Rolls, E.T./Tovée, M.J.: Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. In: *Journal of Neurophysiology* 73, 1995, pp. 713–26.
Rolls, E.T./Cowey, A./Bruce, V.: Neurophysiological mechanisms underlying face processing within and beyond the temporal cortical visual areas [and discussion]. In: *Philosophical Transactions of the Royals Society of London, B* 335, 1992, pp. 11–21.
Samuel, A.G.: Lexical activation produces potent phonemic percepts. In: *Cognitive Psychology* 32, 1997, pp. 97–127.
Smith, L.B./Heise, D.: Perceptual similarity and conceptual structure. In: *Advances in psychology 93 – percepts, concepts, and categories: The representation and processing of information*, ed. by B. Burns, Elsevier 1992, pp. 233–72.
Solomon, R.: *The passions: The myth and nature of human emotions*, New York 1984.
Sterelny, K.: *Thought in a hostile world*, Oxford 2003.
Tillas, A.: Language as Grist to the Mill of Cognition. In: *Cognitive Processing* 16(3), 2015, pp. 219–243.
Tillas, A.: How do Ideas Become General in their Signification? In: *The Baltic International Yearbook of Cognition, Logic and Communication, Vol. 9*, ed. by E. Machery/J. Prinz/J. Skilters, Kansas 2014.
Tillas, A.: *Back to our Senses: An Empiricist on Concept Acquisition*, Doctoral Thesis, University of Bristol, UK 2010.
Tomasello, M./Farrar, J.: Joint attention and early language. In: *Child Development* 57, 1986, pp. 1454–63.
Tomasello, M./Todd, J. (1983). Joint attention and lexical acquisition Style. In: *First Language* 4, 1983, pp. 197–212.
Tovée, M.J./Rolls, E.T./Azzopardi, P.: Translation invariance and the responses of neurons in the temporal visual cortical areas of primates. In: *Journal of Neurophysiology* 72, 1994, pp. 1049–60.
Weiskopf, D.A.: Concept Empiricism and the Vehicles of Thought. In: *Journal of Consciousness Studies* 14 (s 9–10), 2007, pp. 156–83.
Wimmer, H./Perner, J.: Beliefs About Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception. In: *Cognition* 13(1), 1983, pp. 103–128.

Albert Newen, Francesco Marchi

CONCEPTS AND THEIR ORGANIZATIONAL STRUCTURE

Concepts Are Templates Based on Mental Files

1. Introduction

Give their centrality to philosophical accounts of thought and reasoning, our lack of a clear account of the nature of concepts remains something of an embarrassment. Crucial questions still remain in need of an answer: (i) How can or should we characterize the minimal epistemic conditions of concept possession? (ii) Are concepts innate and, if they are, to what extent? (iii) Where is the borderline between non-conceptual representations and conceptual representations? (iv) Do animals possess concepts? (v) How can we best account for conceptual changes in the ontogenetic development of humans? (vi) How can we account for the variety of concept types, including definitional concepts, natural kind concepts, etc.?

In this article we focus on the last two questions. To deal fruitfully with them, we have to presuppose a framework that situates our views concerning questions (i) to (iv). We are not able to offer detailed answers to these questions in this paper, although our discussion bears some relevance to all of them. The new account we propose is situated between empiricist accounts of concepts, on which concepts can be fully analyzed in terms of a network of associated perceptual information (Barsalou 1999, Prinz 2004), and rationalist accounts, on which concepts are radically different in format from perceptual representations and form some kind of abstract symbols. The latter holds for, e.g. Fodor's language of thought (Fodor 1975), Peacocke's theory of concepts (Peacocke 1992), and Dretske's theory of digital representations (Dretske 1983). We set the stage for our proposal by highlighting certain criteria of adequacy for a representational account of concepts and discussing Dretske's view of concepts as digital representations. We will argue that, on the one hand, a strict demarcation between different formats of conceptual representations (digital) and perceptual representations (analog) cannot account for *conceptual transformations* as one criterion of adequacy, and, on the other hand, that concepts need perceptual features for *prototypical categorization*, an important aspect of a conceptual system. Thus, we establish

a negative conclusion: Concepts in general do not have a radically different format from that of perceptual states (though perhaps some special concepts do). Then we develop a positive analysis of concepts which enables us to offer a characterization of conceptual change. We defend three main claims. (1) Concepts can be fruitfully understood as consisting of two components: (a) an integrated associative network used for categorization of some types of entities (e. g. objects) according to properties, and (b) a handling system that organizes this associative network. (2) This account of concepts enables us to adequately describe the transformation of a conceptual system during ontogenetic development: it can be described as a transformation realized by a modification of the type of handling of the relevant network of associated information. (3) This account is flexible enough to do justice to the variety of concepts we observe in natural language (e. g. definitional concepts and natural kind concepts).

2. Setting the stage: Concepts are mental representations essentially based on epistemic abilities

In this paper we presuppose a representational view of concepts. We also presuppose the general idea that any representation can be characterized according to three features: a representational *vehicle*, a representational *content*, and a *format*. The *vehicle* is the substrate that implements the representation, e. g. a brain state in the case of mental representations; the *content* is the information that is represented; and the *format* is, roughly, the way the information is organized.[1]

We can now characterize our proposal with respect to questions (i) to (iv) above, while aiming to remain at as general a level as possible. In answering point (i), the question of minimal conditions of concept possession, we will identify two criteria of adequacy for any theory of concepts. It is a fruitful perspective to characterize concepts as mental representations and as mainly determined *by epistemic abilities*, which can be spelled out as prelinguistic epistemic abilities (Newen, Bartels 2007); some authors argue that, in addition to epistemic abilities, some normative features are constitutive of concept possession (Glock 2008). Outlining minimal conditions, we focus on two *epistemic abilities* relevant to having a concept C, namely *recognition and*

[1] If a mental representation is only combined with one single piece of information, the organizational structure mirrors that it is an atomic unit.

categorization.[2] Recognition abilities include the ability to identify the same property *p* (associated with C) realized in different objects, as well as registering one and the same object as having different properties at once, including *p* (associated with C). Categorization can be illustrated (as a sufficient condition) by the epistemic ability to form a category and register when an object (or property) belongs to it, e. g. representing being red as a colour or representing being round as a shape (for these two cases, see the categorization abilities of the grey parrot Alex presented in Pepperberg, 1999), representing a toy as a toy-animal in contrast to a toy-human (for the latter categorization ability in seven-month-old children, see Pauen, 2000). Both the abilities of recognition and of categorization can be spelled out independently of having natural language available (see Newen, Bartels 2007).

The idea that concepts correlate with specific epistemic abilities has some advantages.[3] For instance, epistemic abilities (and thus the concepts grounded by them) can be easily described not only for humans, but also for animals and robots. Moreover, this allows us to characterize the basic concepts, which we will call characteristic-features concepts, used to explain a level of representation that is more elaborated than perceptual discriminations and yet more fundamental than linguistic representations (see Newen, Bartels 2007). This epistemic view is anchored in the presupposition that a theory of concepts should enable us to predict and understand the cognitive capacities of different systems, including those of robots, prelinguistic children and language-proficient adult humans. These background considerations enable us to highlight *two criteria of adequacy* for a theory of concepts: (1) Any theory of concepts should account for minimal epistemic abilities of recognition and categorization; (2) A fruitful theory of concepts should work with a characterization that enables us to attribute basic concepts to different cognitive systems, including those of certain animals and young infants.

With this background in place, we can adopt a »working stance« on questions (ii) to (iv), while not committing to any specific answers. Concerning (ii), the question of nativism, we wish to remain generally neutral, and exclude only the possibility that all concepts are inborn, since one of our starting points is the observation of some conceptual transformations that we will analyze as changes in the organization of concepts during ontogeny (see below). Concerning (iii), i. e. the borderline between non-conceptual and conceptual representation, we here need only to presuppose that some mental representations are non-conceptual, and that (at least) humans activate or develop some additional conceptual representations, where these come

[2] We exclude Fodor's view of concepts from an epistemic theory of concept possession, since he argues that concept possession is completely independent of epistemic abilities (Fodor 1998, 6).

[3] Otherwise, one might wonder whether concepts are needed at all: see Machery 2009.

with (at least) recognition and categorization abilities (see above).[4] Question (iv), concerning whether animals possess concepts, is answered positively elsewhere (Newen, Bartels 2007), but for present purposes we presuppose a neutral notion of concepts that does not by definition exclude that animals may have them.[5]

In addition to the two criteria of adequacy for a theory of concepts given above, we want to add a further criterion, connected with question (v): How can we account for the transformation from simple conceptual representations into complex conceptual representations? Such transformations happen paradigmatically in ontogeny, when children unfold their linguistic understanding of words. More precisely, we observe a shift from understanding a word *relying on characteristic features* toward a *definitional understanding* (Keil 1992). In our terminology, we want to call this a shift from a *characteristic-features concept* to a *definitional concept* associated with the same word. This transformation within the realm of concepts is our third criterion of adequacy: (3) A theory of concepts should account for the conceptual transformation from *characteristic-features concepts* to *definitional concepts* that happens, at least for some expressions, with the acquisition of language during normal ontogenetic development.

How must concepts be characterized to meet all three criteria of adequacy? In order to develop our own account, the first step is to question the idea that conceptual representations must be radically distinguished from perceptual representations. To address this problem, we start by discussing the influential work of Fred Dretske, who wants to account for concepts by means of the notion of digital representations, where the contrast class for digital representations is analog representations. There are two possible interpretations of Dretske's view. A gradualist interpretation admits that digital representations are less informationally rich than analog representations; a radical interpretation posits a neat distinction between the representational basis of perception and concepts. According to this stronger reading of Dretske, analog and digital representations are distinct classes of representations: all perceptual representations are analog while all conceptual representations are digital (Dretske 1983). We criticize this radical interpretation of Dretske's view, on the grounds that it is incapable of accounting for one of the main features of our categorization abilities, namely prototypical categorization. Then, we develop an alternative account that starts by integrating perceptual

[4] A detailed suggestion is made in Newen & Bartels 2007. The debate about concept possession in animals is a rich independent research area; see also Chater & Heyes 1994; Allen 1999; Glock 2000.

[5] The opposite view is sometimes defended, by defining concepts as being essentially connected with natural language abilities (Davidson 1985).

representations as part of conceptual representations. Furthermore, we will account for abstract concepts as well. Finally, this positive view is defended by arguing that it can account for our third criterion of adequacy, i. e. the conceptual transformation in ontogeny that is captured neither by the change from analog to digital representations nor by purely empiricist theories of concepts.

3. Dretske's proposal: Concepts as digital representations

As we argued above, the cognitive capacities of recognition and categorization are two of the fundamental criteria for having a concept. In this section we discuss whether, in order to account for these two epistemic abilities, we should characterize conceptual representations as fundamentally different from those underlying perceptual capacities.[6]

According to Fred Dretske (1983), there is a fundamental difference between the conceptual representations underlying the capacities of recognition and categorization, and the non-conceptual representations underlying other more basic cognitive capacities without concepts, such as perceptual discrimination. Dretske characterizes this difference in terms of representational formats, distinguishing between *analog representations* and *digital representations*. This distinction is based on how much information is encoded, and is exemplified by the difference between the information encoded in a sentence and the information encoded in a picture.

Dretske claims that nonconceptual representations of perceptual properties, such as colors, lengths, etc. on the one hand, and conceptual representations used to categorize and recognize those properties on the other, are encoded in these different formats, namely analog versus digital representations.[7] The reason for this demarcation is that representations might diverge from one another insofar as one representation might encode only specific information about some object or property – for example, the statement »there is coffee in the cup«[8] encodes only the information about a cup

[6] The result of this discussion, as we shall see, has significant consequences for other relevant topics connected to the nature of concepts, such as animal concept possession and the possibility of cognitive effects on perception. For the purposes of this paper, however, the upshot of this section will provide the starting point from which to develop our account of the development of a human conceptual system and our explanation of the shift in categorization strategies reviewed above.

[7] Dretske originally uses »forms«. For consistency, we adopt the notion of format as introduced above.

[8] The example is Dretske's.

with the property of having coffee in it – while another representation might encode additional information about the same source – for example, a picture of a cup might carry information about color, shape, orientation, quantity of coffee, etc. Dretske takes this difference to be a fundamental one. Looking at a scene he calls the representation that encodes specific information about some aspect of the scene *digital representation*, and the representation that encodes all the additional information about the scene *analog representation*.[9] In short, *analog* representations are information-rich, while *digital* compositions are information-specific and information-poor. In principle, we take it that a digital representation could be information-rich, but this would require such a representation to have a structure similar to that of a very complex sentence, such as »there is a red cup in the center of the white table ... The cup is oriented 25° to the left ... The cup is approximately half full of coffee«, etc. It is immediately clear how such a sentence is already fairly complex, yet still informationally poor compared to a picture-like representation. On the contrary, the process of digitalization is essentially a process of cutting off additional irrelevant information. According to Dretske, this is a necessary computational step in order to allow for generalization and recognition of the source of a signal and its properties across multiple instantiations and contexts. This is the basis of categorization and, thus, of conceptualization.

To be even more precise, according to Dretske a representation in digital format is an informational structure that carries only the most specific[10] information about some thing *s* that has the property of being *F*. That is to say, if a signal that reaches the system is carrying the information that *s* is a red square, for example by seeing a red square, this signal carries information in analog format about *s* being red and *s* being a square. The system can then extract two different digital representations from this signal, namely that *s* is red and that *s* is a square. In both cases, the digitally represented properties are the most specific pieces of information that the signal carries.

To be the most specific pieces of information means, roughly, that these pieces of information are not nested (either nominally or analytically) into some other pieces of information carried by the signal. The representation of *s* being a square also carries information about *s* being a rectangle, quadrilateral etc. The same goes for the representation of *s* being red, which also carries

[9] Along the same line, Peacocke (1986) claims that perception represents in a unit-free manner, that is, it does not represent discrete measure-units, like centimeters for length or kilograms for weight. Conceptual representations, on the other hand, being discrete in format, are bound to represent information according to such measure-units, and are therefore intrinsically different from perceptual representations. Macpherson (forthcoming) discusses this issue in connection to the possibility of cognitive penetration, under the assumption that perceptual experiences have non-conceptual content.

[10] For a criterion of specificity, see below and Dretske 1983, Ch.7.

information about *s* being coloured, etc. Both being a *square* and being *red*, however, are the most specific properties of *s* conveyed by the signal, while *rectangle* and *coloured* are nested in *s*'s being square and red. If the signal had been the written sentence »*s* is a rectangle« then the system could have digitalized *s* is a rectangle as the most specific information about *s*. Moreover, digital representations are insensitive to their causal source: a system can extract the same digital representation from a picture or a written sentence.

If Dretske is right, in order to allow for the capacities of categorization and recognition, information encoded in perception has to be digitalized. Hence, a system that is incapable of this process of transduction of information into a different format will not meet the criteria for concept possession that we have identified. However, as we mentioned above, there are two possible readings of the distinction between analog and digital representations. According to the weak reading, the distinction is only quantitative, and the process of digitalization is just a matter of gradual refinement and exclusion of irrelevant information. According to another more radical reading, the distinction captures two independent ways of encoding information. The process of digitalization transforms representations that necessarily encode additional information about what they represent[11] (analog) into representations that are somewhat akin to syntactically structured symbols (digital). Although the gradual development that is constitutive of the weak reading of Dretske's distinction captures an important aspect of the cognitive development of representations in general, we argue that it is not helpful for characterizing concepts. Furthermore, we believe that the radical demarcation between different formats posited by the stronger reading of the distinction is not an adequate way to characterize concepts either.

We have one main reason for rejecting the stronger reading of Dretske's distinction of formats as a mark of the conceptual, namely that *analog* representations can account for certain categorization capacities with which *digital* representations seem to have problems. We refer to typicality effects, which are categorization effects that depend on the similarity and family resem-

[11] According to Dretske, a signal may always carry information in both analog and digital format. However, his coffee cup example (see above) suggests that he thinks the analog information conveyed by the signal, which could be either an utterance or a written form of the sentence »there is coffee in the cup«, is irrelevant and could be disregarded. This is not true according to modern philosophy of language, because the information transported by the communicative context, and e. g. the intonation of the utterance (»semantic focus«) is not only pragmatic, but also influences the semantic content. But, for the sake of argument, let us ignore these aspects. Then, the question is: is it correct – as Dretske suggests – that the speech signal only carries specific information about there being coffee in the cup, without any additional information and, especially, no additional perceptual information like the color of the cup?

blance of stimuli.[12] During prototypical categorization,[13] i. e. categorization that shows typicality effects, a certain object x is judged to be a better instance of F than another object y, where such a judgment is formed quickly and naturally by subjects. Since digital representations only carry the most specific information that x is F or that y is F, it seems very hard to see how such structures can account for the fact that sometimes x_s are judged to be a better instance of F than y_s.

According to some authors,[14] in order to quickly evaluate whether a given stimulus is more or less typical for a category, a system must adopt different features of comparison, activated with the stimulus, that enable it to highlight more typical exemplars and to memorize these as template for future categorizations. For instance, if a robin is judged to be more typical than a raven for the ›bird‹ category, the multifaceted information that allows for such a judgment should somehow be reflected in the concepts ›robin‹ and ›raven‹. In order to readily make such a judgment, a system must be able to simultaneously compare different characteristic features of robins and ravens to those of previously encountered members of the ›bird‹ category. The point here is that a system is able to categorize a stimulus on the basis of its appearance and degree of resemblance to other previously encountered stimuli.[15] In order to do so, during one such categorization the system must have access to a rich associative network of information, which enables the matching of the perceived stimuli with stored templates, where these will depend upon previously encountered stimuli. According to Dretske, the availability of such informational abundance, which defines analog representation, is lacking in digital representation. Therefore, it seems that such instances of categorization are possible only when information is encoded in analog format. Thus, Dretske's notion of digital format cannot account for any concept that includes a typicality effect, and thus his notion of digital representations is too narrow. Even assuming it is possible to achieve the same result through *digital* representations, in order to do so the system would have to construct an implausibly complex sentence-like representation of all the relevant information. Therefore, even if some important cases of categorization can be achieved by *analog* representations alone, we lose much

[12] See Rosch et al. 1976, Rosch 1978, and Prinz 2004.

[13] We are aware of the Fodorian rejection of prototype theory, and we do not wish to defend a form of prototype theory here. We are interested in the development of categorization and recognition capacities in human subjects, and we hold that typicality effects are an important feature of how categorization works that ought to be explained by an adequate theory of concepts, and not simply added to such a theory as the result of some other cognitive process.

[14] See, for example, Gärdenfors 2000, Rosch 1978, Lakoff 1987.

[15] The detailed differences between prototype-based and exemplar-based explanations of these effects are irrelevant for our present purposes.

of our motivation for the strong reading of Dretske, which posits a necessary step of transducing information into *digital* format in order to reach a level of conceptual representations in the first place. In addition, this provides no central characterization of concepts in general.

What point have we reached so far? Prototypical categorization offers some compelling reasons to abandon the strong reading of Dretske's distinction between formats in the debate about concepts. What about the weaker reading? If, as we hold, the weaker reading is taken to mark a merely quantitative difference, i. e. a gradual change of focus from perceptual features that are analog to very specific and limited informational features that are digital, then Dretske's view is in principle compatible with our own. But this gradual change is not necessary for the acquisition of basic concepts, as we will show. Furthermore, purely digital representations only account for a very limited amount of special concepts, e. g. abstract concepts that are only anchored in specific and limited descriptive information (see below). Digital representations cannot account for concepts in general since, as empiricist theories convincingly highlight, most (if not all) concepts integrate perceptual analog features. Thus, the appeal to *a digital format* is not a helpful way to characterize concepts in general. We need a new perspective in order to account for the core of concepts.

In light of the above discussion, we still want to retain the canonic characterization of representations according to *vehicle*, *content* and *format* that we presented at the outset. In order to do so, in the remainder of the paper we shall adopt a different notion of format than the one discussed by Dretske. We suggest that in addition to the associated information that is integrated to constitute a relevant *cluster of information*, the integrated information receives and needs a *special handling* to form a conceptual and not just a non-conceptual cluster of information. Therefore, we characterize our notion of format as the *handling* of the relevant cluster of information associated with a concept.

How is this new framework related to Dretske's account? In contrast to the strong reading of Dretske, we argue that we do not need to presuppose a change from analog to digital format to account for conceptual representations,[16] and concepts with typicality effects, at least, are essentially analog. Furthermore, in relation to the weak reading of Dretske, we will show that

[16] In the case of linguistically expressed concepts, we hold that the contents of conceptual representations always include at least some additional information about the represented entities and properties, not only the information specifically expressed by the linguistic label. For example, »being a tennis ball« specifically expresses the property of being a tennis ball, but also includes information about being yellow, having a special size and material, and being often used to play with dogs. In this respect, strictly speaking, no conceptual representation is in the Dretskean purely digital format, since, almost always, part of the relevant information remains

a gradual change from analog to digital representations is not helpful in describing the relevant transformation from one type of concept to a new type of concept that happens during ontogeny, from 4 to 9 years of age. Instead, the latter can best be accounted for by our two-factor theory of combining a cluster of associatively activated information (that accounts for the empiricist intuition that we need to ground our concepts, at least ontogenetically, in perception) with a specific way of handling this information cluster (that accounts for the rationalist intuition that concepts can have some specific organization, especially abstract concepts). On this view, a cluster of associated information about an object (events, and so on) can be organized and reorganized by the handling system to form flexible property clusters.

4. Outline of the alternative: concepts as mental files constituted by associated information and a handling system

In the above discussion, we argued that positing a radical distinction between analog and digital formats does not provide a good approach for the purposes of characterizing the foundations of conceptual representations. Nevertheless, we hold that some of Dretske's reasons for preserving certain fundamental differences between perceptual and conceptual representations must be taken into account. We reformulate his main motivation for positing a difference between analog and digital formats, with two observations: firstly, many of our concepts carry specific information best expressed in a natural language. Secondly, some of these are abstract or even scientific concepts, such as the concept GENE. Laypersons often lack an adequate grounding in perception for such concepts, and the relevant representation seems to be solely constituted by descriptions. The Dretskean view may account for this observation by introducing a *digital format*, arguing that the concept GENE consists of some typical descriptive information. Can we account for abstract concepts like GENE without re-introducing a digital format?

In order to develop such an account for abstract concepts, we rely on the framework of *mental files* (See Perry 1990, 2002; Kahnemann et al. 1992; Gordon & Irwin 1996; Newen 2010/11; Recanati 2012). In its basic form, mental files can be thought of as containers which are used to collect relevant information. This way of thinking,however, is too metaphorical and problematic because the mental file is supposed to be the *vehicle* of information but not the content.[17] We want to avoid this problem by understanding a mental file as a unity of relevant information (not just as a vehicle). The information unified as a mental file includes information that can come from at least two different sources. The first source is perception, which delivers us information about basic sensory features (visual, auditory, etc.) of an object or situation and sensory-motor information, as well as representations of affordances.[18] The second source of information is language-based thinking and reasoning, which enables us to memorize descriptive information about an object or situation. In a mental file for an object, property, or situation, both types of information are unified into an object-file, property-file or situation-file. Hence, the content of a mental file can consist in two fundamental types of information: perceptual information (which includes sensory-motor and image-like information as well as affordances) and descriptive information. Such a cluster of information can then be handled in specific ways. But let us first focus on the content of a mental file.

We illustrate this with a mental file of the property of being a chair. This file unifies representations of the primary sensory features of chairs, such as color, texture, shape, etc. which are structured according to typicality.[19] It also includes information about the affordances, i. e. the different actions that can be performed with chairs, like »sitting«, »grabbing« and so on. Finally, if a cognitive system has a natural language, then the *file* may include descriptive information about typical chairs, such as »it has four metal legs«, »it is white«, »it has metal arms«, »it is comfortable«, etc. (more on this below).

Note that the amount and grain of information depends on the knowledge and expertise of the subject with the given object. For instance, a chair designer's *chair-file* will plausibly have more refined content than the layperson's: this holds for both perceptual and descriptive information, since we accept that our perception is shaped by our descriptive knowledge (Vetter, Newen 2014). A mental file of an object, property etc. is a unified package of information about the object, property, etc. If a file of the property of *being a chair* is created on the basis of just one experience of seeing a chair, it may contain very specific and only partial information (perhaps not registering the property of being a folding chair). After having seen many chairs, we are able to memorize rich information, i. e. a combination of features of a few concrete chairs, as well as an average of features that one can expect to be realized by a chair (the most probable of these are established by one's previous experiences). The development from poor to rich files is a relevant fact,

perceptual and thus analog, or is connected with information that is not »nested in« the relevant fact, and is thus by Dretske's definition analog as well (Dretske 1983, 137). In the case under discussion, neither the color nor the regular use of tennis balls for playing with dogs is nested in the relevant property of being a tennis ball.

[17] For a discussion of the problems, see Leonardi 2013.

[18] Considered here as relevant action possibilities in a broad sense.

[19] This information is fundamentally the same, whether someone perceives or imagines the chair.

which is independent of the type of information in the file, be it perceptual or descriptive.

So far, we have described the associative network of information that is one core component of a concept. Now we need to introduce the second component, which is a *specific way of handling* this network of information. Our idea is that the contents of a mental file always come with a certain degree of internal organization, which depends on how the information in a file is handled. Why do we need this dimension of handling information? We propose two reasons: (1) We have already mentioned that we have the ability to develop typical representations for a property, such as being a chair, after having seen some chairs. This is only possible if the associated information is not simply collected, but also combined with a weighting process[20] such that the information unified in a property file of being a chair is structured according to features with graded typicality. (2) We can form natural kind concepts or descriptive concepts on the basis of the same associative network of information. With the word »water«, we express a natural kind concept, i. e. we refer to the substance that comes out of our taps and flows in rivers. This should be understood as a definition of the concept from samples of water. We usually learn to associate surface properties with this substance, e. g. being tasteless, being transparent, being a fluid, being capable of satisfying someone's thirst, etc. The associative network of information that a person associates with the word »water« underdetermines the concept WATER a person has. If the person takes the surface properties to define the concept $WATER_1$ then, following Putnam (1975), a substance with the chemicals XYZ on Twin Earth would also be an instance of water, since in the famous thought experiment the substance XYZ is supposed to have the same surface properties as water on earth while having the different chemical structure XYZ. If someone instead takes *the samples of water on earth* to determine the concept $WATER_2$, then the chemical structure of earth's water, i. e. H_2O, determines the relevant substance. In this case, the substance XYZ would not count as water. The understanding of the word »water« in our language is the second case: the natural kind concept $WATER_2$ as determined by *the samples of water on earth* in contrast to the descriptive concept $WATER_1$ determined by *the surface properties*. The difference between the two concepts is not a difference in the associated information, which can be exactly the same since we can create both concepts independent from any knowledge of chemistry.

[20] It is important to note that the handling system is not supposed to be some independent cognitive feature. We think of it as a characteristic of the representational vehicle that encodes the information, in this case the neurons. It is not possible to expand on this point here, we can only mention that we believe it is possible to describe the weighting process, and thus the handling, according to probabilistic models of neural activity.

The difference is only due to *the handling of this information*, which leads to a different way of determining the reference of the concept and a different role for the surface properties. This observation allows us to clearly mark the difference from Barsalou's (1999) theory. His theory and ours share the first component, i. e. an integrated associative network of information. Barsalou also accounts for a contextual selection of a part of this information in a situation, but this contextual selection does not involve any restructuring or reorganization of the information. And the latter handling process is necessary to account for the difference between descriptive concepts and natural kind concepts which may rely on the same contextual associative network of information.

Let us make an intermediate summary: concepts are determined by (at least) two components, an integrated associative network of information that enables recognition and categorization, and a specific way of handling this information. The result of the interaction of these two components provides templates that are the basis for recognition and categorization. As we will illustrate shortly, recognition and categorization are realized by a process of pattern matching, which is based on these templates. As we shall see, templates that are used for recognition and categorization are the best candidates for *conceptual* representations. To work out the details of the alternative view of concepts as mental files, we will first outline the development of mental files (4.1) and then describe how mental files can be used for recognition and categorization (4.2).

4.1 The development of mental files

(i) *Two types of information in a mental file*

In perceiving an object, information starts to be integrated into an object file when the subject learns how to unify the information as coming from one and the same object. Sensitivity to object permanence (Baillargeon 1987, 1993) is an early test that indicates this competence in children. Information about the object gets systematically enriched alongside cognitive development during ontogeny.[21] On the basis of object permanence (Baillargeon 1987) infants learn step by step to register more features of the objects they perceive:

[21] We are not taking any stance in the debate about nature and nurture, i. e. about which abilities are inborn and which learned. We are impressed by the work of Carey (2004) and Spelke (Kinzler & Spelke 2007), who argue for several inborn capacities that they analyse as core cognition. It is an open debate whether the informational units which are constructed relying on core cognition only deserve to be characterized as conceptual. In this paper, we outline one way to look at the borderline between nonconceptual and conceptual representations, but our

color, shape, the number of the objects in a situation, some functions of the object (starting with sensorimotor affordances) and more and more properties (Arterberry & Bornstein 2001).

At 14 months, infants start to register their own perspective towards an object and distinguish it from the perspectives of others (Sodian et al. 2007), and they learn to understand how a half-constructed object, e. g. a house made of Lego, can and should be completed. They learn to understand not only physical properties, but also social regularities associated with objects: if an adult introduces red blocks as sandwiches and white blocks as soap in a game context, two-year-olds quickly insist on using these objects with the newly introduced regularities, and protest if someone plays differently (Rakoczy 2006; Rakoczy et al. 2008). Furthermore, together (or parallel with) the acquisition of a natural language, children learn to associate linguistic descriptions with objects, which increase in complexity. Descriptive information stored on the basis of linguistic symbols can be (at least partially) characterized by specific and comparably rich inferential processes, e. g. if a child learns to understand the word »not«, then it must learn to understand the inferential connection that »it is not the case that p« is true if »p« is false (in addition to learning that »true« is a feature of states of affairs as being realized). Thus, we can distinguish two types of information in mental files: perceptual information and descriptive information. Both of these are associated with an object in an object-file.[22]

(ii) *The handling of the information in an object file*

If a child has learned to integrate associated information about an object, this means that the child is able to handle or organize this information in order to form object templates, which can then be activated in a situation. The information stored about one object in an object file is selectively activated, dependent on the context. Let us call the selectively activated unity of information associated with the object the (situated) object template. It

focus is not to work out the borderline but to describe the central structure of concepts that we need to presuppose in order to account for the variety of concepts and their transformation.

[22] So far, we have highlighted the acquisition of information about an object, but the same story should be told about properties, if children focus on properties. In parallel with registering object permanence, children must learn to register stability of some properties and also learn to register one and the same property as being instantiated in different objects, i. e. they learn to construct property-files by unifying the information they have of one and the same property. Furthermore, this holds for each type of entity that children learn to focus on, e. g. types of relations, events, processes, and situations. Furthermore, if a child has linguistic competence, then one way of initiating a mental file can be realized by storing a new linguistic label, e. g. a name, with the disposition to unify relevant incoming information with that label.

persists long enough to shape our perceptual experience (Noles et al. 2005), and is the essential basis of recognizing or even categorizing an object. Object recognition can be described as a process of pattern matching using object templates constructed from mental files.[23] Early object-files enable children to activate a template associated with one and the same object, which allows them to search for the object and recognize it.[24] For instance, after a certain age, children expect the same object to be under a cover when they have observed someone in the act of hiding it there (Luo et al. 2003).

In its basic form, at the beginning of development, the activation of a template is nothing more than the activation of all the associative information about an object that is unified in an object-file. However, the organization or handling of the information gets more and more sophisticated as ontogenetic development proceeds. At an early stage, a child may simply integrate the information associated with an object in an object-file without much selection, but as soon as typical features are separated from non-typical features, a weighting process of the collected features is already taking place. This is especially true if the weighting of characteristic features is not only determined by the frequency with which such features are observed. The handling system enables a person to modify an object file from a simple collection of features, whose relevance is only determined by the frequency of observation, to a complex organization of systematically weighted typical features. The latter are sensitive to context and further independent factors. Most importantly – as we will illustrate – this is a framework that also can account for the transformation of a characteristic-features concept into a definitional concept. Last but not least, the handling system enables us to create new concepts, by constructing a new integration of associated information using the information across different, already established files. Thus, concepts can be fruitfully understood as the result of the combination of an integrated cluster of associated information, and its handling.

[23] This is in line with the intensive work of Anne Treisman on the role of object files in perception: Treisman 1992, Kahnemann, Treisman, Gibbs 1992.

[24] The description of object files is compatible with the dual processing theory of perception, especially since in recent years there is more and more evidence to suggest that normal object representation is essentially dependent on the coordination of both pathways (Farivar 2009; Perry, Fallah 2014).

4.2 Recognition and Categorization with Mental Files

Recognition and categorization abilities are (at least) necessary conditions for having concepts. How can we account for them within a theory of mental files? Given the general framework of mental files as a model of mental representation, our idea is that the contents of a mental file always come with a certain degree of internal organization, which is a result of how the information in a file is handled and then used. Let us assume for the moment that the initial constitution of a mental file happens when information of a certain kind is represented in the system for the first time. At this initial stage, the system is supposed not to possess full-blown linguistic abilities. The only kind of information that is represented at this stage is perceptual. During the regular development of an organism, ordinary perceptual interaction with the world involves different sensory modalities and conveys information about different perceptual features that are bound together in a certain way. A given auditory stimulus, for instance, is reliably connected with a certain visual stimulus and not with others. It is precisely this minimally organized way that perceptual signals come to our sensory organs based on fundamental phenomena like reliable co-occurrence that allows the perceptual system to interpret different stimuli as representations of physical objects, sounds, background, etc. Such minimal organization becomes very important for the system to reliably predict the presence of a certain feature when a certain other feature is detected. This basic form of perceptual feature-matching unfolds into a form of prediction, which enables the matching of incoming signals onto object templates generated from stored mental files of objects or properties that have previously been encountered. This enables the basic process of object recognition: *Recognition of an object is realized by a process of pattern matching on the basis of an object-template activated by a minimal handling (or organization) of the information in the corresponding object-file.*

To account for categorization, we have to describe the recognition of properties. It is evolutionarily important for animals in general to register properties, e. g. being a predator, being edible, being of a higher social rank, etc. Of course, animals and human infants first learn to register these properties only on the basis of some sensorimotor contingencies. Afterwards, children (and perhaps some animals) learn to develop enriched property-files that enable them to categorize objects according to properties, e. g. as being of the same shape, the same color, the same material, the same basic function, etc. Then they integrate not only sensorimotor patterns but also affordances attached to properties (e. g. being a knife comes with the affordance of cutting objects). If the property file is rich enough, and has some systematic interaction with other files, it can constitute a concept (Newen, Bartels 2007). Furthermore, if a child acquires natural language it can also

associate language-based descriptive information as part of the property file. *Categorization of an object according to a property is realized by a process of pattern-matching on the basis of a rich property-template activated by a minimal handling (or organization) of the information in the corresponding property-file.* We have now seen how this framework of mental files can account for the central abilities of recognition and categorization.

After enough perceptual encounters with a given object or property, and once the corresponding mental file has sufficient structure, any future perceptual encounter with similar objects or properties will be categorized on the basis of a relatively stabilized and activated template. As we shall see, in the case of the transformation from a characteristic-features concept into a definitional concept, we need to presuppose specific ways of handling information, since the change of concept can take place even if the relevant associated information remains the same. Achieving such a flexibility is precisely the task of the handling system that we have characterized. We shall now address the question of which mental files are rich enough to be classified as concepts.

5. Which Mental Files are Concepts? Concepts are Templates Based on Mental Files with a Specific Structure

Recognition and categorization are clearly necessary conditions for having concepts. But are they also sufficient? Recognition is a very basic cognitive ability, as we have indicated above: if a five-month-old baby passes the object permanence test and has a mental file of her red puppet, then this file may only contain very parsimonious perceptual information, such as some sensorimotor and image-based features that are sufficient to recognize the puppet. If, on the basis of such minimal features, someone takes recognition to be sufficient for having a concept PUPPET, then the ascription of the concept would be explanatorily useless, because it would be identical with the explicit ascription of the basic ability to recognize individual objects. Does the situation change if we add some categorization abilities? Even if the child starts to generalize some properties of the puppet, like being red (as compared to being yellow or blue) and being soft (in contrast to being hard), this does not lead to a fruitful ascription of having a concept PUPPET. The situation changes if the baby learns to recognize the complex property of being a puppet and on this basis categorizes different puppets, as compared e. g. to being a toy dog and being a toy car. This may still be possible independently of natural language, but this is not a question on which we take any position here. Our main point is that the mental file of an object or property can start to unfold with rather parsimonious content, and it is not fruitful to define this as a

concept, even if basic abilities of recognition and categorization are included. At some point, the mental file is systematically enriched, and enables the categorization of the property of being a puppet for several puppets within a sufficiently rich contrast class (of toys), while the puppets also vary at least in some of their features, like colour, size, and solidity. The minimally rich categorization of being a puppet presupposes some adequate way of handling the associated information in the mental file. It is then useful to attribute a concept PUPPET to someone, since this involves the fruitful prediction of a cluster of flexible behavioral dispositions of recognition, categorization, and a sufficiently rich handling of the associated information (for a detailed argument and definition of concept possession which does not presuppose natural language, see Newen, Bartels 2007). The details of this view of concept possession do not matter for the purposes of this paper. It is sufficient to accept that with the acquisition of a minimally mastered natural language (at the latest), a child learns the abilities described above and the mental file becomes rich enough to function as a concept. *Thus, if a stored mental file is activated in a situation as a template which is used for (minimally rich) processes of handling the unified information, and such a template is structured enough to be involved in recognition and categorization, then this template is a concept.*

One upshot of this view is that concepts are *situated* mental representations anchored in mental files. We allow that, loosely speaking, not only the template but also the underlying mental file can be called a concept, because sometimes an object file can be identical with the template (in this case the associative cluster of information is activated without further selection); but strictly speaking the mental files are the background associative information clusters stored in long-term memory which enable us to activate a concept in working memory as a specifically structured template in a situation. Furthermore, the development of mental files offers a framework for characterizing the continuous development from nonconceptual representations to conceptual representations. If a child learns a natural language, then the child starts to include descriptive information in addition to perceptual information into the mental files. Descriptive information comes with a specifically rich network of inferences, while perceptual information may come with some associative relations, but these are nevertheless significantly less rich than language-based inferences. The development of mental files is closely connected with the development of natural language understanding, and thus with language-based concepts. Why should we take language-based concepts to be best characterized by mental files? This question will be answered by discussing the third criterion of adequacy for a theory of concepts: we must be able to account for the conceptual changes within the ontogenetic development of linguistic understanding, in which there is a shift from an understanding relying on

characteristic features to a definitional understanding of words. This will be explained using the example of the transformation of the understanding of the word »island« between the age of 4 and 9.

6. The third criterion of adequacy: conceptual transformation in ontogeny

Frank C. Keil (1989) describes the development of language-based concepts between the age of 4 and 9 as a shift, in three stages, from characteristic to defining features in the understanding of concepts. He investigated dozens of words expressing concepts, such as »Lie«, »Robber«, »Menu«, »Twins«, »Taxi«, »Factory«, »Lunch«, »Church«, »Island«, etc. While 3–4 year-old children understand these words as expressing concepts only on the basis of characteristic features, 9–10 year-olds have learned to understand the concept as based on a definition. Eight-year-olds typically show an intermediate understanding relying on both characteristic and definitional features. In the case of the word »island«, Keil observed the following shift: 3–4-year-olds have an understanding solely based on characteristic features, such as palm trees, sandy beaches, seashells, warm weather, sunshine, and swimming, In contrast, 9–10 year olds rely on the definition »being a body of land surrounded by water on all sides.« In an intermediate state, children (of roughly 8 years) rely on both characteristic as well as the relevant defining features. How can we account for this systematic shift from characteristic to defining features in the understanding of concepts?

The framework of mental files is well suited to meeting this challenge. Since the 3–4-year-olds are able to use the word »island« in a lot of sentences, mainly correctly, we have to presuppose that there is already a conceptual representation at this stage; but then we need to account for the radical transformation from the early concept of ISLAND based on characteristic features to the late concept of ISLAND based on definitional ones. Let us call the early concept a *characteristic-features concept* and the late concept a *definitional concept*. We can describe the development using the tool of semantic frames (Löbner 2015). These frames characterize a concept using themes and values of themes: a 4-year-old child may experience an island during summer in Greece, and to characterize this experience we can use themes (purpose, activity, temperature, etc.) and their respective values.

If a 4-year-old is told that she is taking a holiday on an island, and she experiences the island as having sand, seashells, and being near water, and that it is hot, summertime, and ideal for swimming, than these experiences may be represented by a flat array of features almost devoid of any specific organization. All these features are taken to be characteristic of islands. This

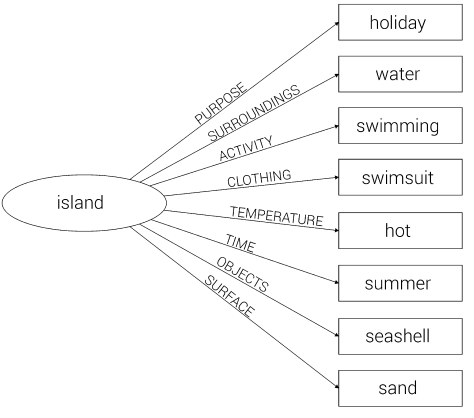

Figure 1: *The chararcteristic-feature concept of ISLAND of a 4-year-old*

is a basic form of a characteristic-features concept, and since it relies on an integration of *characteristic* features into a pattern, we can easily account for the typicality effects observed for certain concepts.

In the further ontogeny, a child learns to separate some characteristic features like relational affordances (e. g. the possibility of taking part in sports) from typical properties of the object, like hotels, beaches, palm trees and being surrounded by water. This is represented by the relational affordances listed on the left hand side, while the objective properties are drawn on the right hand side (see below). The transformation takes place by enriching the information with a more detailed hierarchy of interdependent properties, as well as a specific way of highlighting the relevant characteristic features, in this case the properties of the object, by inhibiting the relational affordances. This process continues until we reach a definitional understanding, if the same associated package of information is organized such that the only characteristic feature that counts is the descriptive information »being a piece of land completely surrounded by water.« Thus, the transformation process of concepts related to the same word from characteristic to defining features can only be adequately described if we account for two factors involved: the associated information combined with the word, and a handling system (operating system) which restructures the information by highlighting some features and inhibiting others, thereby constituting a specific organization of the relevant information that constitutes a new type of concept.

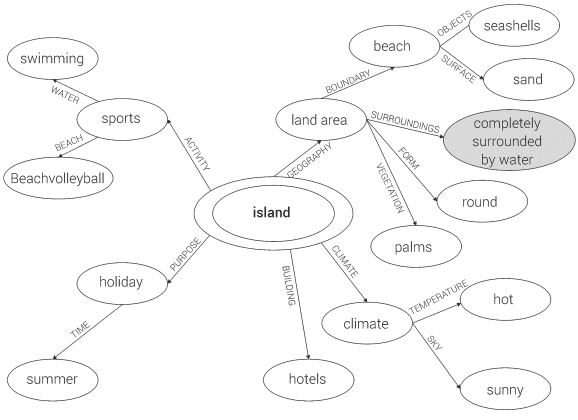

Figure 2: *The definitional concept of ISLAND after the reorganization*

7. Further challenges for the proposal of concepts as mental files: Abstract concepts, natural kind concepts and the flexible use of concepts

Concerning the development of mental files, we have already mentioned that infants start by representing perceptual information associatively. When they learn a natural language, they learn to structure this information and to combine it with (at least minimally) complex inferential roles, in relation to other words and sentences. These inferential roles when combined with a word are the main characteristic of descriptive information associated with the word.[25] If we want to highlight the different roles of perceptual information **(PI)** and descriptive information **(DI)** in a mental file, we can characterize the mental file as having the same associative informational basis, while the handling system produces a different template, activated on the basis of the associated information: one template focuses on the cluster of perceptual information as the individuating part, while the other focuses on an descriptive element as the individuating part. Given this perspective, it makes an essential difference whether the concept is determined by features of the perceptual information or features of the descriptive information. Furthermore, if we presuppose that mental files are usually anchored in the world, i. e. they are related to objects, properties etc., we can also account for natural kind concepts. The following presentation of the concept ISLAND as a mental file includes all three components which can be chosen by a handling system to become

[25] It is important to remember that the introduction of descriptive information does not simply correspond to a change of format, as in the stronger reading of Dretske discussed in section 2.

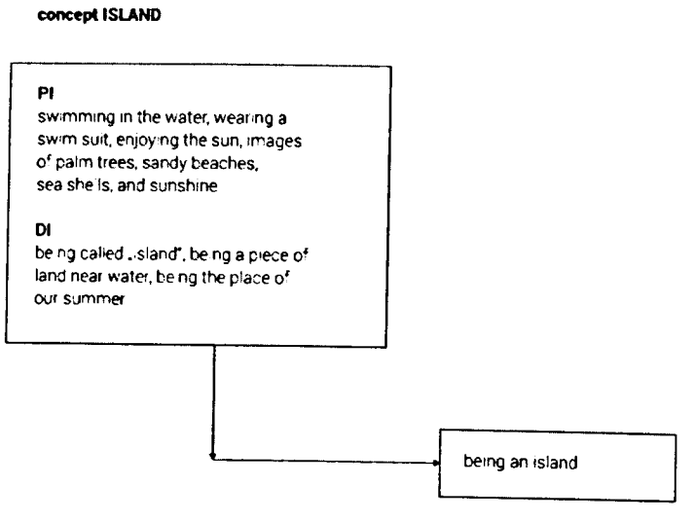

Figure 3: *The concept ISLAND described as a mental file*

dominant: the perceptual information (PI), descriptive information (DI) and the anchoring relation.

If a 4-year-old child understands the word »island« as a characteristic-feature concept, then the objects falling under it are determined by the cluster of perceptual features (PI) that directly allow categorization by pattern matching, which come with typicality effects. If a 9-year-old understands the word »island« as a definitional concept, then the objects are categorized only on the basis of one critical descriptive feature, namely »being a piece of land surrounded by water.« If someone understands the word »island« as expressing a natural kind, then the concept is determined by a causal anchoring relation to concrete samples of islands, which determines the kind as having an essential property. We can now easily reconstruct how one and the same integrated associative information in a mental file can be used as the basis for different concepts, since it can be transformed by the handling system into one of the three concept types: a characteristic-features concept, a definitional concept, and a natural kind concept.

Let us take stock: concepts, we have argued, are templates based on mental files through the activation of associated information that is structured by a handling system. Within this framework, we can account for the third criterion of adequacy, i. e. the transformation from characteristic-feature concepts into definitional concepts. Furthermore, we can account for different types of concepts expressed by the same word, e. g. the natural kind concept WATER is accounted for by the fact that the anchoring relation to samples of water is the determining relation. Using this example, we can also integrate the difference between folk concepts and scientific concepts: this additional

difference (on top of having a natural kind concept) mainly concerns the type of descriptive information involved and activated in the relevant template. Let us presuppose that in our language the concept WATER is understood as a natural kind concept, i. e. the mental file is structured such that some samples of water are taken to determine the concept. Then it is still underdetermined which descriptive information is available and used. If the descriptive information of a natural kind concept involves superficial features only, like coming out of taps, being capable of satisfying someone's thirst, then this is a folk concept. On the other hand, if the descriptive information includes and focuses only on having the chemical structure H_2O, then we have a scientific concept.[26] This is the difference as illustrated for a natural kind concept.

Another aspect of the framework of mental files is that the plurality of concepts expressed by one word of a language is not only a product of long-term changes, but can also take place in the understanding of one sentence in one situation as a result of the flexibility of the handling system: while pointing to a plate with a vegetarian sausage on it, I utter the sentence »This sausage is not a sausage.« The sentence is usually interpreted as expressing that »the object that looks like a sausage« (i. e. the sausage in the characteristic-features sense) is not a sausage in the definitional sense. Thus, by understanding the sentence, we change our understanding of the very same word. In the first instance, we understand »sausage« as a characteristic-features concept, activating the typical perceptual features, while in the second instance, we activate a definitional concept like »being an item of food in the form of a cylindrical object which contains finely chopped and seasoned meat usually stuffed into a prepared animal intestine.« With this situational change, which activates different concepts with the same word in one sentential utterance, we arrive at a non-contradictory interpretation of the sentence. Thus, our theory accounts for two important observations: we can develop a variety of concepts on the basis of one corpus of associated information, and we can use these different concepts rather flexibly, even when they are expressed by the same word.

Furthermore, this framework allows us to account for *abstract concepts*, including concepts expressed by theoretical terms like the concept GENE or ELECTRON. In the case of scientific concepts, as mentioned above, the concept is often predominantly descriptive. We need not exclude perceptual information (PI) from being represented in these scientific concept files; we need only rule out that perceptual information shares the relevance of descriptive information (DI) for the specific inferential network built at this

[26] A detailed general characterization of scientific concepts, in contrast to folk concepts, would go beyond the scope of this article.

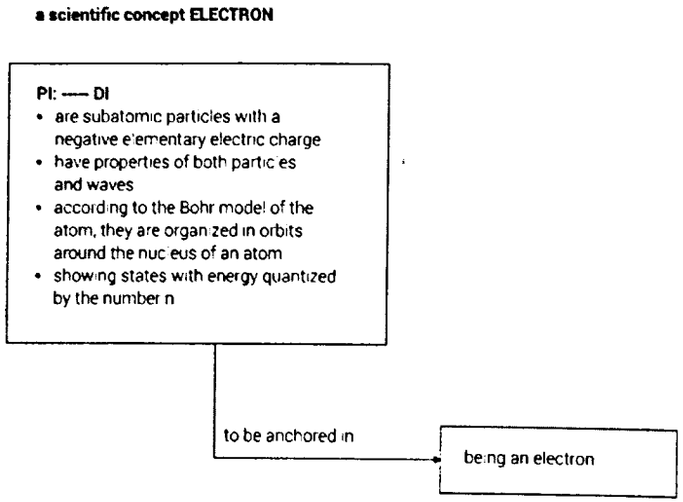

Figure 4: *The concept ELECTRON described as a mental file*

stage.[27] Through scientific theorizing, it is possible to refine the descriptive-information contained in (a set of) mental files. Such refinement allows for an even more complex handling of the stored information, and a consequent further improvement of the overall inferential network in which mental files are embedded. Accordingly, we are able to form concepts that unify only descriptive information in the relevant mental file. Then, the concept is determined by this descriptive information and its role in the inferential network, characterized for natural language representations. An example is the following mental file of the concept ELECTRON, which characterizes the standard representation that occurs when one learns this concept at school.

We may call concepts that rely only on descriptive information »purely descriptive concepts.« In the case of purely descriptive concepts, the possibility of establishing connections among different files no longer depends on the characteristic-features information encoded in such files, as was the case in previous stages. This mirrors a feature of natural language: the inferential network established among linguistically expressed concepts becomes step by step more independent of perceptual features and exponentially more powerful. During ontogeny, children first learn concepts by developing files based on perceptual information, while later the same files involve more and more descriptive information, as we illustrated in the transformation from characteristic-features concepts to definitional concepts. On the basis of a

[27] This could be modeled via Bayesian weighting.

rich inferential network, we can learn to understand descriptive concepts, including purely descriptive concepts like ELECTRON.

8. A comparison with the main alternative theories of concepts: Why prefer the present approach?

As a final consideration, we will discuss how the model developed during the course of this paper offers several advantages over some of the main alternative accounts of concepts in the literature. We set out by claiming that the aim of a theory of concepts is, on the one hand, to account for the cognitive capacities of categorization and recognition of humans and certain animals, and, on the other hand, to explain certain conceptual changes during ontogenetic development.

In advancing our proposal, we have been tracing a viable path between two opposing views. One view holds that concepts are completely different from perceptual representations, a position discussed here in the form of the stronger reading of the Dretskean distinction between the analog and digital representational formats. The opposite view holds that concepts are essentially identical with some types of perceptual representations, where the latter include sensorimotor representations as well as imagistic representations (Barsalou 1999, 2005; Prinz 2004). We believe that both of these opposing views have significant shortcomings that our theory overcomes.

The empiricist view relies on evidence that the understanding of words is correlated with sensorimotor brain activation, e. g. action verbs come with specific sensorimotor activations in the brain, including the specific activations of body parts in the somatosensory cortex, when listening to verbs such as: »to kick«, »to pick«, and »to lick« (Haug et al. 2004). Partially siding with the empiricist stream developed into theories by Barsalou and Prinz, we can criticize the Dretskean notion of a digital format, intended to cover all concepts: this notion is not able to account for experience-based information-rich structures, including perceptual information, which according to recent studies are involved in several concepts expressed in natural language. This information-rich structure is exactly what Dretske wants to account for with the analog format of representation, while analog representation remains – according to Dretske – in the area of nonconceptual representation. Thus, the Dretskean notion of a digital format is far too specialized to account for all concepts; it seems to account only for what we called purely descriptive concepts. Furthermore, the radical version of the distinction between the analog and the digital format sets up an excessively strict demarcation between perception and concepts. If we want to explain certain forms of prototypical categorization based on similarity of stimuli with typicality effects, we need

to allow that the concepts involved will have an information-rich structure, which represents and binds together different perceptual features of the same object at the same time. Again, such informational richness is not part of the digital format of representation, which cannot adequately account for the typicality effects of concepts. Furthermore, this strict demarcation would not allow us to describe the continuous transformation from characteristic-feature to definitional concepts.

On the other hand, a purely empiricist account is also inadequate: first, there is empirical evidence that abstract verbs, such as »to think«, »to comprehend«, and »to consider« do not involve motor areas in the brain, while concrete verbs do to a considerable extent (Rueschemeyer et al. 2007). Moreover, while in normal contexts action verbs (e. g. »to kick«) are correlated with a strong recruitment of motor and premotor cortices, this is not the case if the same action verbs are used in an idiomatic context (e. g. »to kick the bucket«) (Raposo et al., 2009). These data pose a considerable challenge for all radically empiricist accounts.[28] We have argued that we can construct different types of concepts, such as characteristic-feature concepts, definitional concepts, and natural kind concepts, on the basis of the same associative network of perceptual information correlated with a word. Thus, a concept cannot be constituted by a network of associated features alone, and we therefore need to presuppose a handling system that enables us to reorganize the same basic information. Another deficit of empiricist accounts is that they cannot easily account for the construction of abstract concepts (like ELECTRON or GENE) that rely only on descriptive information. Last but not least, the empiricist (and alternative) accounts do not offer a framework for describing the shift in categorization strategies from characteristic-feature concepts to definitional concepts, which happens from the age of 4 to 9. A theory that treats concepts as perceptual symbols (Barsalou 1999, 2008, 2009) does not include definitional concepts, and especially not purely definitional ones. In contrast, our theory delivers a graded view of concepts that reflects cognitive development from infancy to adulthood, and is tightly anchored in both perceptual and inferential capacities at different stages of development. Therefore, we are able to retain similarity as a core feature of conceptual representations and deal with typicality effects, while at the same time explain

[28] The debate about the neural representation of abstract words is not settled. But even the most recent findings on abstract verbs are in line with the minimal claim that abstract verbs are less bounded to sensorimotor representations than concrete verbs (Sakreida et al. 2013). On the other hand, it seems that abstract words remain loosely anchored in sensorimotor representations (Sakreida et al. 2013; Wilson-Mendenhall 2013). But this does not imply that in abstract words these sensorimotor representations still play the same important role as in concrete words.

increasingly complex forms of categorization, including purely descriptive concepts, based on complex inferential reasoning.

9. What is new?

How is this different from other empiricist accounts? In developing Barsalou's account further, we introduce a second component which is constitutive for concept possession, i. e. a handling system that organizes and reorganizes the relevant associative network. In contrast to both Prinz (2004) and Barsalou (1999, 2012), the associative network of information is organized in a mental file which can contain not only *perceptual* information but also *descriptive* information. This allows us to account for the difference between perception-based concepts and abstract concepts like GENE. In contrast to Recanati's theory of mental files (2012), we can not only account for singular concepts but also for general concepts, and we are able to specify the basis of purely perception-based concepts much more clearly than he is able to do.

Someone who accepts our analysis and arguments may wonder whether we are doing any more than reactivating the ideas of inferential role semantics. It is correct that we characterize the descriptive information expressed by natural language terms mainly in terms of inferential relations. This is the same as in inferential role semantics; however, a mental file usually contains perceptual information as well. As we have seen with the 4-year-olds' characteristic-feature concept of ISLAND, a concept can be based on perceptual informaton alone before the transformation takes place. Furthermore, we allow for perceptual generalizations that are more basic than language-based inferences. Finally, we can integrate natural kind concepts like WATER, which include the causal anchoring relation to samples of water, as a constitutive part of the concept. Thus, contrary to inferential-role theories, we do not propose a purely internalistic account of concepts.

10. Conclusion

We have argued for the following main thesis concerning concepts: Concepts are constituted by two components: (i) an integrated network of relevant information associated with an object, property, or other entity; and (ii) a handling system which organizes and reorganizes this associative network. These two components are implemented in a mental file stored in long-term memory. Loosely speaking, concepts are mental files (as a unity of associated information with a minimal structure). More precisely, concepts are templates with a specific organization (i. e. handling) of the relevant associated

information activated in a situation (by working memory) based on mental files. With this account, we offer an alternative to both empiricist and rationalist theories of concepts. In line with empiricist theories (and in opposition to rationalist theories), the framework we propose allows in principle for a close link between conceptual development and perceptual experience. For, according to our view, a fundamental representational basis is shared between the cognitive capacities of recognition and categorization on one hand, and perceptual experience on the other. The shared basis is the associative network of a mental file. In line with rationalist theories, our account has an additional component, i. e. the handling system, which distinguishes concepts from clusters of perception-based features. With our two-component theory of concepts, we can account for recognition and categorization, and also for the transformation of concepts in ontogeny. The latter is a specific advantage of our proposal. We have also illustrated that we can do justice to many of the standard phenomena pertaining to concepts, including abstract concepts, the difference between folk and scientific concepts, and many more. Thus, we hope to have introduced an alternative framework for a fruitful theory of concepts.

References

Allen, C.: Animal Concepts Revisited: The Use of Self-Monitoring as an Empirical Approach. In: *Erkenntnis* 51, 1999, pp. 33–40.

Arterberry, M.E./Bornstein, M.H.: Three-Month-Old Infants' Categorization of Animals and Vehicles Based on Static and Dynamic Attributes. In: *Journal of Experimental Child Psychology* 80(4), 2001, pp. 333–346.

Baillargeon, R.: Object permanence in 3.5 and 4.5-month-old infants. In: *Developmental Psychology* 23, 1987, pp. 655–664.

Baillargeon, R.: The object concept revisited: New directions in the investigations of infants' physical knowledge. In: *Visual perception and cognition in infancy*, ed. by C.E. Granrud. Hillsdale, NJ 1993, pp. 265–315.

Barsalou, L.W.: Perceptual symbol system. In: *Behavioral and Brain Science* 22(4), 1999, pp. 637–660.

Barsalou, L.W.: Continuity of conceptual system across species. In: *Trends in Cognitive Sciences* 9(7), 2005, pp. 309–311.

Barsalou, L.W.: Cognitive and neural contributions to understating the conceptual system. In: *Current Directions in Psychological Science* 17(2), 2008, pp. 91–95.

Barsalou, L.W.: Simulation, situated conceptualization, and prediction. In: *Philosophical Transactions of the Royal Society* 364, 2009, pp. 1281–1289.

Barsalou, L.W.: The human conceptual system. In: *Cambridge Handbook of Psycholinguistics*, ed. by M. Spivey, K. McRae & M. Joanisse. Cambridge, MA 2012, pp. 239–258.

Carey, S.E.: Bootstrapping and the origin of concepts. In: *Daedalus* 133, 2004, pp. 59–68.

Chater, N./Heyes, C.: Animal Concepts: Content and Discontent. In: *Mind & Language* 9, 1994, pp. 209–246.

Davidson, D.: Rational Animals. In: *Actions and Events*, ed. by E. LePore/B. McLaughlin. Oxford 1985, pp. 473–480.

Dretske, F.: *Knowledge and the flow of information*, Cambridge, MA 1983.

Farivar, R.: Dorsal-ventral integration in object recognition. In: *Brain Research Reviews* 61(2), Oct 2009, pp. 144–53. doi:10.1016/j.brainresrev.2009.05.006. Epub 2009 May 28.

Fodor, J.: *The Language of Thought*, Cambridge, MA 1975.

Fodor, J.: *Concepts. Where Cognitive Science Went Wrong*, Oxford 1998.

Gärdenfors, P.: *Conceptual spaces*, Cambridge, MA 2000.

Glock, H.J.: Animals, Thoughts, and Concepts. In: *Synthese* 123, 2000, pp. 35–64.

Glock, H.J.: Animal Minds: Conceptual Problems. 2008. http://www.philosophie.uzh.ch/seminar/lehrstuehle/theoretische2/team/glock/Glock_Animal-Minds.pdf.

Gordon R.D./Irwin D.E.: What's in an object file? Evidence from priming studies. In: *Percept Psychophysiology* 58(8), Nov 1996, pp. 1260–77.

Hauk O./Johnsrude I./Pulvermüller F.: Somatotopic Representation of Action Words in Human Motor and Premotor Cortex; *Neuron* 41 (2), 2004, pp. 301–307.

Kahnemann, D./Treisman, A./Gibbs, B.J.: The Reviewing of Object Files: Object-Specific Integration of Information. In: *Cognitive Psychology* 24, 1992, pp. 175–219.

Keil, F.C.: *Concepts, kinds, and cognitive development.* Cambridge, MA 1989.

Keil, F.C.: *Concepts, Kinds, and Cognitive Development.* Cambridge, MA 1992.

Kinzler, K.D./Spelke, E.S.: Core Systems in Human Cognition. In: *Progress in Brain Research* 164, 2007, pp. 257–264.

Lakoff , G.: *Women, Fire and Dangerous Things.* Chicago 1987.

Leonardi, P.: Review of »F. Recanati: *Mental Files*« Notre Dame Philosophical Reviews 2013. http://ndpr.nd.edu/news/41180-mental-files/.

Löbner, S.: Functional Concepts and Frames. In: *Meaning, Frames, and Conceptual Representation. Studies in Language and Cognition 2*, ed. by T. Gamerschlag, D. Gerland, R. Osswald, W. Petersen. Düsseldorf University Press, 2015, pp. 13–42.

Luo, Y./Baillargeon, R./Breuckner, L.: Reasoning about a Hidden Object After a Delay: Evidence for Robust Representations in 5-Month-Old Infants, *Cognition* 88, 2003, pp. 23–32.

Macpherson, F.: Cognitive penetration and non-conceptual content. In: *Cognitive Effects on Perception: New Philosophical Perspectives*, ed. by J. Zeimbekis & A. Raftopoulos (in press).

Machery, E.: *Doing without concepts.* Oxford 2009.

Newen, A.: Phenomenal Concepts and Mental Files: Phenomenal Concepts are theory-based. In: *philosophia naturalis* 47/48, 2010/11, pp. 155–183.

Newen, A./Bartels, A.: Animal minds and the possession of concepts. In: *Philosophical psychology* 20, 2007, pp. 283–308.
Noles, N.S./Scholl, B.J./Mitroff, S.R.: The persistence of object file representations. In: *Perception & Psychophysics* 67(2), 2005, pp. 324–334.
Pauen, S.: Early differentiation within the animal domain: Are humans something special? In: *Journal of Experimental Psychology* 75, 2000, pp. 134–151.
Peacocke, C.: Analogue Content. In: *Proceedings of the Aristotelian Society* 60, 1986, pp. 1–17.
Peacocke, C.: *A Study of Concepts*, Cambridge, MA 1992.
Pepperberg, I.: *The Alex Studies*, Cambridge, MA 1999.
Perry, J.: Self-Notions. In: *Logos* 11, 1990, pp. 17–31.
Perry, J.: The Self, Self-Knowledge, and Self-Notions. In: *Identity, Personal Identity and the Self*, ed. by J. Perry. Indianapolis 2002.
Perry, J.C./Fallah, M.: Feature integration and object representations along the dorsal stream visual hierarchy. In: *Frontiers in Computational Neuroscience* 8(84), Aug 2014, pp. 1–17. doi:10.3389/fncom.2014.00084. eCollection2014.
Prinz, J.: *Furnishing the Mind*, Cambridge, MA 2004.
Putnam, H.: The Meaning of ›Meaning‹. In: *Minnesota Studies in the Philosophy of Science* 7, 1975, pp. 131–193.
Rakoczy, H.: Pretend play and the development of collective intentionality. In: *Cognitive Systems Research* 7, 2006, pp. 113–127.
Rakoczy, H./Warneken, F./Tomasello, M.: The Sources of Normativity: Young children's awareness of the normative structure of games. In: *Developmental Psychology* 44(3), 2008, pp. 875–881.
Raposo A./Moss H.E./Stamatakis E.A./Tyler L.K.: Modulation of motor and premotor cortices by actions, action words and action sentences. Neuropsychologia 47(2), 2009 pp. 388–396. doi:10.1016/j.neuropsychologia.2008.09.017.
Recanati, F.: *Mental Files*, Oxford, UK, 2012.
Rosch, E./Simpson, C./Miller, R.S.: Structural bases of typicality effects. In: *Journal of Experimental Psychology*, 2(4), 1976, pp. 491–502.
Rosch, E.: Principles of Categorization. In: *Cognition and Categorization*, ed. by E. Rosch & B.B. Lloyd. Hillsdale 1978.
Rüschemeyer, S.A./Brass, M./Friederici, A.D.: Comprehending prehending: neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience* 19(5), 2007, pp. 855–865. doi:10.1162/jocn.2007.19.5.855.
Sakreida K./Scoroll C./Menz, M.M./Heim, S./Borghi, A.M./Binkofski, F.: Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. In: *Frontiers in Human Neuroscience* 7(125), 2013, pp. 1–13. Published online 2013; doi:10.3389/fnhum.2013.00125.
Slater, A./Quinn, P.C.: Face Recognition in the Newborn Infant. In: *Infant and Child Development* 10, 2001, pp. 21–24.
Sodian, B./Thoermer, C./Metz, U.: Now I see but you don't: 14-months-olds can represent another person's visual perspective. In: *Developmental Science* 10, 2007, pp. 199–204.

Treisman, A.: Perceiving and re-perceiving objects. In: *American Psychologist* 47, 1992, pp. 862–875.
Vetter, P./Newen, A.: Varieties of Cognitive Penetration in Visual Perception. In: *Consciousness & Cognition* 27, 2014, pp. 62–75.
Wilson-Mendenhall, C.D./Simmons, W.K./Martin, A./Barsalou, L.W.: Contextual processing of abstract concepts reveals neural representations of nonlinguistic semantic content. In: *Journal of Cognitive Neuroscience* 25, 2013, pp. 920–935.