial, and dsagreement is often expressed using diagonal crosses.

of traits which exclude one another, and they thus form clear meme alleles (Distin, 2005, pp.169–170). Likewise, any combination of three unrelated traits is viable here.

The trajectory which connects varieties of (A) is by far the most detailed example and contains most of the inherent features of the promoted view. Its origin, (a*), is supposedly an overweight human cynic. The notation is different because (a*) is a being to whom descriptions are attached when it becomes represented. (A1) is a private mental conception someone else forms about (a*). The formulation ∃(a*):H(a*)∧[F(a*)∧G(a*)][7] expresses its contents. Here, the existence of a particular entity (a*) with the traits of H, 'being human', F, 'being fat', and G, 'being gullible', is assessed. While (a*) is an unique existence, it is not the only entity which corresponds to this description.

This representation differs from the original object because it ascribes (a*) gullibility instead of cynicism. This error is corrected during the transition from (A1) to (A2). In this case, the host of both (A1) and (A2) remains the same, and the dashed arrow between them denotes this relationship. Thus, (A2) is formed because of repeated exposure to (a*). These circumstances eliminate conception (A1) but because the two share object (a*), this ensures that (A1) should influence the formation of (A2). Since the host remains the same, presenting the transition from (A1) to (A2) as a trajectory which is influenced by the ecological factor of (a*) being present is preferable. Although repeated attempts to communicate by an information transmitting agent do not initiate this correction, this situation resembles effects of teacher in Driscoll's model because repeated exposure influences the posited correction (2011, pp.310–312).

Constructing an interactor generates a new, public representation which belongs to set (1.3). This material entity, be it a sketch or soundwaves, is not mental but embodies publicly available information. While the position that each public representation is a transformed *interpretation* of the intended mental representation (Sperber, 1996, pp.33–34) is accepted, the intentional preservation of relevant traits is nevertheless asserted. The innate expressivity (Heylighen, 1993; 1997; 1999) or salience (Heylighen, 1993, p.4) of some traits might make them more easily expessible. Salience is almost identical with relevance in its everyday sense. Both are context-dependent and thus also rely on the communicators' capacity to satiate their audiences' interests (Ibid.; Sperber and Wilson, 1986/1995, p.270).

7 The general form in this case would be ∃(x):H(x)∧[F(x)∧G(x)] but (a*) denotes a particular element among the set of objects to which variable x applies. Although Russell's notation might also be applied (Russell, 1905), it is refused because people might not actually be aware of all the relevant details and thus not represent them. Asterisks denote particular elements instead of variables throughout this chapter.

Thus, (A3), which is the interactor of (A2), contains some of the relevant features. The example assumes that the dimensions of (1) humanity–alienism, (2) rotundness–robustness, and (3) gullibility–cynicism are sufficiently relevant to become conveyed despite any other differences. However, the taken view would assign the traits being conveyed in actual communication a fitness value (Heylighen, 1999, p.3; Heylighen and Chielens, 2009, p.12) according to the ratio in which they are expressed compared to all public interactors which relate to the same objects. For example, while some entity being an alien may be considered particularly relevant and thus always become expressed, the same being's gullibility might require a more specific context to motivate expressing this trait. Thus, contextual considerations about the ecological factors which are inherent in this information's cultural selection environment would still apply (Sperber, 1996, pp.115–118).

Finally, (A3) is interpreted by various people. These new representations are mental and thus a part of World 2 in the materialistic sense introduced above. Because they are not private in the same sense as (A1) and (A2) due to the public World 3 nature of their predecessor, these tokens are positioned into set (1.2.3). Each presented interpretation of the public representation (A3) differs from the others in some regard. While (A4) retains (a*)'s original traits, (A5) attributes (a*) brawn instead of adiposity, and (A6) denies its humanity in favor of alienism. (A6) being realized is not utterly impossible although only a very specific worldview might enable this interpretation. Nevertheless, each of these three representations still contains the trait of cynicism. Even though there exist radical transformations which would demolish any claims about the replication of representations as such, substructures of these representations can still be retained differentially. If nothing pertaining to the object of interpretation were retained, communication would have failed completely. Although possible, such a situation would require a separate explanation. The criteria discussed by Heylighen (Heylighen, 1993; 1997; 1999; Heylighen and Chielens, pp.17–20) offer a cultural Darwinist explanation pertaining to the unsustainibility of certain structures under the prevailing circumstances.

The example discussed above presents what is assumed to be a relatively normal cultural trajectory: (1) There exists an object, and it is observed by an individual who formulates a mental representation of this entity. While not all such impressions are made public, (2) expressing them generates a public interactor (Hull, 1980, pp.317–320; Hull, 2001, Chapter 1) which enables the propagation of these impressions' information content.

(3) Any further conceptions which are formulated according to a process of interpreting and metarepresenting such public expressions are equally World 2 and World 3 objects.

On the other hand, the rest of the examples depict slight variations of this formula. (B1), (B2) and (B3) constitute a trajectory which is not grounded on an observed object. Rather, they illustrate how creativity helps generate new content. The dashed arrow from (A1) to (B1) is a reminder that while this account allows creativity to surpass existing mental contents, the latter may nevertheless influence the outcome. In this case, (A1) and (B1) share the traits of (1) representing a human and (2) this person being gullible. Previous conceptions might influence the imagination simply because the new representation may have prerequisites determined by its creator, such as denoting an imaginary person. When these requisite traits already exist among held conceptions, further details may be adopted from the latter in accord with the creator's intentions.

Beyond this, the trajectory from (B1) resembles the first. A public interactor is provided in the form of (B2) and further chains of metarepresentation become stranded on intersections (1.3) and (1.2.3). However, the two remaining examples deviate from this pattern. (C1) is naught but an isolated public embodiment. It must have originally been generated by an agent capable of representing it mentally (Popper, 1972/1979, pp.155–156). But if figure 4.2 were to depict only the relationships between simultaneously existing representations,[8] how it neglects (C1)'s predecessors becomes reasonable. This situation is plausible simply because many sorts of public representations like writing can easily outlast their creators and interpreters. The likes of (C1) might become secluded (1) because they do not reach further interpreters or (2) because there no longer exist potential interpreters. The latter scenario would necessitate that there exist no people with the prior conceptions inferring the intended information would require. For example, were (C1) a statue which depicts a stalwart and vaguely humanoid shape endowed with symbolic elements such as its pose expressing gullibility, everyone being unable to recognize the object and the symbolism would abolish this statue's communicative function. Even recognizing it as a product of intentional effort might become hindered. This lack of inferable intention would in turn prevent people from realizing the cultural nature of such public representations and they would become confused with elements of set (1.0).

8 Of course, (A1) is claimed to have been replaced by (A2). Thus, such a snapshot interpretation is slightly contradictory. However, the host of (A1) and (A2) could easily retain a metarepresentation of (A1) which attributes it to their earlier self while retaining the contents of (A1) despite their displacement.

The last example, (D1), is but an unactualized design. It emphasizes the Platonic undertones of World 3 (Popper, 1972/1979, pp.122–123; Popper and Eccles, 1977/1983, pp.43–45) and the Design Space. However, the Design Space is not a real level of existence, nor the world of ideal entities. Dennett stresses how most of the hypothetical designs this abstract thinking tool postulates are unintelligible and but a tangle of incoherent yet possible combinations of the symbols used to express any information structure (Dennett, 1995, pp.108–113, 124–125). In addition to (D1), the contents of all the discussed examples are also available as elements of set (3.0), although their embodiments elsewhere are of primary interest. (D1) merely illustrates how the Design Space transcends the material because the hypothetically possible is not reducible to the actual. No ontological commitments concerning the subsistence of the possible are posited. One rationale for discussing such non-actual designs is mapping the space of possibilities which may be realized by actual designs transforming (Dennett, 1995, pp.125, 128–134, 143–144).

Mapping that space of possibilities in relation to a specified amount of traits may be done by grouping each potential element according to the traits it contains. For any given amount of traits (n), there exist 2^n possible subsets which contain these traits. Designs which contain any amount of unlisted extra traits may be unproblematically incorporated into these combinations of the specified traits because only the specified traits are relevant for this grouping. However, the negations of these traits must also be considered first. Without assuming the inclusion of these negations, there would be no contradictions between the classifications modeled after the subsets. Because the relevant question for each of these (n) traits becomes whether (1) it, (2) its negation or (3) neither is present in any given combination, K, the amount of combinations, S,[9] equals $S=3^n$ where (n) still denotes the amount of traits being studied. The *marked* of traditional structuralism (de Saussure, 1916/1959/2011) involve such distinctions, and is also how meme alleles emerge (Distin, 2005, pp.169–170).

Regardless of the amount of traits chosen, the formula $S=3^n$ yields the amount of all possible classifications which utilize those traits and every entity is included in at least

9 Misters Antti Hallamäki and Lassi Haapanen are to thank for helping solve the actual amount of viable subgroups. Mr. Hallamäki also proposed a more intricate yet allegedly equivalent formula. In that formulation, which is replicated below, (i) denotes the amount of elements in each group. It helps calculate the sum of combinations up to (n).

$$\sum_{i=0}^{n} \binom{n}{i} * 2^i$$

one of these subsets. Even if an entity possess none of the specified traits, the empty set captures it successfully. Likewise, multiple classifications may apply. For example, if object (b) has structures which correspond to the conceptions people have regarding Traits T_1 and T_2, it may be classified into eack kind, K, which incorporates one or both of these traits but not their negations, nor traits object (b) does not possess. The general form these classifications have may be illustrated by utilizing the following formal description: $\forall(x)$: $\{[(T_1(x)\underline{\vee}\neg T_1(x))\underline{\vee}(\neg(T_1(x)\wedge\neg T_1(x))]\vee[(T_2(x)\underline{\vee}\neg T_2(x))\underline{\vee}(\neg(T_2(x)\wedge\neg T_2(x))]...\vee[(T_n(x)\underline{\vee}\neg T_n(x))\underline{\vee}(\neg(T_n(x)\wedge\neg T_n(x))]\}\rightarrow K_m(x)$.[10] The exclusive disjunction which is denoted using operator $\underline{\vee}$ is distinguished from a standard disjunction because both are being used. Effectively, each object which contains a specific combination of the specified traits, T, and their negations may be classified into a kind, K. The amount of possible kinds corresponds to the amount of possible subsets which may be formulated using a certain amount of traits, (n). Thus, (m) has a maximum value equal to the amount of possible subsets: $m \leq 3^n$.

Despite such classifications being cultural constructions because the traits they depend on are arbitrary, they demonstrate how the Design Space functions. Instead of arbitrarily nominated traits, what is mapped are actual structures, and instead of a random amount, all possible dimensions of design are included. Which of these possible classifications actually become utilized depends on whether they actually become represented. Because the habitus and embodied doxa structure other contents (Bourdieu, 1972/1979, pp.164–171; Deer, 2012; Maton, 2012), genuine classifications are a part of their contents.

In summary, utilizing Popper's three worlds to structure the proposed model does not require postulating anything beyond the material to explain the existence and transmission of representations. However, the truly material reality of World 1 also contains states of affairs which embody aspects beyond the plain physical (Sperber, 1996, pp.24–25; Deutsch, 1997, pp.170–193). Admitting these mental and public representations which consist of information does not imply a self-contradictory materialism (Sperber, 1996, pp.10–12). All causality still remains within the sphere of the material despite explanations involving entities which are capable of embodying information (Deutsch, 1997, pp.20–23).

Those representations follow a certain pattern of movement in their involvement with these layers, and it results from the metarepresentative nature of cultural transmission. Primary representations of actual or imaginary objects are mental and private, and either

10 Brackets for the successive disjunctions are omitted here because the truth value of this sentence is not being examined. Effectively, all the traits are equal for present purposes.

pastel or florid (Dennett, 2000). However, those representations which become manipulated using symbols are made florid (Ibid. 20–21) and interactors extending into the public sphere may only be produced for such representations.[11] These interactors do not necessarily utilize a system of symbols like a natural language: an interactor may fulfill its propagative function as long as the intended audience is capable of comprehending the intention behind the produced ostensions (e.g. Sperber, 2000c; Goldin-Meadow, 2006; Goodwin, 2006). These extensions are public instead of mental, but they enable forming new mental representations. While each of those tokens is private, they are all members of a now public type. Instead of being primary representations of entities, these further representations are metarepresentations of the public representations of other mental representations.

While here all metarepresentations which represent publicized representations are defined *cultural* in the loose sense which concerns the relationship of World 2 and World 3 entities, Sperber's distinctions are also applicable. Such representations would be links in a Social Cognitive Causal Chain produced by a publicized personal Cognitive Causal Chain (Sperber, 2001, pp.306–309; 2006, pp.435–436). Becoming a Cultural Cognitive Causal Chain would require that the contents of these representations become stabilized (Sperber, 2001, pp.309–313; Sperber, 2006, pp.436–446). This requires a widespreadness which ensures recognition (Enfield, 2006; Levinson, 2006, p.49; Sperber, 2006, pp.443–444) and enables social corrections (Driscoll, 2011, pp.310–312; Schegloff, 2006, pp.88–89).

While Sperber's rejection of generalizations which contain the formal properties of representations (Sperber, 1996, Chapters 2 and 3) is being criticized here, he correctly emphasizes the role of actual mental and public representations. Merely his merely his de-emphasis on the role abstractions play in explaining cultural evolution (Ibid. pp.62–63) is criticized. This section only contains general outlines for the attempted synthesis because its details must be discussed at the level of mental and public representations. However, the presented formal descriptions also apply to those kinds, and construing typologies helps illustrate the distribution of different traits via comparing their prevalence among various conceptions. The next topics are (1) how these traits become embodied in concrete representations and (2) how those tokens interact with actors and each other to produce the conditions of their cultural selection environments.

11 Distin criticizes Dennett for having mystified the origins and embodiment of memes (Distin, 2005, pp.78–82). This account agrees with her view that minds are the original producers of representations instead of those contents preceding mental production by being embodied in the elements of set (1.0) (Ibid. pp.81–82). Only cognized products are allowed into the subsets with information structures (Audi, 1999, p.435).

4.3 Mental representations

Mental representations are the locus of the replication of culture. Both (1) private, primary representations and (2) metarepresentations of other public and mental representations are included in this category. Next, what must be tackled because of the allegedly materialistic position is how the provided formal descriptions might be instantiated in the brain (e.g. Edelman, 1987b; Delius, 1991; Dennett, 1991; Cosmides and Tooby, 2000). Other topics include the level of experienced certainty (e.g. Peirce, 1955, Chapters 2–4; Wittgenstein, 1969) which, alongside referential opacity, allows people to retain even contradictory representations (Distin, 2005).

Delius notes how meme theorists have been forced to utilize trait descriptions instead of the actual logic of structuring information which brains employ (Delius, 1991, p.84). Because the relationship between that logic of distributing contents and the semantics of natural languages remains quite elusive, the trait descriptions being used here are highly formal. The reason the way semantic contents are embodied remains unclear is seemingly how individual brains are structured into unique wholes during development although they retain a common logic for storing information (Ibid. pp.82–83; Edelman, 1987b; 2006). Whatever this system is, it will not likely correspond to the taxonomies of the natural languages with which a host is proficient (Dennett, 1987, Chapter 6; 1991, pp.302–303). Consequently, the traits being researched are almost wholly arbitrary from the viewpoint of current brain research despite diverse information contents being necessary embodied by different brain structures (Delius, 1991, p.81).

Storing contents which pertain to particular objects may be approached as the elements of memories being distributed across the brain in storing structures such as (1) synapses (e.g. Milner et al, 1998, pp.453–463; Kandel, 2006; Linden, 2008, pp.126–144) and (2) ion channels (Ibid.). In particular, a memory trace is originally imprinted in the hippocampal regions but after some consolidation, the information about specific modalities seems to be stored in the cortical areas which normally process the appropriate types of contents (Ibid. pp.112–116). This highlights the dispersal of contents in the brain because despite this diffusion, contents pertaining to particular objects are recalled quite reliably. While brain science might condemn such simplifications, it will be surmised that the physical embodiments in the brain resemble the format of the used formal descriptions

sufficiently for the objectives of this study. While the available information might not support this interpretation, current knowledge about brain structures would at least seem compatible with such a logic of distributing mental contents.

Whether the hippocampal regions contain a connecting reference point for associations which pertain to certain objects in (Eichenbaum, 2004) or the networks neural connectedness constitutes form substructures which facilitate associative processes (Edelman, 1987b, Chapter 9), the provided formal descriptions should be applicable. In particular, the descriptions of individuals and types may both be construed using similar trait taxonomies. Neither approach differentatiates between the particular structures which are designated to be responsible for representing individuals and archetypes.

Expressing a conception regarding a unique individual requires the quantifier for unique existences, $\exists!$, a named element such as (a*) and a list of all traits, T, being attributed to object (a*) by the conception's host: $\exists!(a^*):T_1(a^*)\wedge T_2(a^*)\wedge T_3(a^*)...\wedge T_n(a^*)$. Such a convention could also distinguish between types because they are essentially acquired ways to describe specific combinations of traits. However, such conventions also require utilizing the universal quantifier, which would imply that the people who host them have an essentialist position towards those typifications. They are posited to be of the form $\forall(x):K(x)\rightarrow(T^K_1(x)\wedge T^K_2(x)...\wedge T^K_n(x))$, where K denotes the type or kind in question and T^K applies to the traits being associated with that type.[12] Though such *folk essentialism* towards both natural and social phenomena may actually be genetically prescribed (e.g. Medin, 1989; Griffiths, 2002; Mahalingam, 2007), the disposition to think that way can also be altered (Dennett, 1995, pp.200–203). At this point, the fact that folk essentialism is widespread is what matters. However, although the simplicity and widespreadness of folk essentialism make using such notation for examples desirable, less essentialistic ways to represent also exist. For example, family resemblences (Wittgenstein, 1953/2001/2009b, pp.67–77) are an example of such a systen for classifying phenomena. Different styles may be at odds but nothing prevents people from utilizing multiple such strategies.

Because actual essentialism is eschewed, both the types and the traits being represented mentally are arbitrary to an extent. These typifications are often inherited, either

12 Here, T^K may represent either the presence or absence of some feature, and they are the associations K engenders. The relationship of this habit of representing with the general form of typifications discussed in section 4.2 is noteworthy because they seemingly form a tautology together. However, T^K are not technically identifiable with the traits discussed earlier because the former are cognitively inferred associations and thus not part of a strictly formal definition.

culturally or as biological dispositions, and they structure the very experience of reality. The representations which correspond to these classifications are contained in World 1 as the material instantiation of their correlates in the brain, but World 1 itself does not submit itself to such circumscribing monikers. Furthermore, these typifications seem to be represented primarily pastelly (Dennett, 2000). Bourdieu's assessment of the doxa would seemingly capture the core of how the existence of such shared classifications actually affects people's worldviews and practices (Bourdieu, 1972/1979, pp.164–171; Bourdieu and Wacquant, 1992, pp.73–74). The folk essentialism such apparently self-evident typifications embody is also observed by Sperber when he contemplates the human capacity to produce and process abstract representations (Sperber, 2000c, pp.127–129, 133–136).

The formalizations enable examining the distributions of different traits among those typifications which pertain to specified objects. The proportion to which some traits are included among different typifications may also be of interest when universal human dispositions are being studied. They help discover psychological attractors which have emerged during the course of human evolution (Sperber, 1996, p.113). Of course, an extremely important ecological factor which affects this situation is whether or not any number of objects included in a type actually contain the traits associated with it. Some circumstances, especially memory lapses caused by (1) misattribution (Schachter, 2002, Chapter 4), (2) suggestability (Ibid. Chapter 5) or (3) biased memory (Ibid. Chapter 6), may nevertheless facilitate the inclusion of false traits. Schachter even argues that the human memory is surprisingly error-prone although most people likely do not realize the extent of their disposition to remember thing incorrectly (Ibid. pp.3–5). Nevertheless, the truly interesting cycles of selection for traits begin at the level of metarepresentations and fiction. At that point, factors like resemblance with the actual object often become much less prevalent and biased misunderstandings start to produce mutations.

When people form primary representations, it would seem that the stance being utilized affects the information their representations will include. This choice is affected by at least (1) conventions, (2) evolutionary dispositions and (3) the context. For example, the intentional stance is assumed naturally towards other people and many other organisms (Dennett, 1989, pp.16–35; Baron-Cohen, 1995, pp.32–38). Conventions include dominant experience-structuring conceptions such as widespread typifications – whether or not they are actually doxastic or merely contextually overpower the alternatives. Lastly, the context

consists of multiple factors such as the actors' concurrent needs and the social expectations affecting them. While geologists might normally assume the physical stance (Dennett, 1989, p.16) towards the rocks they are studying because their research pertains to the physical properties of those objects, they might instead assume the design stance (Ibid. pp.16–17) if a beast suddenly ambushes them. Instead of the composition of a rock being considered, the sharpness of its edges might assume priority when the hapless geologist scrabbles for potential weapons with which to defend their person.

If some traits are bound to different stances, they may only become manifest when those stances are assumed towards the representations of objects which contain those traits. A trait may become bound to a stance simply because traits are all linguistic distinctions and because different stances are linked to different language games (Wittgenstein, 1953/2001/2009b). The molecular constitution of an object such as it being iron may hardly be translated to the language of design unless iron objects are assigned specific functions. Likewise, mentioning how an object has a bladed form might seem like a description of its physical qualities but blades always exist for some purpose. If the object is attributed bloodthirstiness, none would assume that its physical properties are being discussed. All these traits, (1) 'being iron', (2) 'being bladed', and (3) 'being bloodthirsty', belong to different language games but each applies to a bladed iron object being handled by a potentially murderous person.

The contextual factors which affect the assumed stance are highly similar to those which determine the experienced relevance of different interpretations of the communicative ostensions intentional agents produce. Thus, while the entity being scrutinized often possesses no intentionality, the interpreting agent might utilize the same capacities as when inferring the meaning of ostensive communication. This conjecture amounts to an extension of the relevance theory of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012). What the agents does is (1) discerning their current context, (2) realizing the level of analysis appropriate for that context, and (3) utilizing available strategies to observe the features appropriate to that level of analysis. While this list may seem overtly intellectualizing, the actual processing need not be as elaborate – let alone conscious (Augusto, 2010; 2013a; 2013b). Changes between levels are likely controlled genetically in adaptively significant contexts such as when one is ambushed. Folk psychology (Dennett, 1987, pp.50–57) is another such example.

The second part, realizing the appropriate level of analysis, is particularly intellectualizing but it merely means choosing a stance. Dennett emphasizes how different stances are appropriate for different objects and in different contexts (Dennett, 1989, pp.16–18). He even explicitly states what is being claimed here as well: 'what we come to know, normally, are only the *relevant* truths our sensory histories avail us.' (Ibid. p.18). The assumed stance guides our perception towards certain features of the objects we encounter. When one is being assaulted, what becomes relevant are those designs which enable weaponizing an object such as it being easily graspable, or which might facilitate escape such as the observed ease of climbing some tree.

Consequently, there should be a module responsible for such inferences and it would likely precede the capacities interpersonal communication requires if metarepresentational capacities truly predate linguistic communication (Sperber, 2000c, pp.121–127). However, such capacities may be but an implementation of a more general ability to formulate representations. Both capabilities are seemingly guided by similar considerations of contextual relevance but metarepresenting also requires additional capacities such as detecting intentionality and representing the minds of others (Baron-Cohen, 1995). Baron-Cohen lists four distinct mental capabilities which enable modern interpersonal communication: (1) the Intentionality Detector (ID), (2) the Eye-Direction Detector (EDD), (3) the Shared-Attention Mechanism (SAM), and (4) the Theory-of-Mind Mechanism (ToMM) (Ibid. Chapter 4). The conjectured brain mechanism responsible for the capacity to infer the relevant traits of objects and actions may not resemble these mechanisms but they likely contribute to it. Particularly the EDD and the SAM should facilitate discovering which features of objects other people who are present deem relevant. However, the sufficiency of these four mechanisms in actually inferring which features are relevant seems questionable. They merely draw attention to certain ostensive stimuli – eye-direction in particular – which further orient people towards specific features of the world. While discerning another's intentions using the ID and ToMM may easen inferring what that person perceives as relevant, those two mechanisms are dependent on the proximity of other individuals. The general capacity to detect relevant features cannot be constituted by such limited applications and thus likely precedes them.[13]

13 The alternative that mental simulations about how others would perceive an object are primary resembles the claim that talking to oneself simulates the speech of others to activate certain neural networks (Dennett, 1991, pp.195–199, 301–303). However, a prior capacity to form primary representations would still be required, and it would necessitate the ability to discern the relevant features of observed objects.

Metarepresentations seemingly pose less problems in this regard than primary representations because they fit the Sperberian framework unproblematically, and Sperber does indeed stress their importance (e.g. Sperber and Wilson, 1986/1995; Sperber, 1996; 2000c; 2001; 2006; Wilson and Sperber, 2004; 2012). The framework Baron-Cohen conjectures (Baron-Cohen, 1995) also prioritizes successfully representing what others already represent. What truly differentiates metarepresentations from primary representations is how the former also depend on which of their parts the hosts of the re-represented primary representations express. This addition increases the amount of levels of selection (Heylighen, 1999; Heylighen and Chielens, 2009, pp.10–13), and the forces of elimination and mutation become much more prevalent and significant when an object becomes less and less available to be observed directly.

The referential opacity this breeds causes a huge disparity between primary representations and metarepresentations. Including propositional attitudes (Dennett, 1989, pp.174–202; Distin, 2005, pp.169–170) begins affecting the truth-values of the contents of representations. Thus, the contents of metarepresentations which reference the propositional attitudes carried by the hosts of the re-represented conceptions are strictly speaking different from the contents of those beliefs. However, the trait-based account this study promotes should provide fruitful insights into this matter.

This discussion requires introducing the knowledge and belief operators epistemic logic utilizes (Hintikka, 1962/2005). Because the subject is the conceptions people hold about the beliefs of others, the knowledge operator merely denotes the *assumption* that someone knows. K_p and B_p denote that a person (p) either knows, K, or believes, B, the specified contents. For example, $K_p(A)$ implies that the representing subject accepts the contents of (A) as knowledge held by (p) and thus as true. Here, (A) represents all those contents to which K_p applies. Likewise, $B_p[K_p(A)]$ would express person (p) being represented as believing that they know, and the representing subject would not commit themselves to the truthfulness of the contents of (A).

Thus, when the re-represented belief of a person (p) that some object (a*) is a member of a kind, K, which entails traits T^K_1 and T^K_2 and also a third trait of interest, T_3, is formalized, the result would be $B_p\{[\forall(x):K(x)\rightarrow(T^K_1(x)\wedge T^K_2(x))]\wedge(\exists(a^*):K(a^*)\wedge T_3(a^*))\}$. This formalization expresses that person (p) believes both (1) the general rule for typifying that kind K engenders traits T^K_1 and T^K_2, and (2) that object (a*) both (I) belongs to the

specified type and (II) possesses trait T_3. The contents of such a formula might seem overtly complicated, yet they are actually quite simple as the following example should illustrate: ”(I am under the assumption that) Kim believes that all philosophers are eccentric and bearded, and that Sokrates is both a philosopher and dead.” The falsity of Kim's belief is obvious but it could nevertheless be held by someone unacquainted with the likes of Nietzsche and Arendt. To be precise, that person *could* host representations of Nietzsche and Arendt but those representations would not include all the relevant facts. They would either not include the fact that (1) neither Nietzsche, nor Arendt had a beard, or (2) these people would not be represented as elements of the set of philosophers. Alternatively, the representations of that person might contain false information such as the claim that both Nietzsche and Arendt have beards. In Arendt's case, such a false supposition might also imply a misunderstanding regarding her gender.

Being aware of one's own conceptions also poses a problem, as noted by Hintikka (1962/2005, pp.41–45). Perfect self-awareness is implausible because it would lead to (1) an infinite regress of metarepresentation and (2) to being able to infer the contents of what one does not know (Ibid.). However, the reality that people floridly re-represent their own conceptions requires a form of self-awareness since those representations are available for mental manipulation (Dennett, 2000). Formalizing such references to oneself is achieved by simply distinguishing a particular element of the set of represented persons. This person, (i), is the individual a host identifies with their self. Thus, B_i and K_i would mean 'I believe (that)' and 'I know (that)', respectively. Becoming aware of some of one's own beliefs generates the corresponding metarepresentations which these operators describe. Similarly to how beliefs about the non-existence of flamingoes in the moon can become manifest despite normally remaining unconscious (Sperber, 1996, pp.86–87), conscious self-reflection allows re-representing any belief the reflecting agent hosts. While this also concerns the consciously processed metarepresentations themselves, a vicious regress is avoided simply because those further links in this chain are inferable but only emerge when actually inferred (Ibid.).

Thus, those conceptions the representing subject (i) hosts without ever having become aware of them remain pastel indefinitely (Dennett, 2000). The example utilized above, $B_p\{[\forall(x):K(x)\rightarrow(T^K{}_1(x)\wedge T^K{}_2(x))]\wedge(\exists(a):K(a)\wedge T_3(a))\}$, would be one such conception. In contrast, $K_i\{B_p\{[\forall(x):K(x)\rightarrow(T^K{}_1(x)\wedge T^K{}_2(x))]\wedge(\exists(a):K(a)\wedge T_3(a))\}\}$, which that

same subject might form after having become aware of their former conception, would be florid (Ibid.). Further levels of such metarepresentations may be cognized, but any given person doing so is far from necessary. This lack of self-awareness concerning many of one's own beliefs is what enables the doxa to emerge and thrive (Bourdieu, 1972/1979, pp.164–171; Deer, 2012). Particularly the folk essentialist (Griffiths, 2002; Mahalingam, 2007) conceptions people regularly utilize would seem potentially doxastic because people are predisposed structure their experiences by using such distinctions (Ibid.; Bourdieu, 1972/1979, pp.164–165; Medin, 1989; Deer, 2012, pp.114–117).

The most important feature of such metarepresentations is how the quoted contents are retained despite their referential opacity. Although neologisms are the worst vice of both contemporary philosophers and meme theorists, this phenomenon shall be dubbed *epimemetic content*. *Epigenetic* describes inheritable changes in gene expression despite the actual genes remaining the same (Taylor, 2012, pp.213–224). Accordingly, *epimemetic* would mean inheritable additions which do not alter the actual memetic code. Metarepresenting is concerned because the layers of attributed propositional attitudes do not alter the contents themselves. Acknowledging that another person believes in a certain proposition also implies realizing the contents of that proposition to the extent that they are being represented as the belief of another. In effect, the primary representation a metarepresentation re-represents may be extracted. If a person knows about the beliefs of a Feenomanist (Dennett, 1991, pp.82–85), the contents of those beliefs may be propagated as such, without referring to any particular Feenomanist's conceptions. When asked what do they know about Feenoman, this person could reference the beliefs of another to formulate an answer along the following lines: 'Feenoman, huh? Well I know someone who believes in that forest god. To my understanding, this supposed deity can heal the sick.'

Such an answer would still refer to information about Feenoman as something ulterior, as the beliefs of another person. Indeed, the person in this example does not believe in Feenoman and thus would not attribute actual godliness and supernatural healing capabilities to what they probably consider a figment of the imagination. Choosing an expression is a wholly pragmatic matter, and what is communicated is that while the speaker has a conception about Feenoman, they do not commit themselves to it by believing in its contents. If the speaker could reliably presume that their audience will not misidentify them as a Feenomanist despite their choice of words, they could merely refe-

rence their conception: 'Oh, Feenoman? He is a forest god with the power to heal the sick.' Indeed, the relevance theory (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) would predict as much.[14]

In both cases, what is communicated includes sufficient cues to reconstruct a conception of Feenoman which includes his divinity and healing capabilities. This representation may contain epistemic operators which attribute belief in those contents to some individual. The conceptions about the type of 'godly beings' the host of this new metarepresentation already has will affect how they conceive details regarding Feenoman and these extra traits may become disseminated when the host's actions provide ostensions which imply their existence. Because the conception regarding Feenoman has indeed been reconstructed as something quite different from what was originally communicated (Sperber, 1996; 2000a), the propagation of different traits and typifications is being suggested as the real object of research for studies on cultural evolution.

These conjectures are not completely unfounded because firstly, as Distin argues, propositional attitudes are applicable to memes and thus, all the memes a person hosts do not necessarily contribute to their set of personal beliefs (Distin, 2005, pp.169–170). Secondly, it would seem possible that the neural basis of metarepresentations contains *source tags* which are comparable to the operators being addressed (Tooby and Cosmides, 2000). Storing these representations would consist of distributing their contents, including the source of the representation being metarepresented between areas of the brain. Those contents could then also be accessed separately and addressing all details such as the source would become unnecessary. This picture seems to fit what was said about the possible neuronal basis of representations earlier (e.g. Edelman, 1987b; 2006; Milner et al, 1998; Eichenbaum, 2004). Hypothetically, another epimemetic factor would be the information pertaining to appropriate contexts of application – the proper fields – because that information is applied to instructions which could regardless produce the corresponding action patterns.

Distin addresses this question about propositional attitudes by distinguishing (1) active memes (Distin, 2005, p.89–91), and (2) passive or recessive memes (Ibid. pp.44–45, 58–59). Meme tokens are active when their hosts are motivated to share them (Ibid. p.90),

14 Incidentally, this would also seem to answer Moore's problem of 'saying and disbelieving' (Hintikka, 1962/2005, pp.9, 50–54). The contents of a metarepresented belief of another may simply be uttered without committing oneself to them beyond that. Hintikka's own solution is very similar and includes the same distinction between referring to one's own conceptions and those of another (Ibid. pp.52–54).

and having integrated the contents into one's own worldview increases this level of motivation considerably (Distin, 2005, pp.44–45). The term 'passive meme' is slightly misleading because what Distin means are the retained information contents in which the host does not personally believe (Ibid.). 'Recessive meme' conveys this idea better because the passivity being discussed does not imply infertility. Rather, it signifies how the host of a passively retained meme is unmotivated to actively spread it by expressing its contents (Ibid.). Distin utilizes what she calls the '"flat earth" meme' to demonstrate this point (Ibid. p.44): While most appropriately schooled people are aware of the existence of the conception that 'the Earth is flat', most believe this proposition to be patently false. However, they do retain this information as a metarepresentation which contains the affixed epistemic operators which attribute this belief to others (Ibid. pp.44–45).

This difference between active and recessive conceptions pertains to different levels of experienced certainty. What is accepted as sufficiently certain becomes affixed into the set of personal beliefs – a personal worldview. Depending on the actual level of certainty involved, the person will either consider the information knowledge or a 'just a belief' although even the latter involves some commitment.[15] Confirming observations and fine argumentation obviously increase this level of certainty but other factors also affect it. Hence, information may become *represented* as knowledge regardless of its objective epistemic status, and these representations are the subject of this current inquiry. The epistemic significance of subjective certainty has been previously addressed by Wittgenstein (1969) and Peirce (1955, Chapters 2–4) among others.

The minimum level of certainty when discussing integration should be called *acceptance* (Grice, 2001). All personal beliefs are accepted in the sense that their hosts judge them to be either desirable or plausible (Ibid. 44–50). However, qualitatively more powerful forms of certainty also exist. Wittgenstein illustrates this beautifully with his river metaphor (Wittgenstein, 1969, p.337). Clauses §96–99 of *On Certainty* (Wittgenstein, 1969) depict how some accepted propositions structure the rest like a solid riverbed guides the flow of a river's waters (Ibid. p.337). These structuring beliefs are resilient enough to maintain their position despite the continuous input of information, depicted as the rushing

15 However, this only applies to florid representations – those which are accepted consciously (Dennett, 2000). Much of what people assume remains unconscious (Augusto, 2010; 2013a; 2013b), including their doxastic presuppositions (Bourdieu, 1972/1979, pp.164–171; Deer, 2012). Having a disposition to react in some particular fashion does not require being aware of the reason for those actions. The halo effect Augosto discusses (Augusto, 2013a, pp. 658–659; 2013b, p.22) is a good example of this.

waters of a river. Under standard circumstances, there exists some erosion which may shift the river's course ever so slightly over time. However, it is also possible that the input overwhelms these boundaries and reforms the cognitive landscape. Likewise, the structuring basis might become reorganized suddenly, although this would likely still demand an input which the existing system of beliefs cannot accommodate (Wittgenstein, 1969, p.337). Exactly these cases motivate harnessing either selectionism or the Jump model which was introduced in section 3.3 as an explanation, although leaps in the design of held representations need not always be this overwhelming.

Similarly, Peirce discusses the guiding principles of thinking – habitual ways of drawing inferences (Peirce, 1955, pp.8–9). Both these principles and the riverbed Wittgenstein describes (Wittgenstein, 1969, p.337) may be identified with the habitus of an actor (Bourdieu, 1972/1977, p.214), and such actors are not limited to human subjects but also include social aggregates (Maton, 2012, pp.52–54). Peirce also acknowledges that the development of certainty is an interpersonal matter which is highly affected by the lack of alternatives among peers as well as the communally accepted sources of belief (Peirce, 1955, pp.12–22). Thus, how such certainty parallels Bourdieu's exposition of the doxa (Bourdieu, 1972/1979, pp.164–171) and the conformist bias (Boyd and Richerson, 1985, Chapter 7; Henrich and Boyd, 1998) should be obvious because they all require a communal consensus. Nevertheless, individual people are the locus of discussion here.

The fixation of beliefs consists of the assimilation of information and particularly its retention (Heylighen, 1999; Heylighen and Chielens, 2009, pp.10–13). Various propositional attitudes which are linked with different levels of experienced certainty may be classified as different styles of retention. A conviction – a belief with an extremely high certainty value – is retained much better than a whim because its high experienced certainty makes it a hinge which supports other accepted beliefs and around which they revolve (Wittgenstein, 1969, pp.346–347)[16]. It is probably also expressed more likely and more often, if Distin is correct (Distin, 2005, pp.89–91).

In summary, mental representations consist of a variety of elements, many of which are epimemetic and thus do not affect the propositional content of those representations. Despite the lack of supporting evidence, the presented account should at

16 This refers specifically to clause §152 of *On Certainty* (Wittgenstein, 1969). While most of Wittgenstein's works are constituted by such distinct clauses which specify the intended portions better than page numbers, page numbers are used in-textto maintain a uniform style.

the very least be compatible with current neuroscientific information regarding the way representations are stored in the brain (e.g. Edelman, 1987b; 2006; Milner et al, 1998; Tooby and Cosmides, 2000; Eichenbaum, 2004). It is being claimed that (1) representations consist of associative descriptions which contain (I) typifications and (II) traits, (2) primary representations are generated by a relevance-based search guided by the host's stance towards their objects, and (3) secondary representations contain epistemic operators which include the represented propositional attitudes of their sources. Additionally, the contents of such metarepresentations may be detached from such source tags (Tooby and Cosmides, 2000) and expressed individually. Whether or not this is actually done depends on the relevance of expressing the sources (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) as well as the propositional attitude the host personally has towards the contents (Distin, 2005, pp.44–45, 58–59, 89–91). This personal propositional attitude manifests as a level of assumed certainty, although even the representations the person holds most certain need not be accessed consciously (Bourdieu, 1972/1979, pp.164–171; Augusto, 2010; 2013a; 2013b).

The proposed formal presentations of mental representations primarily describe what could be dubbed *folk essentialism* (Griffiths, 2002; Mahalingam, 2007), a naturally occurring way to structure experiences (Ibid.; Medin, 1989). Although sustaining such a naive essentialist approach is not a genetically determined necessity (Dennett, 1995, pp.200–203), escaping its influence may prove surprisingly challenging (Ibid. p.39). Alternative approaches include family resemblances (Wittgenstein, 1953/2001/2009b, pp.67–77) and population thinking (Dennett, 1995, p.193; Boyd and Richerson, 2000). Whatever approach is adopted towards particular phenomena, that manner of structuring consists of cultural information itself (Bourdieu, 1972/1979, pp.164–171; 1979/1984/2010; Deer, 2012; Maton, 2012) and is thus subject to selection. Likewise, manners of gaining information are evaluated according to the standards of a community (Peirce, 1955, pp.12–22). All these factors affect the fitness of particular conceptions.

The next section addresses the interactors (Hull, 1980; 2001) hosting these mental representations may engender. While this section is primarily concerned with the stages of assimilation and retention, the dynamics of expression and transmission will become elaborated next (Heylighen, 1999; Heylighen and Chielens, pp.10–13). The influence of Sperber is emphasized when what exactly becomes expressed and why is being discussed.

4.4 Public representations

Earlier, public representations have also been dubbed public productions because it is being argued that they are the interactors (Hull, 1980; 2001) of mentally stored information packages. These interactors produce the circumstances which enable the replication of the information which is responsible for these extensions (Hull, 1980, pp.318–320; 2001, pp.22–24). The transmission of cultural information may indeed consist of frequent transformations when new representations are construed according to ostensive cues (Sperber, 1996; 2000a). However, while the contents of the representations may not remain identical, the patterns for structuring experience they contain can still be transmitted.

What is examined in this section are types of public representations and how exactly do they affect the mental representations of receivers. These topics involve (1) minding one's audience, including manifest conceptions about their prior beliefs, (Sperber and Wilson, 1986/1995; Grice, 1991; Wilson and Sperber, 2004; 2012) and (2) an elaboration of the reason the transformations Sperber emphasizes (Sperber, 1996, pp.100–106; 2000a) do not falsify cultural selectionism. Thus, what is examined are the consequences of shifting the focus of research from representations as such towards their constituting traits. However, these representations themselves still frame research because the distribution of trait descriptions is of little interest unless it is proportioned in relation to a distinct set. Nevertheless, comparisons between types of representations are still quite possible and may even reveal predispositions humans have towards certain types of information. At both levels, the effect of convergent evolution are prioritized above exact lineage-tracking (Dennett, 1995, pp.355–357).

The first issue is whether or not an interactor needs to be unambiguous and thus strictly encoded. The second question is if it should embody all the information its replicator contains. The relevance theory of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) describes how successful communication is possible without information having been encoded, and thus an affirmative answer to either question should be refused. However, the existence of encoded messages is not denied because this framework can also assimilate such cases (Sperber and Wilson, 1986/1995, p.182). Likewise, the theory should enable explaining the replication of thinking patterns and the contents of representations even when such information is not fully explicated.

These two problems are intertwined because they both pertain to the explication of intended contents. The key to answering them is how people expect certain prior conceptions to be hosted by their audiences (Sperber and Wilson, 1986/1995, pp.39–42). Communicators have these manifest metarepresentations which are based either on prior interaction with the audience or derived from their concurrent behavior because of the intentional stance (Dennett, 1989, pp.17–22). They successfully enable the communicator to merely supplement or connect the audience's prior conceptions in order to communicate the intended contents.

Thus, the contents already being assumed to be accessible to the audience need not be conveyed. Despite the redundancy of doing so, the assumed prior information may be expressed because saying too much is a more effective communication strategy than saying too little. The experienced certainty of those metarepresented beliefs being accurate might also affect communication because any doubt pertaining to whether or not a given conception is held by the audience might motivate the communicator to express those contents in order to secure successful communication. Of course, the relevance theory emphasizes the optimality of contributions (Sperber and Wilson, 1986/1995, p.270; Wilson and Sperber, 2004, p.612; 2012, p.6) for a reason. Within that framework, expressing too much may also cause a misunderstanding when the reason the communicator repeats what the audience already knows is not manifest to the members of that audience. The communicator being uncertain about their audience's previous beliefs would be one plausible explanation but receivers might also construe other explanations which would affect what they infer. For example, the audience might conclude that their intelligence is being insulted (intentionally or not), or they might produce intricate explanations about why what is being said actually contains a more elaborate, previously unknown meaning.

Hopefully, the following example clarifies this account. The communicator, person (c), is attempting to persuade an audience which consists of person (a) about the truth of conception $(\exists!(o^*):[K_P(o^*)\wedge K_G(o^*)]\wedge B(o^*))$ which pertains to object (o*). What is being claimed is that the unique object (o*) is a member of the sets of 'professors', K_P, and 'geniuses', K_G. Additionally, (o*) is being attributed trait B, 'baldness'. Hence, what (c) is trying to communicate to (a) is that 'There is this genius professor, Dr. O, who is also bald.'

The question is, how will (c) express this sentiment to (a), and answering it will require further specifications. The example quote above is a realistic utterance which

would convey what is being communicated. However, using it to successfully transmit all the traits being attributed to (o*) by (c) would demand that what both (a) and (c) associate with the classes of 'professors' and 'geniuses' is identical. The differences in people's typifications are a major cause of transformations in the exact contents of transmitted representations. What a typification such as K_P expresses is a generalization. Despite the names to which these typifications correspond being the same across people and thus expressions such as 'professor' being understandable, the actual inferences being made may differ radically because the contents of the represented kinds differ between people. In effect, what is required is common ground (Enfield, 2006) beyond the expressions being used. The doxa (Bourdieu, 1972/1979, pp.164–171; Deer, 2012) or otherwise shared parts of the habitus (Bourdieu, 1979/1984/2010; Maton, 2012) are major sources of such classifications which are shared by many.

The example utterance would thus identify (o*) as a professor and a genius but it would not transmit the intented content linked to these typifications. Let (c)'s conception about professors be expressed thus: $\forall(x):K^{C}_{P}(x)\rightarrow(I(x)\wedge E(u,x))$. In this representation of the set of professors which (c) hosts, K^{C}_{P}, all elements have the quality of intelligence, I, and their relation to some university, (u), is being its employee, E. All this is included in what (c) is attempting to express about (o*) to (a). However, (a)'s conceptions about professors might not be identical. Instead, let (a) be under the following assumptions: $\forall(x):K^{A}_{P}(x)\rightarrow(O(x)\wedge E(u,x))$, and $\exists(x):(K^{A}_{P}(x)\wedge I(x))$. Thus, (a) believes that all professors are old, O, and employed, E, by an university, (u), and also accepts that there exist some intelligent, I, professors although their stereotype about professors does not include this.

Now, the example expression 'There is this genius professor, Dr. O, who is also bald', if uttered by (c), would not necessarily evoke the intended associations. Presumably, both attribute intelligence to geniuses. Thus, that part of (c)'s conception about professors would be transmitted because such overlap generates degeneracy (Edelman and Gally, 2001; Edelman, 2006, p. 33–34). Because (a) already accepts that a professor may be intelligent, there is no contradiction with previously held beliefs. However, a further problem does arise since (a) presupposes that all professors are old. While (c) does not intend to communicate this association, (a)'s representation of (o*) would include it unless further specifications are made by (c). Indeed, the fact that (a) could at some level associate baldness with old age would reinforce the interpretation that (o) is old.

To avoid this misunderstanding, it is required that the differences between (c)'s and (a)'s conceptions regarding professors are manifest to (c) or (a) or both. If (c) realizes the details concerning (a)'s beliefs about professors and thus generates the metarepresentation $B_A[\forall(x):K^A{}_P(x)\rightarrow(O(x)\wedge E(u,x))]\wedge[\exists(x):(K^A{}_P(x)\wedge I(x))]$, (c) may then deny the claim that all professors are old and affirm the intelligence of this particular professor, (o*). Of course, (c) might not be motivated to purposefully avoid implicating that (o*) is old. Professor (o*) could indeed be old and (c) might be well aware of this despite not intending to communicate it at the time. In this case, drawing the inference that (o*) is old would likely be allowed because it would not be considered misinformation. However, if (o*) were relatively young despite their hair loss, and (c) were aware of this, the motivation to also communicate this fact might emerge. The potential misunderstanding might also be deemed harmless by (c) despite their awareness that it would contradict the facts. Only if the age of (o*) is deemed *relevant* by (c) in the present context, is it communicated.

Likewise, if (a) realizes how (c)'s conceptions about professor differ from theirs, they may manifest the metarepresentation $B_C[\forall(x):K^C{}_P(x)\rightarrow(I(x)\wedge E(u,x))]$. However, manifesting this conception is insufficient to avoid the interpretation that (o*) is old. What would be required is that (a) infers that $B_C\{\exists(x):[K^C{}_P(x)\wedge(\neg\ O(x))]\}$. This belief being attributed to (c), that there are professors who are not old, would contradict (a)'s own conception about professors, though. What determines the probability of it being conceived is how *certain* (a) considers their impression that all professors are old. This is naturally affected by many facts about (a)'s worldview, including whether or not they believe that the truth is manifest (Peirce, 1955, pp.11–13, 56–59; Popper, 1963/2002a, pp.6–11). If the supposition that all professors are old were to have been integrated in (a)'s worldview, especially because it is part of the doxa to which they are exposed, this necessary metarepresentation concerning (c)'s allowance of relatively young professors would most definitely not occur to (a). In the end, the beliefs which constitute a doxa are held absolute and universal (Bourdieu, 1972/1979, pp.164–167) and only what is held certain is likely to become part of a habitus (Bourdieu and Wacquant, 1992, pp.18–19). Thus, while the necessary information about (c)'s beliefs might become manifest to (a) without being explicated by (c), the probability of this happening would be quite low. Communicators are effectively responsible for realizing the differences between them and their audiences and for formulating an appropriate utterance to convey what they intend to communicate.

How this example answers the two questions raised above may now be summarized. The strict encoding of meaning seems unnecessary if communicating the intended traits may be achieved reliably within the framework of the relevance theory which eschews the necessity to encode meaning in the first place (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012). On the condition that the provided example is realistic in this regard, accomplishing this seems possible because the combination of (1) recognizable cues such as concepts represented by the receiver in some form and (2) the recognition of another's prior conceptions should apparently suffice for communication (Sperber, 2000c). Such ostensive stimuli and manifest conceptions are exactly what is emphasized by the relevance theory. Thus, the example would be an acceptable illustration of how communication may reliably replicate information without anything beyond what Sperber would allow being postulated.

The second issue should also be resolved by the above example. Because repeating the information already being attributed to the receiver is unnecessary, every detail about what is being communicated need not be expressed. Oftentimes, cultural common ground (Enfield, 2006) and shared doxa (Bourdieu, 1972/1979, pp.164–171; Deer, 2012) ensure that there is plenty of overlap between the prior conceptions of the communicator and the receiver. In such cases, the mutual terminology being utilized will provide the intended associations unproblematically. This also reinforces the claim that what should be researched are not representations as such but the traits they contain because Sperber correctly claims that sometimes public products actually convey no new information because the presented type is already familiar to the audience (Sperber, 2000a). While being exposed to a public interactor of a mental representation different from one's own might not eliminate the previous conception, it might easily influence and complement it, conveying the connection to some previously unassociated traits. Traditional accounts of cultural evolution would describe this situation as the occurence of a mutation but here it is naught but the successful propagation of certain associations.

There would thus seem to be multiple types of expressions which may affect the same representations. How those representations become modified pertains to (1) the associations connected to different classifications, or (2) the descriptions of specific objects. Thus, the associations a representation of an object produces are affected both by (1) the traits being attributed to that object and (2) what those traits are held to entail. For

example, when object (o*) is attributed K_P, 'being a professor', and K_G, 'being a genius', as well as B, 'baldness', the exact associations made depend on what each of these descriptions entails according to the host. The representation may also be modified by adding descriptions such as K_M, 'being a member of Mensa', and C, 'having cancer'. Associations may become interconnected if they incidentally overlap: treating cancer causes loss of hair so if (o*) is attributed both B and C, the cancer will likely be held causally responsible for their baldness unless a better explanation is available. Likewise, contradictions which become acknowledged consciously will likely cause modifications which restore consistency. These contradictions may apply to either the associations of two different typifications which are applied together to at least one object or to the associations of a typification and a further description. An example of the latter was provided above when (a) associated being a professor with being old but (o*) was postulated to (possibly) be both a professor and not old.

Thus, expressions may be construed according to the assumed differences between what the communicator and the receiver(s) associate with different objects and different descriptions. The replication of a conception like $\exists!(o^*):[K_P(o^*)\wedge K_G(o^*)]\wedge B(o^*)$ which is intended to actually convey $\exists!(o^*):[(I^+(o^*)\wedge E(u,o^*))\wedge(T(o^*)\wedge F(o^*))]\wedge B(o^*)$ is accomplishable in a multitude of ways. This degeneracy (Edelman and Gally, 2001; Edelman, 2006, p. 33–34) is important because it increases the likelihood of the intended associations being drawn despite differences between people's conceptions. Here, the extraordinary intelligence, I^+, associated with geniuses supplants merely being intelligent, I, which professorship implies because the former also implicates this fact about (o*) . The other traits being associated with geniuses are being talented, T, and being somewhat eccentric, F, while being a professor is associated with employment, E, by a university, (u). Baldness, B, is simply an additional description being affixed to (o*). Thus, the communicator, (c), would be under the following impressions about professors and geniuses $\forall(x):K_P(x)\rightarrow(I(x)\wedge E(u,x))$ and $\forall(x):K_G(x)\rightarrow[I^+(x)\wedge T(x))\wedge F(x)]$.

The interactors bearing $\exists!(o^*):[(I^+(o^*)\wedge E(u,o^*))\wedge(T(o^*)\wedge F(o^*))]\wedge B(o^*)$ are thus only required to embody those traits the audience would not be expected by the communicator to associate with object (o*). If the audience contains some person (d) who believes that all professors are geniuses, $\forall(x):(K_P(x)\rightarrow K_G(x))$, and if this becomes manifest to communicator (c) who in turn formulates the belief $B_D[\forall(x):(K_P(x)\rightarrow K_G(x))]$, (c) need not

explicate that (o) is a genius simply because the same associations are regardless conveyed to (d) by attributing (o*) professorship. However, this would still require that (d)'s conceptions of geniuses and thus professors include the intended associations of extraordinary intelligence, I+, talentedness, T, and eccentricity, F. If a metarepresentation pertaining to (d)'s exact associations with geniuses is not manifest to (c), expressing the intended traits being attributed to (o) might feel necessary. Also, (c) might be unable to conceive how (d)'s conceptions could differ in this regard, and (d) would be attributed (incorrectly or not) beliefs identical to (c)'s own. If (c) were to err in this regard, some of the intended traits would fail to be replicated. An attempt to either explicitly state the intended traits instead of relying on typifications or an attempt to explicitly convey one's own typification scheme would still function as an interactor communicator (c) could construct for the representation $\exists!(o^*):[(I^+(o^*) \wedge E(u,o^*)) \wedge (T(o^*) \wedge F(o^*))] \wedge B(o^*)$ or some of its parts. The likelihood that all associations, particularly for uncommon classes, another may form become manifest to the communicator is quite small at any rate.

A glaring flaw in what has been presented thus far is how it only applies to very specific types of representations. Even some previous examples such as Dawkins's instructions to make Chinese junks (Dawkins, 2003, pp.144–146) would hardly fit into the framework provided in this chapter because it seemingly applies only to representations which consist of quite broad descriptions. This discludes instructions which instead are normally constituted by successive directives. Representations where the details of its object's structure are crucial such as song lyrics would also hardly fit the present account. While such limitedness should remain permissible if there do indeed exist compatible explanations for different cultural items, a general outline is also sketched for those cases because it seems unlikely that how they propagate would follow an utterly different logic. At most, formalizing their contents may provide a challenge.

The level of detailedness such distinctive cases contain distinguishes these special cases from previous examples. This requirement generates the illusion that both instructions and structure-emphasizing representations fail to replicate unless their contents are copied perfectly fidelitously. However, describing instructions using a way which remains compatible with the general account should remain achievable. Attempting to tentatively formulate such descriptions yields directives like 'any person (p) may produce result (r), which contains Trait A, by executing Action B', which may be formalized along the lines of

$\forall(p)\forall(r):(A_B(p,r)\rightarrow T_A(r))$.[17] In turn, the standard way used thus far may be applied to the actions and traits such a description contains. A combination of multiple such advice would constitute the instructions to produce artifacts, practices and other cultural items with the desired traits (Dawkins, 2003, pp.144–146).

Similarly, most people hardly remember every detail about the lyrics of the songs they can nevertheless recognize. Dennett specifically notes how remembering the first four notes of Beethoven's Fifth Symphony (*Fate*) is much more prevalent than being able to replicate the rest of that symphony (Dennett, 1995, p.344). Likewise, people often remember the themes certain songs embody even when that cannot recall the exact lyrics of those songs. While most of the provoded examples of such structure-dependent representations relate to music, the same logic should apply to all parallel instances such as quotes, poetry, and blueprints. All these instances are related because they embody the requirement to utilize certain symbols accurately. Most of these examples are constituted by language, and the rest by other systems of symbols which are used to express the relations of the things the symbols they contain represent.

Such cases where only general features are remembered do not require the general theory to adapt. The actual representations which are present in those cases do indeed attribute traits such as 'being about love' and 'containing explicit language' to their objects. Nevertheless, when the exact lyrics of a song are remembered, an interesting phenomenon is illustrated. These types of detailed information may indeed be stored and recalled but such rigorous memorization seems to require a special attitude towards the objects. This attitude resembles the physical stance (Dennett, 1989, p.16) but instead of emphasizing the physical structure of objects, it accentuates the syntactic properties of entities which consist of symbols. It is thus nominated the *syntactic stance* because the design stance (Dennett, 1989, pp.16–17) would not suffice either since what is memorized are not functions per se but the actual (syntactic) structures of the objects. Forcing the representations generated this way into the earlier framework does them violence but it should be accomplishable. Calling the contents of such descriptions something besides traits might be appropriate but their formula would be 'object (o) contains substructures (α), (β), and (γ)'. The arrangement of these substructures could then be denoted by expressing their relations.

17 Details pertaining to the exact quantifiers, for example, may vary according to what is being claimed by the instructions. For example, it could be claimed that there exist some persons with certain qualities which allow always succeeding in producing the intended product. Here, what is claimed is that any person (p) who executes Action B will succeed in producing Trait A to token (r).

For example, $A_O(\alpha,\beta)$ would mean that in the case of some object (o*), substructure (α) is anterior, A, to substructure (β). These substructures could then be dissected further until the level at which people actually make no further divisions would be reached.

Because public representations are being examined, it might be appropriate to recall the distinction between *copying-the-instructions* and *copying-the-product* (Blackmore, 1999, pp.61–62, 213–215; Distin, 2005, pp.93–95) which pertains to how one interacts with a public representation to assimilare the information it embodies. This distinction also explicitly states how there exist at least two functionally different types of interactors which are capable of emboding mental representations: (1) instructions and (2) products. Ironically, these two types greatly resemble the abovementioned deviant cases, and this connection does complicate matters.

The distinction between copying the instructions and copying the product is primarily applicable to the knowledge for producing specific artifacts or performances. Although Blackmore (1999) and even Dennett (1991; 1995) would sometimes prefer to identify mental representations with nothing but such capacities, no such impoverished behaviorism is being advocated within the confines of this study. Because most features of Distin's distinction between memes and the minds which process them are accepted (Distin, 2005, pp.77–102, 168–173), the transmission of descriptions is not posited to be limited to copying the instructions and products. Instead of people learning how to produce specific utterances or how to imitate heard utterances, what has been posited is a rational process of interpreting and distilling meanings (Sperber, 1996, pp.33–41; Distin, 2005, pp.106–108). Such semanticity is hardly digestible for a simple division between copying either instructions or products, but Blackmore's distinction (Blackmore, 1999, pp.61–62, 213–215) still remains applicable to descriptions which are limited to either physical or syntactic properties. Thus, the deviant cases introduced above thus remain within this division's jurisdiction while the rest are captured better by the relevance theory.

Explaining how copying the instructions occurs requires but a relatively straightforward application of the logic for constructing representations of instructions which was introduced above. Because the ostensions available for the purpose of communicating and allowed by the relevance theory are not limited to verbal utterances (Sperber and Wilson, 1986/1995, pp.49–54; Sperber, 2000c, pp.121–127), even showing instead of telling may be accommodated. However, successfully copying the product relies on the prior availa-

bility of specific mutualist memes. Basically, specific schemas for typifying phenomena and the instructions to produce the observed features constitute this foundation which copying the product successfully requires. Thus, this process may be compared to generating primary representations. Because certain background information is required, new typifications and instructions hardly become transmitted. Instead, such products may propagate the specific combination of existing information which is necessary for reproducing them. When these products symbolize other entities, they may also convey new traits which are being attributed to those entities. For example, a bronze statue of Charles Darwin might transmit the instructions which are required for producing replicas of it to the people who already know enough about casting bronze. It might also transmit new information about Darwin to those who recognize him as being the represented object.

In summary, public representations propagate patterns of thinking but while they do not encode the contents of those mental representations as such, they may be formulated to complement or contradict beliefs the audience is held to already host. The contents of such metarepresentations may be false, but doxa (Bourdieu, 1972/1979, pp.164–171; Deer, 2012) or other beliefs which comprise common ground (Enfield, 2006) between the participants increase the reliability of communicating the intended meanings despite this fallibilism. Likewise, the connectedness of different typifications through common associations engenders degeneracy (Edelman and Gally, 2001; Edelman, 2006, p. 33–34) which further enhances the reliability of reproducing the intended associations. While such conduct still risks misrecognition, all intended traits may also be expressed explicitly when it is recognized that the audience might not draw the intended conclusions otherwise.[18]

Although the presented framework may at first seem superficially incompatible with (1) instructions and (2) memorizing the exact structure of specific products, it may be supplemented to integrate these cases successfully. Similarly, expressions and products function differently as interactors but this may be surpassed by analyzing these differences. Copying products successfully requires prior knowledge pertaining to instructions on how to achieve certain results. Either type may also have an expressive function which requires the audience to be able to interpret the utilized symbolism successfully.

18 Of course, the associations those expressions produce face the same problem but at least one such pitfall is avoided and thus the mitigating factors such as common ground (Enfield, 2006) and degeneracy (Edelman and Gally, 2001; Edelman, 2006, p. 33–34) may be applied once more. It seems likely that at some point, either the communicator must (1) accept some transformations of the intended message as irrelevant and let them occus or (2) reach a level at which both participants have the exact same associations. In practice, the first scenario seems more likely although it often occurs inconspicously.

4.5 Conclusions

This study has aimed to clarify how representations merge and how exactly do associations with certain traits propagate. The adopted approach had been demonstrating the compatibility of selectionism and certain features of Dan Sperber's epidemiology of beliefs (Sperber, 1996; 2001; 2006). The side of selectionism has been addressed primarily with the help of meme theorists (e.g. Dawkins, 1976/1989/2006b; Hull, 1982; 2001; Dennett, 1991; 1995; 2006; Blackmore, 1999; Distin, 2005; Heylighen and Chielens, 2009) despite some allusions to the Dual-Inheritance Theory (e.g. Boyd and Richerson, 1985; 2000; Henrich and Boyd, 1998; 2002; Henrich and McElreath, 2003) having also been made. Although the meme theory has glaring problems, it emphasizes an important aspect of this question: the propagation of representations is not completely under human supervision. Rather, people are equally affected by the conceptions they embrace, although they are also responsible the selection pressures which apply to those representations. In turn, Sperber offers the relevance theory of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) and a welcome critique (Sperber, 1996, pp.100–106; 2000a) of the naively replicationist views of most meme theorists. The latter motivates assuming a new and hopefully fertile perspective to selectionism.

What was provided first was an introduction to universal Darwinism and orthodox meme theory. In combination with Dan Sperber's critiques of such cultural selectionism (Sperber, 1996, pp.100–106; 2000a), they helped establish a tentative framework which could integrate Sperber's viewpoint without compromising selectionism. The basis for this account had already been provided by Kate Distin argument that selectionism is plausible despite meaning being reconstructed during communication (Distin, 2005). Her claims were elaborated by arguing that what replicates are trait descriptions which are applied to different objects and categories. Although this view already deviates from the analogy with natural selection because it de-emphasizes the importance of distinct lineages, studying the distribution of different traits strikingly resembles to the gene-centric view on evolution (Dawkins, 1976/1989/2006b; 1982/1999).

Details pertaining to Sperber's own theoretical framework were examined next in order to elaborate the the emerging theory. The relevance theory of communication (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012) was introduced first,

and it was accepted as an excellent description of human communication. The role metarepresentations play within Sperber's framework is indispensable and thus the human capacity to metarepresent was also dissected. Sperber's views on the matter were complemented using Dennett's intentional systems theory (Dennett, 1989, pp.16–35; 2009a). Finally, the attractors (Sperber, 1996; Claidière and Sperber, 2007) and causal chains of representation (Sperber, 1996; 2001; 2006) which define Sperber's epidemiology of reprentations were brought under scrutiny. It was concluded that Driscoll's attempt to integrate selectionism and epidemiologic approaches (Driscoll, 2011) provides important insights regarding the roles of (1) adaptive significance, (2) the people who distribute specific ideas, *teachers*, and (3) the distribution of attractors. She convincingly claims that these factors delimit the effects of attractors (Driscoll, 2011, pp.307–313). In contrast, it was also concluded that when representations transform, their content may undergo even greater changes than either Sperber or the analogy with evolution should allow but that these leaps are facilitated by researchable processes such as imagination. The incongruity between this *Jump model* and selection was deemed resolvable because humans are able to cull unsustainable ideas (Popper, 1978) which prevents these macromutations from filling culture with such monsters (Dennett, 1995, p.287). Thus, this second deviation from biological evolution should also prove innocuous, although the exact processes which enable such jumps need to be identified to support that conclusion (Sperber, 2001, p.298).

Before the framework could be recapitulated, it needed to be complemented with Pierre Bourdieu's reflexive sociology (e.g. Bourdieu, 1972/1979; 1979/1984/2010; Bourdieu and Wacquant, 1992) which provides further tools to analyze the dynamics of the cultural selection environments and their effects on people's minds. Concepts like (1) doxa (Bourdieu, 1972/1979, pp.164–171; Deer, 2012), (2) habitus (Bourdieu, 1979/1984/2010; Maton, 2012), (3) capitals (Bourdieu, 1979/1984/2010; Moore, 2012), and (4) fields (Bourdieu, 1979/1984/2010; 1984/1988; Thomson, 2012) contribute to this end greatly because they encapsulate the effects personal resources and communal conventions have. They thus enable a much more detailed analysis of the relationship of public and private opinions as well as the effects the context has on the assimilation of beliefs. Likewise, Sir Karl Popper's three worlds (Popper, 1972/1979; Popper and Eccles, 1977/1983) were used to categorize the ontological status of different representations. In particular, these categories allowed demonstrating how a materialist account may accommodate cultural

representations without postulating a truly separate universe for them. The relationship between mental, public and cultural representations and how the first two comprise the third could also be demonstrated using Popper's framework.

The final outline of the theory which has been drafted on these pages is constituted by the following claims: (1) The selection of culture occurs in multiple phases which include (I) assimilation, (II) retention, (III) expression, and (IV) transmission (Heylighen, 1999). (2) This selection applies to multiple types of information structures and primarily affects which traits become associated with specific (I) objects and (II) categories. (3) This whole process of selection is mediated by the rational reconstruction processes on which Sperber and Wilson claim communication to be based (Sperber and Wilson, 1986/1995; Wilson and Sperber, 2004; 2012). Considerations of experienced certainty (Peirce, 1955, Chapters 2–4; Wittgenstein, 1969) and particularly propositional attitudes (Hintikka, 1962/2005; Distin, 2005, pp.169–170) complicate this basic account further. The latter's connection to communication and metarepresentations (Sperber, 2000c) requires cumbersomely deliberating what the people who participate in communication and in the propagation of culture may actually know about each other's opinions and how their fallibilism in this regard affects the reliability of transmitting their conceptions.

Because of propositional attitudes, people may assimilate and retain even conceptions they consider untrue because those representations are inherited as metarepresentations which attribute the corresponding beliefs to others. Such contents may however become expressed even by the people who do not personally adhere to them. Oftentimes, they become expressed in a form which includes *epimemetic contents* such as a mention of the original sources (Tooby and Cosmides, 2000). However, even such cases propagate the original information in some form, and the level of (un)certainty affixed to them as an epimemetic addition may be expunged by new hosts with whose existing positions those contents are more compatible.

Likewise, repeating one's beliefs in the form they are being hosted is unnecessary because people may infer one another's positions using the intentional stance (Dennett, 1989, pp.17–22; Baron-Cohen, 1995). When they metarepresent someone else's position manifestly, it allows communicators to construe expressions which help modify the corresponding representations to match the intended interpretation in relevant aspects, and this suffices to propagate associations. Because people are capable of deriving primary

representations from the entities they encounter, even public representations which hardly convey information such as artifacts may modify existing representations in a way which may be categorized as replication. What this requires is (1) sufficient prior knowledge about the appropriate practices and systems of symbols, and (2) assuming the appropriate stance (Dennett, 1989, pp.16–35; 2009a) towards those objects. While the provided example formalizations which were used to describe the contents of representations hardly apply to all different forms of representations, the same basic logic would nevertheless seem to apply universally. Basically, while the exact contents of different descriptions may contain different elements such as directives, they may all be formulated by referencing particular objects and classes of objects in a similar fashion.

To summarize, the amalgam of cultural Darwinism and inferential communication introduced in this study claims that associations contained within representations of different concepts or objects adhere to universal Darwinism. Their propagation is unhindered by the transformations which endanger more traditional models of cultural selection (Sperber, 1996, pp.100–106; 2000a) because those transformations occur at the level of associations which allows reconceptualizing them as a form of selection. The claim that such associations replicate is based on and thus compatible with inferential models of communication (Sperber and Wilson, 1986/1995; Sperber, 2000c; Wilson and Sperber, 2004; 2012). The presented viewpoint requires considering the manifestness (Sperber and Wilson, 1986/1995, pp.39–42) of another's conceptions as well as the role played by propositional attitudes as a part of those metarepresentations.

Further research is required to clarify the exact nature of the Design Space (Dennett, 1995, Chapters 3 and 6; 2013, Chapter 38) as well as to identify all the different types of representations the human brain actually processes. Ideally, formal descriptions akin to those provided on these pages may be formulated for all those types. Actually challenging the provided framework by testing the hypotheses which may be derived from it would also be most desirable in the long run for this theory has been intended as a serious inquiry into the relationship of the mental and the cultural. Thus, its compatibility with neuroscientific research and other related fields of inquiry must be assessed because at this point, this relationship could only be evaluated rudimentarily.

5. REFERENCES:

Audi, R. ed. 1999. *The Cambridge Dictionary of Philosophy*. Second edition. Cambridge: Cambridge University Press.

Augusto, L. M., 2010. Unconscious Knowledge: a survey. *Advances in Cognitive Psychology*, Vol 6, pp. 116–141.

Augusto, L. M., 2013a: Unconscious Representations 1: Belying the Traditional Model of Human Cognition. *Axiomathes*, Vol 23, pp. 645–663.

Augusto, L. M., 2013b: Unconscious Representations 2: Towards an Integrated Cognitive Architecture. [online] *Axiomathes* (Published February 2013). Available at: <http://download.springer.com/static/pdf/949/art%253A10.1007%252Fs10516-012-9207-y.pdf?auth66=1386246015_9b070457faddcff7ab8739f44f1e8ffc&ext=.pdf> [Accessed 3 December 2013].

Aunger, R. ed. 2000. *Darwinizing Culture: The status of memetics as a science*. New York: Oxford University Press.

Baron-Cohen, S., 1995. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press.

Bendall, D. S. ed. 1983. *Evolution from molecules to men*. Cambridge: Cambridge University Press.

Blackmore, S., 1999. *The Meme Machine*. New York: Oxford University Press.

Blackmore, S., 2009. Dangerous Memes; or, What the Pandorans Let Loose. In: Dick, S. J. and Lupisella, M. L. eds. 2009. *Cosmos & Culture: Cultural Evolution in a Cosmic Context*. [online] Available at: <http://www.nasa.gov/pdf/607104main_CosmosCulture-ebook.pdf> [Accessed 24 January 2014]. Ch.7.

Bloch, M., 2000. A well-disposed social anthropologists's problems with memes. In: Aunger, R. ed. 2000. *Darwinizing Culture: The status of memetics as a science*. New York: Oxford University Press. Ch.10.

Bourdieu, P., 1977. *Outline of a Theory of Practice*. Translated from French by R. Nice. Cambridge: Cambridge University Press.

Bourdieu, P., 1988. *Homo Academicus*. Translated from French by P. Collier. Stanford: Stanford University Press.

Bourdieu, P., 1991. *Language and Symbolic Power.* Translated from French by G. Raymond and M. Adamson. Cambridge: Polity Press.

Bourdieu, P., 1994. *In Other Words: Essays Towards a Reflexive Sociology*. Translated from French by M. Adamson. Cambridge: Polity Press.

Bourdieu, P., 2010. *Distinction: A Social Critique of the Judgement of Taste*. Translated from French by R. Nice. Routledge Classics edition. London: Routledge.

Bourdieu, P. and Wacquant, L. J. D., 1992. *An Invitation to Reflexive Sociology.* Cambridge: Polity Press.

Bourdieu, P. and Passeron, J.-C., 1990. *Reproduction in Education, Society and Culture*. Second Edition. Translated from French by R. Nice. London: Sage Publications.

Boyd, R. and Richerson, P. J., 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Boyd, R. and Richerson, P. J., 2000. Memes: Universal acid or a better mousetrap? In: Aunger, R. ed. 2000. *Darwinizing Culture: The status of memetics as a science*. New York: Oxford University Press. Ch.7.

Brodie, R., 1996. *Virus of the Mind: The New Science of the Meme*. Carlsbad: Hay House Inc.

Brown, A., 1999. *The Darwin Wars: The Scientific Battle for the Soul of Man*. London: Simon & Schuster.

Claidière, N. and Sperber, D., 2007. The role of attraction in cultural evolution. *Journal of Cognition and Culture*, Vol 7, pp. 89–111.

Cosmides, L. And Tooby, J., 2000. Consider the Source: The Evolution of Adaptations for Decoupling and Metarepresentation. In: Sperber, D. ed. 2000b. *Metarepresentations: A multidisciplinary perspective*. New York: Oxford University Press. Ch.4.

Dawkins, M. S., Halliday, T. R. and Dawkins, R. eds. 1991. *The Tinbergen Legacy.* London: Springer.

Dawkins, R., 1983. Universal Darwinism. In: Bendall, D. S. ed. 1983. *Evolution from molecules to men*. Cambridge: Cambridge University Press. Ch.20.

Dawkins, R., 1999. *The Extended Phenotype: The Long Reach of the Gene*. Revised edition. Oxford: Oxford University Press.

Dawkins, R., 2003. *A Devil's Chaplain: Selected Essays*. London: Phoenix.

Dawkins, R., 2006a. *The Blind Watchmaker.* London: Penguin Books.

Dawkins, R., 2006b. *The Selfish Gene.* 30th anniversary edition. New York: Oxford University Press.

Deer, C., 2012. Doxa. In: Grenfell, M. ed. 2012. *Pierre Bourdieu: Key Concepts*. Second edition. Durham: Acumen. Ch.7.

Delius, J. D., 1991. The nature of culture. In: Dawkins, M. S., Halliday, T. R. and Dawkins, R. eds. 1991. *The Tinbergen Legacy*. London: Springer. Ch.6.

Dennett, D. C., 1987. *The Intentional Stance*. Cambridge, MA: MIT Press.

Dennett, D. C., 1991. *Consciousness Explained*. London: Penguin Books.

Dennett, D. C., 1995. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster Paperbacks.

Dennett, D. C., 1996. *Kinds of Minds: Toward an Understanding of Consciousness*. New York: The Perseus Books Group.

Dennett, D. C., 2000. Making Tools for Thinking. In: Sperber, D. ed. 2000b. *Metarepresentations: A multidisciplinary perspective*. New York: Oxford University Press. Ch.2.

Dennett, D. C., 2005. *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. London: MIT Press.

Dennett, D. C., 2006. *Breaking the Spell: Religion as a natural phenomenon*. London: Penguin Books.

Dennett, D. C., 2009a. *Intentional Systems Theory*. [online] Available at: <http://ase.tufts.edu/cogstud/papers/intentionalsystems.pdf> [Accessed 29 November 2013].

Dennett, D. C., 2009b. The Evolution of Culture. In: Dick, S. J. and Lupisella, M. L. eds. 2009. *Cosmos & Culture: Cultural Evolution in a Cosmic Context*. [online] Available at: <http://www.nasa.gov/pdf/607104main_CosmosCulture-ebook.pdf> [Accessed 24 January 2014]. Ch.3.

Dennett, D. C., 2013. *Intuition Pumps and Other Tools for Thinking*. London: Penguin Books.

Deutsch, D., 1997. *The Fabric of Reality*. London: Penguin Books.

Deutsch, D., 2011. *The Beginning of Infinity: Explanations that Transform the World*. London: Penguin Books.

Dick, S. J. and Lupisella, M. L. eds. 2009. *Cosmos & Culture: Cultural Evolution in a Cosmic Context*. [online] Available at: <http://www.nasa.gov/pdf/607104main_CosmosCulture-ebook.pdf> [Accessed 24 January 2014].

Distin, K., 2005. *The Selfish Meme: A Critical Reassessment*. Cambridge: Cambridge University Press.

Driscoll, C., 2011. Fatal Attraction? Why Sperber's Attractors do not Prevent Cumulative Cultural Evolution. *British Journal for the Philosophy of Science*, Vol 62, pp. 301–322.

Edelman, G. M., 1987a. CAMs and Igs: Cell Adhesion and the Evolutionary Origins of Immunity. *Immunological Reviews*, Vol 100, pp. 11–45.

Edelman, G. M., 1987b. *Neural Darwinism: The Theory of Neuronal Group Selection*. New York: Basic Books.

Edelman, G. M., 2006. *Second Nature: Brain Science and Human Knowledge*. New Haven: Yale University Press.

Edelman, G. M. and Gally, J. A., 1964. A Model for the 7S Antobody Molecule. *Proceedings of the National Academy of Sciences of the United States of America*, Vol 51(5), pp. 846–853.

Edelman, G. M. and Gally, J. A., 2001. Degeneracy and complexity in biological systems. *PNAS*, Vol 98(24), pp. 13763–13768.

Edelman, G. M. and Tononi, G., 2000. *A Universe of Consciousness: How Matter Becomes Imagination*. New York: Basic Books.

Eichenbaum, H., 2004. Hippocampus: Cognitive Processes and Neural Representations that Underlie Declarative Memory. *Neuron*, Vol 44, pp.109–120.

Enfield, N. J., 2006. Social Consequences of Common Ground. In: Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg. Ch.15.

Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg.

Floridi, L., 2005. Is Semantic Information Meaningful Data? *Philosophy and Phenomenological Research*, Vol LXX(2), pp. 351–370.

Gabora, L., 1997. The Origin and Evolution of Culture and Creativity. Journal of Memetics: Evolutionary Models of Information Transmission, Vol 1(1), pp. 29–57.

Gabora, L., 2004. Ideas are Not Replicators but Minds Are. *Biology and Philosophy*, Vol 19(1), pp.127–143.

Gabora, L., 2005. Creative Thought as a non-Darwinian Evolutionary Process. *Journal of Creative Behavior*, Vol 39(4), pp. 65–87.

Gabora, L., 2008. The Cultural Evolution of Socially Situated Cognition. *Cognitive Systems Research*, Vol 9(1–2), pp. 104–113.

Gabora, L., 2011. Five Clarifications about Cultural Evolution. *Journal of Cognition and Culture*, Vol 11, pp. 61–83.

Gabora, L. and Aerts, D., 2009. A model of the emergence and evolution of integrated worldviews. *Journal of Mathematical Psychology*, Vol 53, pp. 434–451.

Goldin-Meadow, S., 2006. Meeting Other Minds through Gesture: How Children Use their Hands to Reinvent Language and Distribute Cognition. In: Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg. Ch.13.

Goodwin, C., 2006. Human Sociality as Mutual Orientation in a Rich Interactive Environment: Multimodal Utterances and Pointing in Aphasia. In: Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg. Ch.3.

Grenfell, M. ed. 2012. *Pierre Bourdieu: Key Concepts*. Second edition. Durham: Acumen.

Grice, P., 1991. *Studies in the Way of Words*. Cambridge: Harvard University Press.

Grice, P., 2001. *Aspects of Reason*. Oxford: Oxford University Press.

Griffiths, P. E., 2002. What is innateness? *The Monist*, Vol 85(1), pp.70–85.

Happé, F., 1993. Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, Vol 48(2), pp. 101–119.

Hawking, S., 2011. *Does God play Dice?* [online] Available at: <http://www.hawking.org.uk/does-god-play-dice.html> [Accessed 29 November 2013].

Henrich, J. and Boyd, R., 1998. The Evolution of Conformist Transmission and the Emergence of Between-Group Differences. *Evolution and Human Behavior*, Vol 19(4), pp. 215–241.

Henrich, J. and Boyd, R., 2002. On Modeling Cognition and Culture: Why cultural evolution does not require replication of representations. *Journal of Cognition and Culture,* Vol 2, pp. 87–112.

Henrich, J. and McElreath, R., 2003. The Evolution of Cultural Evolution. *Evolutionary Anthropology*, Vol 12(3), pp. 123–135.

Heylighen, F., 1992. 'Selfish' Memes and the Evolution of Culture. *Journal of Ideas*, Vol 2(4), pp. 77–84.

Heylighen, F., 1993. Selection Criteria for the Evolution of Knowledge. *Proc. 13th Int. Congress on Cybernetics* (Association Internat. de Cybérnetique, Namur), pp. 524–528.

Heylighen, F., 1997. Objective, Subjective and Intersubjective Selectors of Knowledge. *Evolution and Cognition*, Vol 3(1), pp. 63–67.

Heylighen, F., 1999. What makes a meme successful? Selection criteria for cultural evolution. *Proc. 15th Int. Congress on Cybernetics* (Association Internat. de Cybérnetique, Namur), pp. 418–423.

Heylighen, F. and Dewaele, J., 1999. Formality of Language: definition, measurement and behavioral determinants. *Internal Report, Center "Leo Apostel", Free University of Brussels*. [online] Available at: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.30.6280&rep=rep1&type=pdf> [Accessed 3 February 2014]

Heylighen, F. and Chielens, K., 2009. Cultural Evolution and Memetics. In: Meyers, R. A. ed. 2009. *Encyclopedia of Complexity and Systems Science*. Springer. Ch.190. [online] Available at: <http://web.archive.org/web/20130310122911/http://pespmc1.vub.ac.be/Papers/Memetics-Springer.pdf> [Accessed 31 January 2014]

Hintikka, J., 2005. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Reprinted edition. London: Cornell University Press.

Horn, L. R. and Ward, G. eds. 2004. *The Handbook of Pragmatics*. Oxford: Blackwell Publishing.

Hull, D. L., 1980. Individuality and Selection. *Annual Review of Ecology and Systems*, Vol 11, pp. 311–332.

Hull, D. L., 1982. The Naked Meme. In: Plotkin, H. C. ed. 1982. *Learning, development, and culture: Essays in evolutionary epistemology*. New York: Wiley. Ch.8.

Hull, D. L., 2001. *Science and Selection: Essays on Biological Evolution and the Philosophy of Science*. New York: Cambridge University Press.

Jorgensen, J., Miller, G. A. and Sperber, D., 1984. Test of the mention theory of irony. *Journal of Experimental Psychology: General*, Vol 113(1), pp. 112–120.

Kandel, E. R., 2006. *In Search of Memory: The Emergence of a New Science of Mind*. New York: Norton.

Kuper, A., 2000. If memes are the answer, what is the question? In: Aunger, R. ed. 2000. *Darwinizing Culture: The status of memetics as a science*. New York: Oxford University Press. Ch.9.

Levinson, S. C., 2006. On the Human ”Interaction Engine”. In: Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg. Ch.1.

Linden, D. J., 2008. *The Accidental Mind: How Brain Evolution Has Given Us Love, Memory, Dreams, and God*. Cambridge, MA: Harvard University Press.

Liszkowski, U., 2006. Infant Pointing at 12 Months: Communicative Goals, Motives, and Social-Cognitive Abilities. In: Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg. Ch.5.

Mahalingam, R., 2007. Essentialism, Power, and the Representation of Social Categories: A Folk Sociology Perspective. *Human Development*, Vol 50, pp.300–319.

Maton, K., 2012. Habitus. In: Grenfell, M. ed. 2012. *Pierre Bourdieu: Key Concepts*. Second edition. Durham: Acumen. Ch.3.

McAdams, H. H. and Arkin, A., 1999. It's a noisy business! Genetic regulation at the nanomolar scale. *Trends in Genetics*, Vol 15(2), pp. 65–69.

McElreath, R., Boyd, R. and Richerson, P. J., 2003. Shared Norms and the Evolution of Ethnic Markers. *Current Anthropology*, Vol 44(1), pp. 122–130.

Medin, D. L., 1989. Concepts and Conceptual Structure. *American Psychologist*, Vol 44(12), pp.1469–1481.

Milner, B., Squire, L. R. and Kandel, E. R., 1998. Cognitive Neuroscience and the Study of Memory. *Neuron*, Vol 20, pp.445–468.

Moore, R., 2012. Capital. In: Grenfell, M. ed. 2012. *Pierre Bourdieu: Key Concepts*. Second edition. Durham: Acumen. Ch.6.

Nettle, D. and Dunbar, R. I. M., 1997. Social Markers and the Evolution of Reciprocal Exchange. *Current Anthropology*, Vol 38(1), pp. 93–99.

Von Neumann, J. and Burk, A. W., 1966. *Theory of Self-reproducing Automata*. Chicago: University of Illinois Press.

Peirce, C. S., 1955. *Philosophical Writings of Peirce*. New Dover edition. New York: Dover.

Plotkin, H. C. ed. 1982. *Learning, development, and culture: Essays in evolutionary epistemology*. New York: Wiley.

Plotkin, H. C., 2000. Culture and Psychological Mechanisms. In: Aunger, R. ed. 2000. *Darwinizing Culture: The status of memetics as a science*. New York: Oxford University Press. Ch.4.

Popper, K., 1978. Natural Selection and the Emergence of Mind. *Dialectica*, Vol 32(3–4), pp. 339–355.

Popper, K., 1979. *Objective Knowledge: An Evolutionary Approach*. Revised Edition. Oxford: Oxford University Press.

Popper, K., 2002a. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge Classics edition. London: Routledge.

Popper, K., 2002b. *The Logic of Scientific Discovery*. Routledge Classics edition. London: Routledge.

Popper, K., 2002c. *Unended Quest: An Intellectual Autobiography*. Routledge Classics edition. London: Routledge.

Popper, K. and Eccles, J. C., 1983. *The Self and Its Brain: An Argument for Interactionism*. Routledge edition. London: Routledge.

Premack, D. and Woodruff, G., 1978. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, Vol 1(4), pp.515–526.

Russell, B., 1905. On Denoting. *Mind*, New Series, Vol 14(56), pp. 479–493.

Ryle, G., 2000. *The Concept of Mind*. Penguin Classics edition. London: Penguin Books.

De Saussure, F., 2011. *Course in General Linguistics*. Translated from French by W. Baskin. New York: Columbia Univertsity Press.

Schachter, D. L., 2001. *The Seven Sins of Memory: How the Mind Forgets and Remembers*. New York: Houghton Mifflin Company.

Schubert, J. D., 2012. Suffering/symbolic violence. In: Grenfell, M. ed. 2012. *Pierre Bourdieu: Key Concepts*. Second edition. Durham: Acumen. Ch.11.

Schegloff, E. A., 2006. Interaction: The Infrastructure for Social Institutions, the Natural Ecological Niche for Language, and the Arena in which Culture is Enacted. In: Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg. Ch.2.

Sperber, D., 1996. *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell Publishing.

Sperber, D., 2000a. An objection to the memetic approach to culture. In: Aunger, R. ed. 2000. *Darwinizing Culture: The status of memetics as a science*. New York: Oxford University Press. Ch.8.

Sperber, D. ed. 2000b. *Metarepresentations: A multidisciplinary perspective*. New York: Oxford University Press.

Sperber, D., 2000c. Metarepresentations in an Evolutionary Perspective. In: Sperber, D. ed. 2000b. *Metarepresentations: A multidisciplinary perspective*. New York: Oxford University Press. Ch.5.

Sperber, D., 2001. Conceptual Tools for a Natural Science of Society and Culture. *Proceedings of the British Academy*, Vol 111, pp. 297–317.

Sperber, D., 2006. Why a Deep Understanding of Cultural Evolution is Incompatible with Shallow Psychology. In: Enfield, N. J. and Levinson, S. C. eds. 2006. *Roots of Human Sociality: Culture, Cognition and Interaction*. London: Berg. Ch.16.

Sperber, D. and Hirshfeld, L. A., 2004. The cognitive foundations of cultural stability and diversity. *TRENDS in Cognitive Sciences*, Vol 8(1), pp. 40–46.

Sperber, D. and Wilson, D., 1995. *Relevance: Communication & Cognition.* 2nd ed. Oxford: Blackwell Publishing.

Taylor, K., 2004. *Brainwashing: The science of thought control*. New York: Oxford University Press.

Taylor, K., 2012. *The Brain Supremacy: Notes from the frontiers of neuroscience*. Oxford: Oxford University Press.

Thomson, P., 2012. Field. In: Grenfell, M. ed. 2012. *Pierre Bourdieu: Key Concepts*. Second edition. Durham: Acumen. Ch.4.

Van Overwalle, F. and Heylighen, F., 1995. Relating covariation information to causal dimensions through principles of contrast and invariance. *European Journal of Social Psychology*, Vol 25(4), pp. 435–455.

Werndl, C. S., 2009. *Philosophical Aspects of Chaos: Definitions in Mathematics, Unpredictability, and the Observational Equivalence of Deterministic and Indeterministic Descriptions*. Ph.D. Thesis. University of Cambridge: U.K.

Wikipedia, 2005. *Henohenomoheji.png*. [image online] Available at: <http://upload.wikimedia.org/wikipedia/commons/1/11/Henohenomoheji.png> [Accessed 6 February 2014]

Wilson, D. and Sperber D., 2004. Relevance Theory. In: Horn, L. R. and Ward, G. eds. 2004. *The Handbook of Pragmatics*. Oxford: Blackwell Publishing. Ch.27.

Wilson, D. and Sperber D., 2012. *Meaning and Relevance*. New York: Cambridge University Press.

Wittgenstein, L., 1922. *Tractatus Logico-Philosophicus*. Translated from German by C. K. Ogden. In: Wittgensten, L., 2009a. *Major Works: Selected Philosophical Writings*. New York: Harper Perennial.

Wittgenstein, L., 1969. *On Certainty*. Translated from German by D. Paul and G. E. M. Anscombe. In: Wittgensten, L., 2009a. *Major Works: Selected Philosophical Writings*. New York: Harper Perennial.

Wittgensten, L., 2009a. *Major Works: Selected Philosophical Writings*. New York: Harper Perennial.

Wittgenstein, L., 2009b. *Philosophical Investigations*. 4th revised edition. Translated from German by G. E. M. Anscombe, P. M. S. Hacker and Joachim Schulte. Chicester: Blackwell.

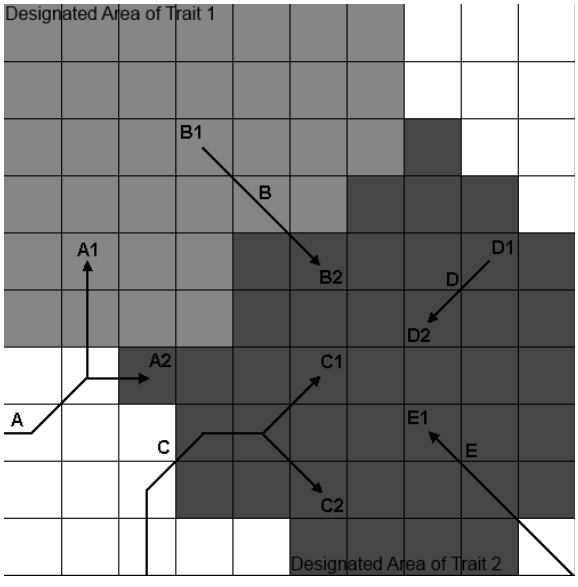
APPENDIX I: TABLES AND FIGURES

Chapter 2: Selectionism

stages/selectors	**Objective**	**Subjective**	**Intersubjective**	**Meme-centered**
Assimilation	distinctiveness	novelty simplicity coherence	authority formality	self-justification
Retention	invariance controllability	coherence utility	conformity	self-reinforcement intolerance
Expression			expressivity	proselytism
Transmission			publicity	proselytism

TABLE 2.1. Heylighen's selection criteria (Heylighen, 1999, p.5).

FIGURE 2.1. The *Henohenomoheji*. Picture acquired from Wikipedia (Wikipedia, 2005).



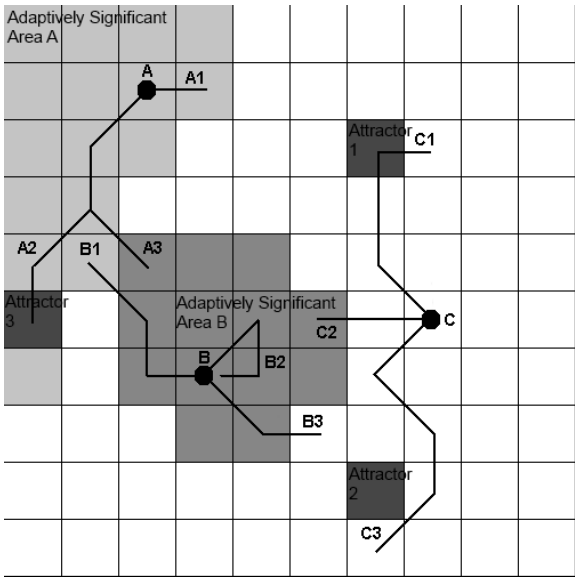
Explanations:

1. Emphasized areas:
a) Designated Areas of Traits 1 and 2
- Contain designs which share the specified traits

2. Drawings:
a) Arrows A, B, C, D and E
- Trajectories of change between designs
b) Cells A1, A2, B1, B2, C1, C2, D1, D2 and E1
- Example designs which contain the specified traits

FIGURE 2.2. Demonstration of the relevance of generalizations questioned by Sperber (Sperber 1996, pp.29, 118) using the grid he also favors (Ibid. pp.109, 112).

Chapter 3: Epidemiology of representations



Explanations:

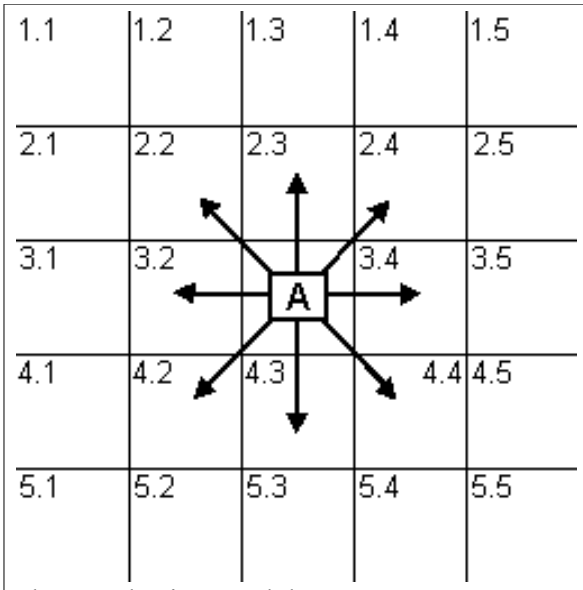
1. Emphasized areas:
a) Adaptively Significant Areas A and B
- Contain all equally adaptively beneficial iterations

b) Attractors 1–3

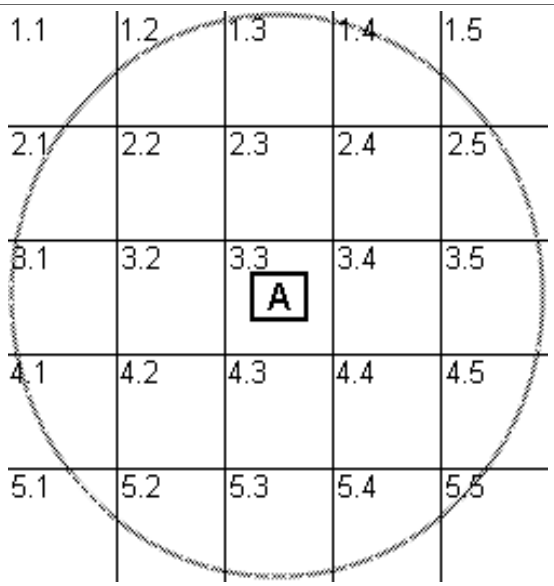
2. Drawings:
a) Dots A, B, and C
- Teachers in Driscoll's sense (Driscoll 2011)

b) Trajectories A1, A2, A3, B1, B2, B3, C1, C2 and C3
- Examples of causal chains of reprenting

FIGURE 3.1. Demonstration of Driscoll's arguments (Driscoll, 2011) by applying it to Sperber's grid (Sperber, 1996, pp.109, 112).

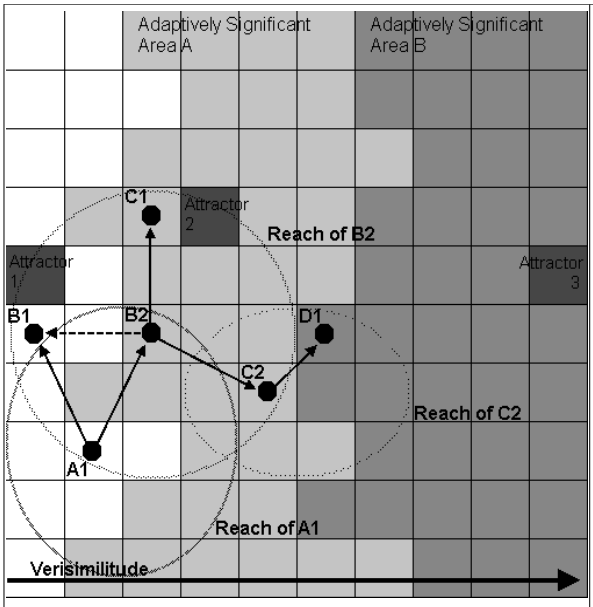


The Sperberian model



The Jump model

FIGURE 3.2. A comparison of Sperber's model of representational descent (Sperber, 1996, p.111) and a model integrating the possibility of bigger transformations, or *jumps*.



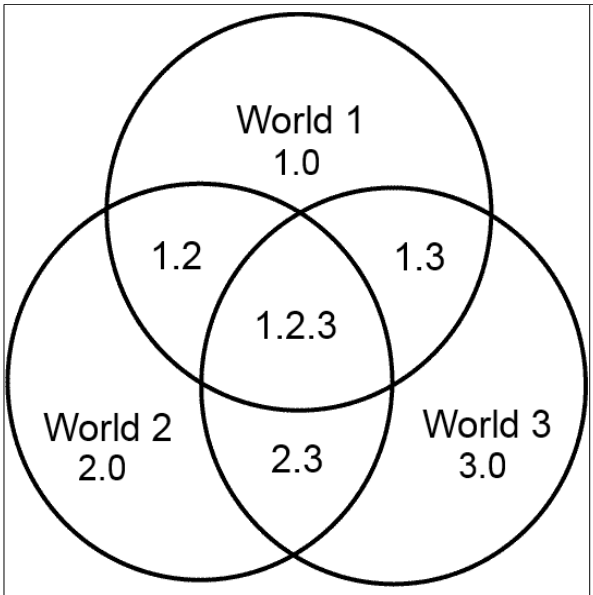
Explanations:

1. Emphasized areas:
a) Adaptively Significant Areas A and B
   - Contain all equally adaptively beneficial iterations
b) Attractors 1–3

2. Drawings:
a) Dots A1, B1, B2, C1, C2 and D1
   - Examples of representations
   - Alphabet denotes generation
b) Dotted perimeters
   - Reaches of designated representations
c) Arrows between dots
   - Transitions between forms

FIGURE 3.3. Combination of Driscoll's hypothesis (Driscoll, 2011) and the Jump model.

Chapter 4: Outline of a theory of associations

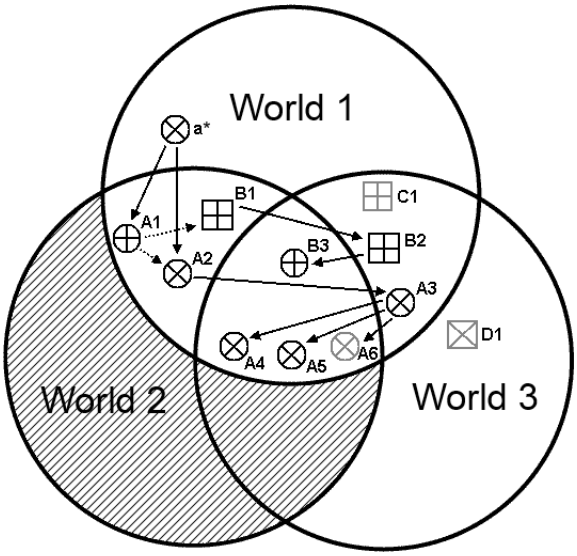


Explanations:

1. Worlds 1, 2, and 3
   - The three worlds conjectured by Sir Karl Popper (Popper 1972/1979; Popper and Eccles, 1977/1983)
   - (1.0), (2.0) and (3.0) denote the sets of entities existing only in single worlds.

2. Intersections (1.2), (1.3), (2.3), and (1.2.3)
   - Denote the sets of entities which exist in all the worlds listed by the numbers in these intersections' names.

FIGURE 4.1. A Venn diagram demonstrating the possible relationships between Sir Karl Popper's three worlds (Popper 1972/1979; Popper and Eccles, 1977/1983).



Explanations:

1. Worlds 1, 2, and 3
   - The three worlds conjectured by Sir Karl Popper (Popper 1972/1979; Popper and Eccles, 1977/1983)
   - Diagonal stripes: empty sets

2. Drawings:

a) Tokens (A1), (A2),... (A6), (B1), (B2), (B3), (C1), and (D1); and Object (a*)
   - Color (black/grey): type of object
   - Shape (circle/square): Allele 1
   - Interior (cross/check): Allele 2

b) Arrows
   - Trajectories of descent between tokens
   - Dashed when merely influencing

FIGURE 4.2. An illustration of the view about cultural representations being promoted.