Philosophical Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cphp20

Conceptualization for intended action: A dynamic model

Mauri Kaipainen, Antti Hautamäki & Joel Parthemore

To cite this article: Mauri Kaipainen, Antti Hautamäki & Joel Parthemore (06 Jan 2023): Conceptualization for intended action: A dynamic model, Philosophical Psychology, DOI: 10.1080/09515089.2022.2164263

To link to this article: https://doi.org/10.1080/09515089.2022.2164263

Published online: 06 Jan 2023.

PHILOSOPHICAL PSYCHOLOGY
https://doi.org/10.1080/09515089.2022.2164263

Conceptualization for intended action: A dynamic model

Mauri Kaipainen[a], Antti Hautamäki[b] and Joel Parthemore[c]

[a]Department of Cognitive Science, University of Helsinki, Helsinki, Finland; [b]Department of Social Sciences and Philosophy, University of Jyväskylä, Jyväskylä, Finland; [c]Department of Cognitive Neuroscience and Philosophy, University of Skövde, Skövde, Sweden

ABSTRACT

Concepts are the building blocks of higher-order cognition and consciousness. Building on Conceptual Spaces Theory (CST) and proceeding from the assumption that concepts are inherently dynamic, this paper provides historical context to and significantly elaborates the previously offered Iterative Subdivision Model (ISDM) with the goal of pushing it toward empirical testability. The paper describes how agents in continuous interaction with their environment adopt an intentional orientation, estimate the utility of the concept(s) applicable to action in the current context, engage in practical action, and adopt any new concepts that emerge: a largely pre-intellectual cycle that repeats essentially without interruption over the conceptual agent's lifetime. This paper elaborates utility optimization by establishing three constraints on concept formation/evaluation – non-redundancy, distinctiveness, and proportionality – embedding them in a quasi-mathematical model intended for development into a formal logic. The notion of a distinctor – a quality dimension of the conceptual space in focus at any given time, used for making what we call a *difference distinction* – is key. The primary contribution of the revised ISDM is the way it relates concepts to action via utility optimization/actualization and the way it describes the emergence of quality dimensions through trial-by-action (trial and error), something previous presentations of CST have failed to address.

ARTICLE HISTORY
Received 12 October 2021
Accepted 27 December 2022

KEYWORDS
Concepts; conceptual agency; conceptual dynamics; distinctors; Iterative Subdivision Model; Conceptual Spaces Theory

1. Introduction

This paper characterizes concepts as the prelinguistic building blocks of systematically and productively structured thought, facilitating a flexible response to one's environment based on consideration of past experiences and anticipation of experiences to come. How they do so can be described in part via an iterative cyclical model that captures something of the dynamics of conceptualization: the process by which concepts emerge, develop, and eventually decay or get replaced, all the while regulated by agents' intentions

CONTACT Joel Parthemore joel.parthemore@his.se University of Skövde, Daltorps skola LGH 1004, Skövde 54192, Sweden

in their changing circumstances. A conceptual agent is simply any agent able to reason about her environment in a systematically and productively structured fashion: i.e., conceptual agents are a subset of cognizing agents. "Intention" as we use it in this paper is midway between the two standard philosophical uses of the term: Franz Brentano's (2015) notion of mere directedness or aboutness and (self-)consciously guided motivation (Parthemore, 2011, p. 44).

We see intention as a general pointer to the motivating forces of an agent, given her embeddedness (or *situatedness*) in a particular physical and social context and embodiment in a particular physical form, deeply rooted in her values, goals, and personal history. It captures her ability to mean something, achieve something (*conation*), and direct her attention toward something with purpose.

John Searle (1980) proposes two components to actions – mental and physical – where the mental part is intention. Our interest here lies specifically with the *intention to act to bring about change in interaction with one's environment*. Intention is every bit as much of a process as a thing; what it is *not* is a static entity, and attempts to reify it should generally be resisted. In the following discussion, we use "intention" as shorthand for "intention to act".

The paper proceeds by fleshing out details of Mauri Kaipainen and Antti Hautamäki's (2019) *Iterative Subdivision Model* (ISDM), which assigns intention to a key role in conceptual agency. It regards concepts, on the one hand, largely as *artifacts* of intended action, on the other as *instruments* for it (to the extent that the two roles can be separated). Following Kaipainen and Hautamäki, we consider the main function of concepts to be instruments for flexibly chosen action: means to target actions to sets of similar concrete or abstract objects (think of handles that allow grasping by hand), where similarity as per Nelson Goodman (1972) is not determined *a priori*. Rather, perceived similarity depends on perspective, optimized for situation-specific,[1] intentionally driven, moving goals. We describe this optimization for intention in terms of continuously (albeit imperfectly) maximized utility for the intended action, with feedback from previous actions (re-)shaping subsequent actions.

Like any model (Giere, 2006), the ISDM is based on theoretical assumptions – axioms, if you will, including the central role of intention and the inherently dynamic nature of concepts. These assumptions are drawn primarily from conceptual spaces theory (CST; Gärdenfors, 2004, 2014)–itself based on prototype theory (Rosch, 1973)–and action theory (Anscombe, 2000). By their nature, no models offer a complete description of a phenomenon, but good ones demonstrate reasonable similarity with regard to relevant aspects (Hautamäki, 2020).

1.1. The dynamics of conceptualization

Our grounding assumption is that concepts – along with the conceptual frameworks they form part of – are dynamically evolving entities, *pace* Immanuel Kant, with his fundamentally static idea of pure (*a priori*) concepts, and Jerry Fodor, with his informational atomism account (Fodor, 1998) whereby concepts cannot change, because what makes a concept is what it tracks in the mind-independent world. By contrast, Joel Parthemore (2019, p. 86) writes that concepts are "in a state of *continuous* (if often only incremental) change," relatively yet not *too* stable, while those that "completely cease to be open to change are, metaphorically speaking . . . dead" (emphasis original). What this means is that concepts are inherently context-dependent, adapting to each new context they encounter, always different from the last. Observe how this extends the relatively weak notion of dynamics in CST, where "the concepts generated by such a categorization mechanism [one that creates categories based on a small number of examples, as does the ISDM] are *dynamic* in the sense that when the agent observes a new item in a category, the prototype for that category will, in general, change somewhat . . ." (Gärdenfors, 2014, pp. 42–43; emphasis original).

This opens an opportunity for empirical testing. If sufficient circumstantial evidence can be found of concepts that do *not* change, that would tell strongly against our version of CST. So far, claims about fixed concepts (PI is often offered as a prime candidate) have remained solely the domain of philosophical debate, if not simply presented as fiats.

While it is generally acknowledged within cognitive science that most concepts *can* change, evolution is rarely made a central characteristic (though see Vosniadou, 1994), never mind the possibility considered that change may, as *per* Parthemore (2019), be obligatory – if often invisible to the observer and only recognizable by reconstruction. The challenge we take up is to strike our own balance between the conceptual extremes of overly capricious (think Barsalou, 1987, with his temporary concepts) and overly stable: that is, to account for the dynamic nature of concepts such that they are agile enough to adapt to contextual changes, stable enough to apply across unboundedly many contexts (though, as Gärdenfors (2014, pp. 41–42) notes, "the cost of generality is the increase of error"), while providing enough continuity to facilitate long-term cognition and intersubjective communication.

Assuming that concepts are essentially dynamic, what steers their evolution? Our model assumes the complex interaction of top-down- and bottom-up-driven processes such that the results cannot, even in principle, be linearized except for purposes of simplified modeling, as we attempt here. Our focus is on a particular top-down process that progressively narrows the target of concepts arrived at via what we call *difference distinctions*: intentionally established quality dimensions establishing finer- and finer-grained – or altogether new – distinctions that make a difference, driven by circumstances.

Our approach follows a methodology common in cognitive science, relying on rigorous philosophical argumentation backed wherever possible by empirical findings from psychology. Intentional orientation drives optimization of concepts for utility in practical action, involving an essentially phenomenological element. First-person embodied experience of action and consequences evaluates conceptual utility, regulating adoption or adjustment of concepts accordingly to provide intentional orientation going forward, *ad infinitum*. The first half of the cycle, intentional orientation – where we focus in this paper – leans on empirical evidence concerning the nature of categorical perception (the perception of distinct categories within what closer observation reveals to be a smooth continuum) and concepts-as -prototypes elaborated since Rosch (1973)–a vast domain over which Harnad (1987) provides a critical overview.

While the claim to dynamicity may indeed open the door to empirical investigation, empirical investigation of theories of concepts in general remains a tricky proposition. The problem is that, on most accounts, concepts are creatures of thought and not directly observable: one can reflect on them, reason about them, but – at least as things presently stand– one cannot *see* or measure them.[2] That reduces one to observing them by their effects, such that an accumulation of indirect, circumstantial evidence tells in favor or against a theory (Parthemore, 2015).

Gärdenfors' version of CST at least makes the explicit effort of opening itself to empirical investigation, which is to be lauded given the lack of such efforts from other accounts. That said, CST as a whole – in any of its incarnations that we're aware of – is not presently empirically testable, never mind falsifiable.[3] Gärdenfors' CST makes one clear empirically testable claim that we're aware of – that concepts, understood as prototypes, should show evidence of convexity in most instances – which has found support in a cross-cultural study of color terms (Jäger 2010). As Parthemore (2015) notes though, that claim is implicit in most if not all prototype-based theories of concepts; the genius of CST lies in making the claim explicit through the language of geometry.

By insisting on the inherently dynamic nature of concepts (a point on which we take Gärdenfors to be agnostic), giving a context-sensitive account of similarity rather than taken similarity as given (as Gärdenfors appears to do), and elaborating on the previously offered ISDM, our goal is to push CST in a more empirically testable direction (not, as an overly quick read might suggest, to offer a competing account). The assumption here is that the more detailed an account, the more opportunities it presents for being demonstrably wrong.

1.2. Concepts' sensorimotor basis

There is clear evidence for the brain as a system that maintains dynamic traces of sensorimotor activations such that these activations form a necessary foundation for mind. Based on those activations, *memories* (which one might interpret equally as *concepts*) emerge as procedural patterns of activity rather than fixed representations retrieved from storage (see, e.g., Versace et al., 2014, pp. 282–283). Indeed, our approach is at heart *non*-representational. In our thoughts and in our communications, concepts point out their referents in the manner that a pointing gesture ostensively defines its target, with direct continuity from the latter to the former (Gärdenfors, 2014, pp. 81–82).

We intend the term "sensorimotor" in the now common way to refer to that which is both sensory experience and motor action, such that the two are inextricably intertwined: rather than sensory experience followed (or preceded) by motor action, there is a complex, circular web of sensorimotor engagements between agent and environment; see e.g., (Bishop & Martin, 2014). One of the upshots of the sensorimotor approach is that even the most abstract and "high-level" of cognition is grounded in (and generally not far removed from) simple, immediate sensorimotor engagements; see, e.g., Barsalou and Goldstone (1998).

A key assumption here is that qualities are derived from experiential qualia[4] translated into conceptual quanta via sensorimotor experience. In the pursuit of understanding matters that ultimately outstrip human capacity to understand – insofar as they would require the human mind to capture an understanding of itself *within itself* completely and consistently – it simplifies matters greatly to assume a matrix of quantitative data standing in for accumulated qualitative experience (hereafter *experience data*).[5] This experience data is the raw material of what the consciously reflective mind can begin to recognize as cognition. Such a simplifying approach seems to us to be the only way forward.

Body-related dimensions – projecting the body into the world – have long facilitated the quantification of experience (Dehaene & Brannon, 2011). Consider the ancient units of inches (the length of a finger joint), feet (the length of a foot) and fathoms (the fingertip-to-fingertip length of outstretched arms) suggest.

Evan Thompson (2007, p. 7), writing of the original computers (as the term was used), who were people of a certain profession, not artifacts, has this to say:

> This kind of physical symbol system is a sophisticated and culturally specific form of human activity. It is embodied, requiring perception and motor action, and embedded in a sociocultural environment of symbolic cognition and technology. It is not bounded by the skull or skin but extends into the environment. The environment, for its part, plays a necessary and active role in the cognitive processes themselves . . . .

The roots of our approach lie in characterizations of holistic mind and body as one, ultimately indivisible whole, itself part of one, ultimately indivisible whole with its environment — as well as the philosophical tradition of American pragmatism represented by (among others) Charles Sanders Peirce, John Dewey, and William James; for a good overview of Dewey's pragmatism, see (Dewey, 1925); for that of James, see (Trigoni, 2015). Cognition extends from the brain through the body into the environment via a phylogenetic continuum from ancestor species to *homo sapiens* (see Gärdenfors, 2003; Omicini et al., 2006).

Ontogenically, our account borrows from Jean Piaget (1972), whereby the development of conceptual agency proceeds from highly concrete and immediate to increasingly abstract and distal operations. If one entertains some version of Recapitulation Theory – famously popularized by Ernst Haeckel with the phrase "ontogeny recapitulates phylogeny" (see Richards, 2008)–one may be tempted to reason that *homo sapiens* retains something of the cognitive signature of its predecessors, including *homo erectus*: that is, both are expressions of "*homo faber*," the creative handyman (Ihde & Malafouris, 2019). What *homo sapiens* adds is a seemingly vast deepening of the capacity to reflect before acting; as Peter Gärdenfors (1995, p. 3) writes: "consider the high jumper who mentally penetrates his bodily movements before actually performing the jump".

1.3. The four Es[6]

Consider embeddedness/embodiment, extended mind, and enactivism as a set of matryoshka dolls. Cognition is never free floating but always ***embedded*** in an environment that helps define it – cognition and cognitive environment establish each other like figure and ground. Concepts have no meaning outside the context in which they exist and are applied.

Cognition is ***embodied***, constantly (re-)defining and evaluating the utility of action in physical engagement with one's environment, the results of that engagement serving as the benchmark against which the mind evaluates its concepts. Embeddedness and embodiment are two sides of one coin, each defining the other. Cognition free of body has no meaning. A key inspiration here is George Lakoff and Mark Johnson's (1980) Cognitive Metaphor Theory for the way it grounds the abstract metaphorical nature of thought (a necessary precursor to linguistic metaphor, according to Lakoff and

Johnson) in the embodied agent's spatial and otherwise concrete physical relations to her environment.

Cognition is ***extended*** into its environment. The Extended Mind Hypothesis (Clark & Chalmers, 1998) incorporates but goes well beyond embeddedness and embodiment in finding no clear line between where mind stops and world begins. Indeed, it is precisely at the limits of embeddedness and embodiment, "skin and skull," that critics of Extended Mind (e.g., Adams & Aizawa, 2008; Rupert, 2009) wish to stop. This is not necessarily to say that mind extends in the ways Clark and Chalmers often seem to have in mind, into such tangible extracranial devices as notebooks and mobile phones; our position is more in line with an argument Clark (2008, p. 34) makes in the second chapter of *Supersizing the Mind*: namely, that "profoundly embodied agents . . . are able constantly to negotiate and renegotiate the agent-world boundary itself"–one that may be as much social as physical. It is likewise aligned with Robinson's (2013) approach to Extended Mind, which attempts to move beyond tedious debates over whether mind "literally" extends into notebooks, mobile phones, and so on. Robinson (who embraces Extended Mind) focuses on pre-linguistic artifacts – cranial or extracranial – serving the mind's conative expressions, based on embodied feeling and physical movement.

Finally, cognition is ***enacted***; Humberto Maturana writes (Maturana & Varela, 1992, p. 255): "I have proposed the term enactive to . . . evoke the idea that what is known is brought forth, in contraposition to the more classical views of either cognitivism or connectionism." Just as Extended Mind incorporates but goes well beyond embeddedness and embodiment; so, too, does enactivism incorporate but go well beyond Extended Mind in seeing an underlying continuity between *all* aspects of agent and environment, each bringing the other forth in an act of co-creation. The individual agent as something discrete and separable from that continuity is like a center of gravity: a useful, even necessary way of looking at the world but one without ontological commitments to what the precognitive world is like. What Gärdenfors evocatively describes as a "meeting of minds" at the fixed point – one person's concepts aligning with another's – we see more as an *overlapping* of minds.

Going beyond the four Es, the inspirations for our conceptually dynamic approach include, among other things, Marx's (1941) dialectics, whereby conceptualization proceeds from the concrete to the abstract and back again. Marx builds his approach on Hegel (see, e.g., Fox, 2005), whose dialectics, applying the method of analysis and synthesis, are similarly based on an interplay between abstract and concrete. From there, a discernible line of thought leads to the articulation of General Systems Theory (GST) by von Bertalanffy (1973), Ashby (1956) and others. Among the most basic implications of GST is the essential non-linearity of systems' progression.

2. Conceptualization as endless progression of difference distinctions

Consider the conceptual agent's environment as a source of experience data, some portion of which is conceptualized into objects that are encountered (or imagined to be encountered), events that take place (or might take place), and properties (or *qualities*) of both (Parthemore, 2014c, p. 155)– some of which in turn come to be lexicalized. In the language of CST, those properties can be conceived of as *integral quality dimensions* of *similarity spaces* occupied by concepts that may be concrete or abstract objects, concrete or abstract events, or the properties that describe these objects and events.[7] In these similarity spaces, proximity amounts to similarity: the smaller the metric distance between two points in any given space, the more similar they are taken to be.[8] The shape of these spaces is by no means static but rather in flux depending on the other spaces to which a given space is currently linked: i.e., similarity is strongly dependent on context (Yearsley et al., 2022).

In contrast, the framework offered by CST is largely static, as Gärdenfors (2004, p. 31) openly acknowledges. This leaves the conceptual system it purports to describe unable to adapt adequately to changing circumstances. He offers a way to make the framework dynamic, talking in general terms about processes operating on conceptual representations; we suggest another, more readily dynamic model applicable in the first instance to the nature of the integral quality dimensions defining a similarity space: those dimensions are not in any way pre-given (as Gärdenfors appears to assume) but intentionally chosen in an iterative process that the ISDM attempts to capture. As described by the ISDM, difference distinctions are made at each step of conceptualization within the concept(s) presently in focus by choosing as the next subdivision criterion the quality dimension (henceforth *distinctor*) among potential alternative dimensions expected to serve the active intentions best.

2.1. Top-down intentionality

Both abstractly and in concrete ontogenetic terms, the process starts from broad, generic distinctions within a minimally partitioned conceptual framework (Parthemore, 2017, p. 41) with a minimal number of minimally structured similarity spaces. For the most part, it moves toward increasingly narrow and more focused distinctions, resulting in a concept hierarchy.[9] Figure 1 illustrates the evolving process of iterative difference distinctions whereby object qualities typically (though not always) take graded values

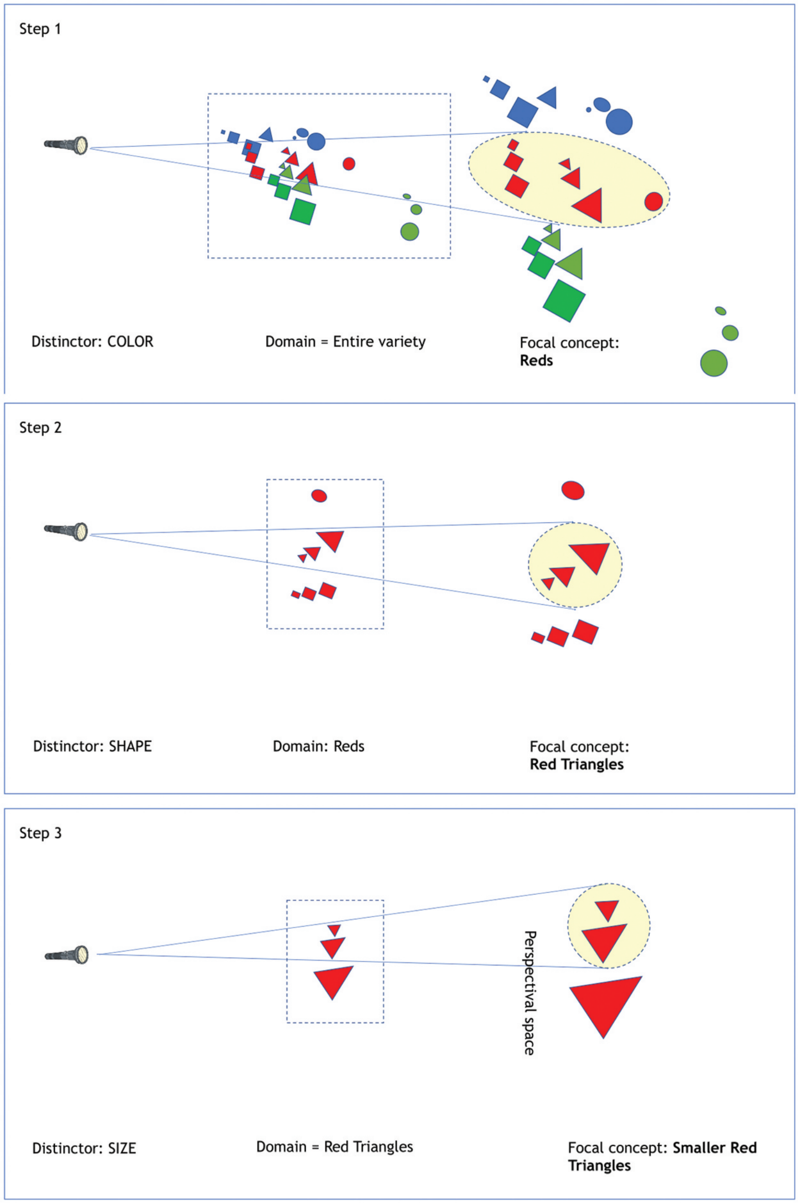

Figure 1. Iterative difference distinctions. Note: Iterative difference distinctions elaborating on the figure from Kaipainen and Hautamäki (2019, p. 106), each step assuming a unique vantage point. Step 1: the distinctor COLOR is applied to the original domain to distinguish BLUES, REDS, and GREENS. Step 2: the distinctor SHAPE is applied to the domain of REDS to distinguish ELLIPSES, TRIANGLES, and RECTANGLES; RED TRIANGLES is chosen as the focal concept. Step 3: The distinctor SIZE is applied to red triangles; SMALLER RED TRIANGLE is chosen as the focal concept.

subject to idiosyncratic judgment. The rationale for distinctor choice will be discussed in the sections on utility optimization (3.2–3.3).

The method is not fundamentally different from that of Swedish botanist Carolus Linnaeus (1735) who attempts to describe the entirety of nature by breaking it down into kingdoms, which he then divides into increasingly finer levels of taxonomy (*taxa*): phyla, classes, orders, families, genera, and species. There are two crucial differences between Linnaeus' approach and ours. First, we assume that the distinctions are motivated by intention in relation to context, whereas Linnaeus takes the category-defining distinctors on each level as given and the categories themselves as something like so-called *natural kinds*: any intelligent extraterrestrial, on encountering terrestrial life for the first time, should come to a similar arrangement. Second, Linnaeus' model is strictly hierarchical; ours, as noted, is not.

The ISDM has been elaborated over the course of a number of publications as a procedural interpretation of what Hautamäki (1986, 2016, 2020) calls *contextual points of view, selections of quality dimensions*, or (*per* Kaipainen & Hautamäki, 2015) *perspectives*. Priorities among quality dimensions are determined via iterative difference distinctions (Kaipainen & Hautamäki, 2019) rather than – as *per* Igor, Douven and Gärdenfors (2018, p. 4) or Kaipainen and Hautamäki (2015)–*prominence weights*. Current perspective – which orients intentions and consequent actions – traces the path of difference distinctions that the conceptual agent has taken so far. It is, as Francisco Varela, Evan Thompson, and Eleanor Rosch (1992) suggest, to "lay down a path in walking." That is to say, there *is* no path until it is walked; each step, in combination with previous steps, determines the next.

Like CST, the ISDM and the framework we build on it here regard concepts as *essentially prelinguistic entities*: they do not require but rather facilitate naming/labeling; this means that lexical concepts are a subset of all concepts. Where CST sees a direct translation from concepts to words of a language ("our words express our concepts"; Gärdenfors, 2014, pp. 21–22), we see a much more complex relationship.

We embrace the language of geometry in which the core elements of CST are framed. CST provides an elegant yet versatile way of picturing concepts as the mediators between, on the one hand, symbolic (including linguistic) and sub-symbolic (or *iconic*) representations; and, on the other, representations of any kind and "mere" associations – just as concepts can be seen as the mediators between *knowledge that* and *knowledge how* (Ryle, 1949), occupying an intermediate level of knowledge beholden neither to the one nor the other (Parthemore, 2011, pp. 25, 37). The most fundamental explanatory tool of CST is that of the aforementioned similarity spaces (Gärdenfors, 2004, p. 193, following Smith & Heise, 1992, p. 252) defined by one or more quality dimensions, each describing a range of values of the quality with respect to which objects occupying the space vary.

Again, in accordance with CST, we take concepts as (mainly) convex regions in similarity spaces (there are important exceptions, such as GENTILE meaning anyone who is not a Jew), building on the same empirical foundations by which CST attempts to justify itself. Those similarity spaces take the form of Voronoi tessellations. However, while CST assumes similarity to be given in some way – if not, in fact, intrinsic in the concepts' non-conceptual referents – our approach attempts to explain how both similarity and similarity spaces come about, and how they evolve as the result of continuous, contextually situated cognitive labor. Specifically, we propose that similarity is determined *solely* by distance within similarity spaces of ever-evolving metrics where, with respect to any given intention, one distinctor is chosen at a time in a difference-distinctive fashion, driving the construction of (in general) increasingly fine-grained and homogeneous similarity spaces in the process of intention-driven conceptualization.

From where does conceptualization begin? It seems safe to assume that certain protoconcepts – basic building blocks prerequisite to concept formation – must be (in some fashion) hardwired (Parthemore, 2014). Human beings seem predisposed to carve up the world in terms of object-type entities, action- or event-type entities, and property-type entities, *regardless of whether the world itself is structured that way*. Human beings are incapable of conceiving themselves and their environment any other way.

Ontogenetically, the initial "true" concepts $\boldsymbol{c_{0a}}$ and $\boldsymbol{c_{0b}}$ might plausibly be a prelinguistic version of SELF and OTHER (Parthemore, 2014, pp. 206–207), forming the basis for all the more complex and precise conceptual distinctions to follow. In any case, what seems clear is that "by the time a child can speak, she . . . has at least the beginnings of a conceptual structure on which to base speech" (LeDoux, 2020, p. 356). As Gärdenfors writes (Gärdenfors, 2014, pp. 65–66), "what happens developmentally is that one domain after another is separated out and can be attended to, indeed, as a separable set of dimensions".

In geometric terms, each concept c_f can be conceived as:

- a single point in a similarity space,
- a (generally) convex subarea of that similarity space, or
- *itself* constituting a similarity space in a mereologically arranged system of embedded similarity spaces, not unlike an endless succession of matryoshka dolls . . .

. . .Both as a subspace of a higher-level space and as the super-space to any number of subspaces distinguishable from one another by means of distinctors, where each distinctor is an intentionally chosen quality dimension. The emergent conceptual framework – a kind of "space of spaces" (Parthemore, 2013b) – can be regarded as a tree-like branching hierarchy of similarity spaces (albeit one with recurrent connections between levels) where each node is simultaneously a concept and a distinctor specifying further subspaces.

So far, we have focused on intentional, top-down-driven conceptualization. However, this apparently linear progression constitutes but one part of what should be understood as not just a causally circular but highly dynamic model. What follows is an attempt to explain at least part of the bottom-up dynamics: namely, how the consequences of embedded, embodied action provides the ultimate evaluation of concepts, qualifying only the most practical as instruments for further use and steering subsequent intention.

2.2. Top-down intention meets bottom-up evaluation

The ISDM (Kaipainen & Hautamäki, 2019) applies Georg Henrik von Wright's (1971, p. 96) notion of *practical syllogism* to describe the iterative process of conceptualization. Consider how two premises lead to an action: (1) agent *A* intends to bring about outcome *p*. *A* considers that she cannot bring about *p* unless she does action *a*. Therefore, *A* determines to do *a*. However, von Wright's syllogism, meant to describe the logic of one-time intentional action, does not lend itself well to describing a continuously iterating process. Its essential shortcoming is that it does not postulate a further step to create a syllogistic chain: in ISDM terms, to evaluate the emerging concept and thereby control its adoption, rejection, or modification.

How then to adapt the syllogism? Assume that the agent intends to reach some goal – a change in her environment – by means of the concepts at hand and the action(s) they point her toward. She performs the action, observes its effects, and *by that observation* filtering up from her perceptions to her pre-reflectively conscious and self-conscious mind, modifies her subsequent intentions: a process of no longer linear but circular causality (Parthemore & Morse, 2010, p. 297) whereby what is cause and what effect is a matter of perspective. Perhaps she concludes that the original goal is impossible, perhaps that it has been achieved, perhaps that it requires modification. She proceeds, one action at a time, *reflecting on her actions only where circumstances compel her to do so* and otherwise proceeding in pre-reflective mode: the waking version, if you will, of sleepwalking.

More concretely, we suggest modifying von Wright's syllogism into a four-step iterative loop (Figure 2) that, properly speaking, does not start or stop anywhere but is continuously in motion from the time the agent is recognizably a conceptual agent:

(1) *Intentional orientation*. Agent *A* intends to bring about goal *g* by focusing on applicable concept c_f (some concept at the appropriate

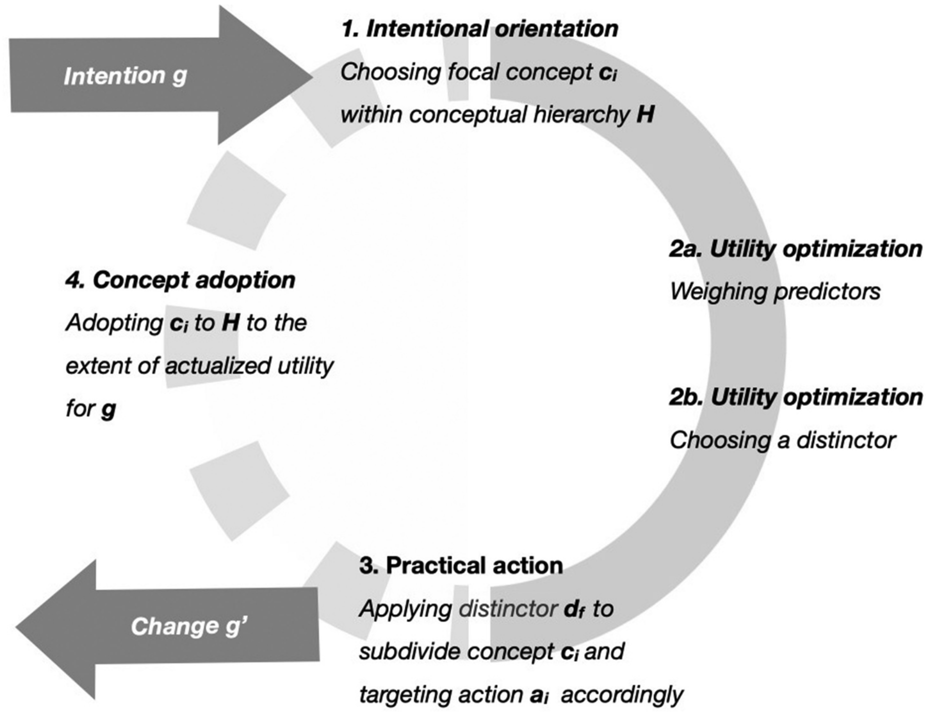

Figure 2. The proposed iterative conceptualization loop.

level of granularity) and considering it for application or refinement. It can be one among the subdivisions resulting from Step 3 or some other concept elsewhere in her overall conceptual framework (because her attention has shifted).

(2) *Utility optimization*. Given a focal concept c_f, A estimates – based on accumulated experience data – whether c_f has sufficient utility for g; and, if not, which distinctor d_f, if any, would allow A to perform further subdivision of concept c_f with the highest predictive utility to bring about g. This goes beyond the merely intuitive inferential consideration suggested by earlier presentations of the ISDM, as elaborated in Section 3.2. This can be broken into steps: considering the possible distinctor(s) to apply and then choosing one.

(3) *Practical action*. A applies d_f to subdivide c_f in the most practical way given the distribution of experience data, targeting action a_i. It may turn out that d_f does not justify a new subdivision so that c_f is applicable as is. It may also turn out, of course, that a poor choice of focal concept has been made, and a new focal concept must quickly be decided upon. After all, conceptual agents clearly *do* make mistakes – frequently!—with what they focus on.

(4) *Concept adoption*. To the extent that c_f proves its expected utility for g in Step 3, A incorporates c_f and its sub-concepts, if any, into

> conceptual hierarchy H. In the event that no subdivision proves useful, c_f's utility may be considered poor and c_f rejected. In case c_f is not at one of the end points of H, it will be in conflict with an earlier traversal of the hierarchy, which will need to be amended – amounting to a more radical re-conceptualization (Parthemore, 2013a, pp. 74*ff*.). Re-conceptualization corresponds to the collapse of some portion of H and the rise of a replacement structure in a manner reminiscent, on a societal level, of a paradigm shift (Kuhn, 1962); the insight is that paradigm shifts can be individual as well as collective. The possibility of such a collapse – indeed, the likelihood of such collapse happening periodically – renders any strictly linear difference-distinctive progression impossible. Re-conceptualization responds to a perceived discrepancy between conceptually anticipated and actual utility for action, breaking the present concept hierarchy at that point and possibly forcing the reconsideration of higher-level distinctions.

1) Intentional orientation accepts intention g as input. 2) Utility is optimized. 3) Practical action a_i is carried out, with consequent change *of* g' in the environment. This serves to evaluate the concept for adoption or rejection, which – depending on circumstances – might or might not result in re-conceptualization. 4) The new concept (if one arises) is incorporated into H, as regulated by its judged utility for g. The dashed line on the left side of the loop signifies the preliminary state of explanation for this predominantly bottom-up, experiential part of the loop.

We now discuss each phase in detail.

3. The iterative ISDM loop

3.1. Intentional orientation

From the point of view of evolution and biology, survival (of life itself if not one's species, of one's species if not one's progeny, of one's progeny if not oneself) appears to be one of the most basic foundations on which conceptual agents build their intentions. One way to approach intention is to regard it as a strategy that has at its core a drive for survival, in a game one plays to win and keep winning for as long as one can. The relevant intentions may be more distally focused and abstract or more immediate and concrete; our focus here is on the latter: bringing about some immediate, observable change in one's environment.

Within the practical-inference framework we are sketching, intention aims at a goal in a largely pre-intellectual way by turning attention[10] to an existing concept within one's concept hierarchy and setting it as the focal concept to which another concept may be compared or contrasted. In principle, any concept in the hierarchy can be focused on; but, in the case of ongoing iterative subdivision, the focal concept is often one of the most recent concepts derived by subdivision: i.e., near the site of the most recent changes in the agent's overall conceptual framework. The process may be highly creative, as when test subjects (Hampton 2017, p. 113) are asked to imagine a bird that is also a kitchen utensil (focal concept) or a fruit that is also a piece of furniture (focal concept). More often than not, they succeed.

Consider economics, where it is customary to divide people into categories with income as the distinctor, thereby establishing (say) concepts of LOWER CLASS, MIDDLE CLASS, and UPPER CLASS.[11] Depending on whether one's intention is to boost the economy of the middle class, facilitate social opportunities for the lower class, or give tax breaks to those in the upper class, the focus concept is chosen accordingly, leading to new subdivisions as needed: say, LOWER MIDDLE CLASS (or WORKING CLASS) and UPPER MIDDLE CLASS.

3.2. Utility optimization: weighing predictors

The next step in the cycle comprises inferences to estimate the potential utility of each quality dimension $Q_1, \ldots, Q_m$ for achieving the intended goal – via a distinctor setting out a potential new subdivision of the previously determined focal concept, paired with a dataset (set of experience data) consisting of perceived objects referred to by the concept: i.e., its referents. The quality dimension predicted to have the best utility for action aiming at the intended goal is adopted as the next distinctor. Each difference distinction defines a uniquely distinguishable cognitive event; the chain of events that subsequent distinctions leave behind constitutes a sort of skeletal narrative: the path one lays down in walking.

We model utility optimization following the principle suggested by Luce and Raiffa (1956, p. 31): "given that a subject's preferences can be represented by a linear utility function, then *he behaves as if he were a maximizer of expected values of utility*" (emphasis original). We suggest three primary predictors of utility: *non-redundancy, distinctiveness,* and *proportionality*. These do not correspond to the design constraints of an optimal conceptual system, as proposed by Douven and Gärdenfors (2018); they can better be conceived as processual constraints driving the course of pragmatic conceptualization – conceptualization that only ever aims at "good enough for the moment".

3.2.1. Non-redundancy

In experience data, as in most multi-dimensional datasets describing human cognition, there is a high degree of interdependence among observable qualities: often they are reflections of one and the same phenomenon. To make sense of the whole, it is important to learn which dimensions co-vary and can therefore be connected via associative learning. Conceptualization facilitates this by seeking meaningful distinctions within the established focal concept for purposes of dividing it into subregions, each with more internal similarity than its parent concept. This is the creation of something more specific from the more general.

Consider an infant, immersed in the "great blooming, buzzing confusion" of which James (1890, p. 488) writes, who has learned to distinguish herself conceptually from others, needing to distinguish the others (focal concept) she can trust – most notably, perhaps, her mother – from strangers she cannot be sure of. The non-redundancy constraint comes to her assistance. Since only certain others return her smile, SMILE-RESPONSIVENESS does not correlate well with the previous self/other distinction. A new distinctor is required, one that should be as unrelated as possible to the dimensions co-varying within the focal (parent) concept.

Kaipainen and Hautamäki (2019) suggest calculating the non-redundancy of a candidate distinctor *Q* as 1 minus the maximum value of its correlations with other quality dimensions, with index *i* running from 1 to the number of quality dimensions in the space. The result approaches 1 as *Q* achieves maximal distinctiveness from the other dimensions.

$$nr(Q) = 1 - max\{corr(Q, Qi)\}$$

If Pearson's (1895) product moment correlation coefficient *r* is applied, it follows that $1 > nr(Q) > 0$ in all cases: i.e., the value of $nr(Q)$ is normalized to the range [0,1], where $nr(Q)$ is the quality dimension facilitating the greatest possible distinction to be made within the focal concept.

Consider the focal concept to be AFFLUENT CITIZENS. It would not contribute any utility to subdivide the concept by attempting to use WEALTH as the distinctor, since WEALTHY AMONG THE AFFLUENT amounts to the juxtaposition of two (near) synonyms, no more useful than looking for the affluent among the wealthy. It would be more useful to divide the conceptual space with a non-redundant distinctor that better contributes to the intended analysis of (for example) INHERITED WEALTH as opposed to NEW WEALTH. The non-redundancy principle would promote AGE-OF-WEALTH (how long the money has been in the family) as the new distinctor, since it does not directly correlate with affluence.

To take another example, consider the concepts of DEVELOPED NATION and DEVELOPING NATION. Both can usefully be subdivided, but probably not based on the average level of progress, however that is calculated, since DEVELOPED/DEVELOPING *imply* level of progress: i.e., on that dimension, the two concepts are relatively internally homogeneous.

The non-redundancy constraint, systematically applied, yields internally more homogenous sub-concepts and so sharpens distinctions between concepts. Each applied distinctor implicitly represents a potential number of other dimensions that co-vary with it to some extent. This kind of dimension clustering (or reduction) is reminiscent of such statistical procedures as principal component analysis (Pearson, 1901) or factor analysis (Child, 2006), with the difference that the optimization criteria emerge via intention instead of from the data themselves. We suggest that the result contributes to the "optimally designed similarity space" discussed by Douven and Gärdenfors (2018, p. 6) in the way that the dimensionality reduction leads toward increasing parsimony.

Preliminary neuroscientific data suggests that the non-redundancy principle indeed reflects processes going on in the mind and brain. In a systematic review, Chance (2014) concludes that the "development of more orthogonal dimensions … is associated with more sophisticated cognitive discriminative ability." He highlights Taylor et al. (1999), who found that activation in response to words in the left hemisphere is more focused and rapid than the corresponding processing in the right hemisphere, leading the authors to conclude that the right hemisphere typically uses more dimensions than the left to represent semantic maps.

3.2.2. Distinctiveness

A quality dimension is *distinctive* if the distribution of objects along it distinguishes two or more concept-grounding regions via gaps or sparse regions within the distribution, hinting at dependence on some other dimension. In contrast, a normal or even distribution leaves no such gaps and affords no such distinctiveness. It should be clear that a quality dimension without gaps is not a good candidate for being a distinctor.

In the present discussion, any measure of distinctiveness based on distribution serves the purpose. A straightforward measure is the broadest gap, $\max\{|x_{i-1} - x_i|\}$, where i indexes data points arranged in ascending or descending order, assuming a normalized range of [0,1]. Another suitable measure is relative sparsity.

The distinctiveness criterion serves Douven and Gärdenfors (2018, p. 4) design principle of contrast: "conceptual structure should be such that prototypes of different concepts can be so chosen that they are easy to tell apart"; distinctiveness should contribute to learnability, "required since varying environments preclude that all relevant concepts are initially provided." The higher the distinctiveness, the more distinct clusters that the domain mapped by the quality dimension affords.

Rosling et al. (2018, pp. 25–26) offer examples of distributions that do (Figure 3, left) and do not (Figure 3, right) support distinctiveness, relating to the familiar distinction between developing and developed countries. In 1965, the countries of the world clearly divided into two clusters with

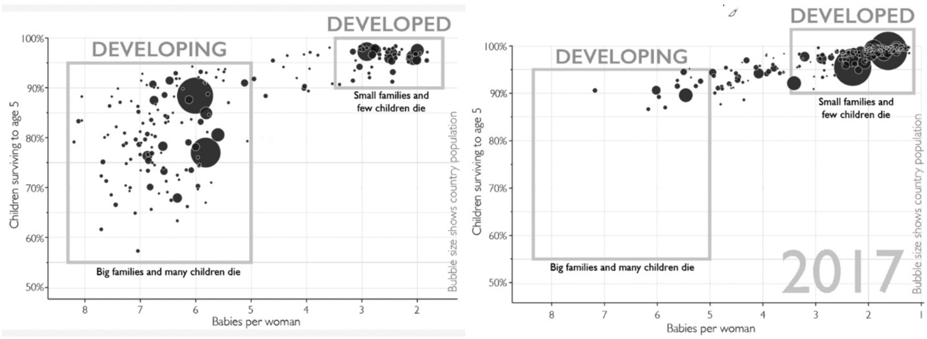

Figure 3. Scatter plots of two common measures of social progress. Note: Scatter plots by country of two common measures of social progress: *babies per woman* (*x*-axis) and *children surviving to age five* (*y-axis*). Images from Rosling et al. (2018, p. 25) (left) and 26 (right); based on free material from https://gapminder.org.

respect to *babies per woman* and *children surviving to age five* (Figure 3, left). The 2017 distribution looks quite different (Figure 3, right).

We do not mean to imply that the distinctiveness criterion – in contrast to the intention driving the ISDM loop – needs to be a matter of remotely conscious inference; indeed, most of the time it probably is not. Consider again the infant, equipped with the concept of I, observing “others” to determine which ones to trust. Her genetic heritage provides her with a tool for this purpose: the smile-mirroring reflex (Salzen, 1963; Simpson et al., 2014), which embodies the distinctiveness criterion, whereby some “others” respond to her smile while others do not. There is no reason to think the infant is in any way reflectively aware of what she is doing, even as she learns to tell TRUSTABLE OTHERS from LESS TRUSTABLE OTHERS–long, long before she can assign verbal labels to these clusters. Of course, the smile-mirroring response only works because of how the environment responds: it would have no utility if the “others” were uniformly lacking the abilities to smile back, like the cybernetic Cybermen of science fiction: human beings whose facial expressions have literally been frozen in place.

3.2.3. Proportionality

In order that a sub-concept may be distinguishable from its parent, it should normally be expected to cover a non-trivial proportion of instances of the parent concept, *but by no means all*. Outliers should *ceteris paribus* be ignored. The principle of proportionality (PQ) determines the minimum allowable proportion of objects in the smallest subdivision to be distinguished within the focal concept (N_{min}) vs. that allowable in the largest subdivision (N_{max}): i.e., $PQ = N_{min}/N_{max}$. A value of 1 indicates that no subdivisions need be made and small, outlier concepts are effectively not permitted; the closer the value is to 0, the more they are not just permitted but encouraged. The value of PQ should neither be too low nor too high but pragmatically set somewhere in the middle. Too low, and one has a concave concept like the concept of OUTLIER itself; too high, and the child concept is effectively identical to the parent.

The optimal value for proportionality will covary somewhat with the distinctiveness of the quality dimension used to distinguish the sub-concept from its parent. Nevertheless, the distribution shown in Figure 3, right, still usefully can be divided into two (or three) groups based on level of development, even though those groups are not so distinct from each other (in terms of babies per woman and infants surviving to age five) as previously.

The proportionality constraint is best approached by means of a statistical analysis (e.g., Howlin & Dziuban, 2019). It is obvious that the conditions dictating a concept's ideal coverage of items out of those covered by the super-concept is highly relative to context and intention, making it not just a challenge but a logical impossibility to find a universal rule or algorithm – as Nnamoko and Korkontzelos (2020) found in their machine-learning study.

3.2.4. Idiosyncratic constraints

There is no guarantee that the non-redundancy, distinctiveness, and proportionality constraints suffice to describe the conceptual distinctions made by a given individual in a given context. Heuristics relating to the individual's personality, expertise, beliefs, or gut feelings may not fit easily within any of the preceding constraints. These pragmatically driven constraints need not represent what is ordinarily conceived of as rational. They may even be deliberately tailored to mislead. *False inductibility* is an example of such a constraint. Consider MEXICAN RAPIST: a concept apparently formed by choosing a quality dimension RAPIST to divide persons of Mexican nationality into RAPISTS and NON-RAPISTS; but, instead of using the representative majority (NON-RAPISTS), it selects what amounts to an extreme outlier minority (RAPISTS) to make connotations about the majority, hiding the obvious misrepresentation whilst encouraging the fear that such a concept is likely to evoke. As noted earlier though, the same process can be used in creatively positive ways.

While it is almost certainly impossible to exhaust all constraints on an individual's conceptualization, their systemic contribution may nevertheless be described as follows.

3.3. Utility optimization: choosing a distinctor

The idea that categorization proceeds via iterated similarity/dissimilarity judgments is well established in psychology (see, e.g., Stewart & Morin, 2007, Yearsley et al., 2022): a vast literature that has often been used as a justification for prototype-based theories of concepts, such as Gärdenfors', which find their natural home in psychology. The purpose of the preceding constraints is to estimate the anticipated utility of each quality dimension as a distinctor. Pragmatism dictates that the choice of distinctor be non-deterministic: the cognitive agent makes the final choice on which constraints to prioritize. For a given agent in a given context, not all constraints have a role to play. High non-redundancy may – for most communicative purposes – have high analytic value, but it is less relevant or even irrelevant in some contexts: e.g., in selecting words for a poem. The contribution each constraint makes depends, among other things, on the size of the dataset, which must be large enough to allow utility estimates to be made in the first place.

Assume k processual constraints $C_1,\ldots C_k$ for the choice of distinctor among quality dimensions Q defining the focal domain. Each constraint C describes a statistical indicator associated with one aspect of utility that Q is expected to have for the intended action. We propose that this utility can be expressed by a value in the normalized range [0,1]: e.g., $C_{non\text{-}redundancy} = 1$ is fully orthogonal to all other potential distinctors; $C_{\text{distinctiveness}} = 1$ divides the parent concept into two, maximally distinct subgroups; $C_{\text{proponrtionality}} = 1$ assumes an even distribution within the parent concept.

To denote her prioritization among constraints, we suggest that the cognitive agent assigns an intentional regulator p_i to each constraint C_i. expressed as a real number in the normalized interval [0,1] where $p_i = 1$ represents maximal influence on the evaluation and $p_i = 0$ no influence. Taken together, the values constitute a list of regulator weights $[p_1, \ldots, p_k]$, one for each constraint C_i. The utility prediction value U_i for each candidate Q_i is calculable as:

$$U_i = (C_{1Qi})^{p1} \ldots (C_{kQi})^{pk}$$

... Where each C_{Qi} expresses the contribution of criterion C toward Q_i. The quality Q_i for which U_i is highest is chosen the next distinctor: i.e.,

$$Q = max\{U_i\}$$

While concepts are elaborately (re-)shaped for their intended purpose, the choice of distinctor itself need not in any way be a conscious act, any more than the weighting of any one or another constraint; recall the earlier discussion under the distinctiveness criterion. The concepts themselves must, at some level, be conscious (see Parthemore, 2017, p. 37, which argues that concepts and consciousness are two sides of one coin), but the mechanisms behind them need not and, in many cases, logically cannot be. Neither are we claiming that each difference distinction must be computational, in the manner we have presented here.

Consciousness raises a bundle of deep issues, as does the relationship between concepts and language. Suffice to say that practical action involves decisions that relate to physical orientation, including grasp, that initially involve only the domains of embodied knowledge and automated sensorimotor skills but may translate, in certain circumstances, into propositionally structured thought and onward to communicable language.

The utility estimate resulting in the choice of distinctor is based on accumulated experience data that is, and can only be, a somewhat arbitrary sample of the total repertoire of experiences possible. There is no guarantee of the distinctor's applicability to the intended action. The choice of distinctor may be interpreted as a kind of hypothesis (Gregory, 1980, who writes of "perception as hypothesis"[12]) that the distinctor has utility for the intended action. This "hypothesis" is valid in a broadly Popperian sense, having the potential to be falsified (Popper, 1934) by the test of practical action: i.e., it may turn out not to be useful at all, and another distinctor or another focal concept altogether may be chosen.

3.4. Practical action

Concepts are evaluated through practical action. Practical action is the embodied engagement of the agent with the environment in which she is embedded (and with which we understand her to be ultimately continuous; see Section 1.3). The action is shaped by an interplay of factors, of which concepts are often – though not always – a critical one. We make the simplifying assumption that the role of concepts is to allow the agent to grasp (Kaipainen & Hautamäki, 2019) the object or objects intended as her target by applying the distinctor arrived at via utility optimization.

If the utility estimate results in a sufficiently good choice of distinctor, it will make a difference relevant to the intended goal. To the extent the non-redundancy constraint is satisfied, the resulting distinctor allows the agent to approach the object or issue at hand from the most cognitively and ergonomically effective angle – think Gibson's (1979) *affordances*, potentials of action that the environment facilitates – allowing her to take full advantage of prior conceptualization while avoiding unnecessary repetition of action.

To the extent that the distinctiveness constraint is optimized, the chosen distinctor suggests effective cutlines. Whether in carpentry or conceptual agency, the optimal cutline requires the least effort to cut along. In an abstract domain such as the abortion debate, the perfect cutline corresponds to the most defendable line of argumentation, economizing the amount of work: be it in terms of vocal-tract muscles, words composed or mental effort exerted. Going back to Rosling et al.'s (2018, p. 26) example, while BABIES PER WOMAN may have been an obvious distinctor between developing and developed countries in 1965, by 2017 it was not.

To the extent that the proportionality constraint is optimized, the chosen distinctor should identify a nontrivial subset of instances of the parent concept. This should not be taken to imply that a concept with only a single referent is disallowed.

The choice of distinctor does not determine how it is to be applied. It determines the available options, with final determination left to the implementation of practical action, given the circumstances at hand. Constraint resolution is deeply dependent on context, and idiosyncratic constraints may – as discussed in Section 3.2.4—override the usual expectations.

Consider a group of school children evaluated for mathematical achievement. There may or may not be a gap in the distribution between high and low achievers; the distribution might be a bell curve, with the only apparent gaps toward the ends, and low distinctiveness in the center. All of that may not ultimately matter. Regardless of the distinctiveness criterion, the teacher may have a practical need to split the group into three for focused instruction. In that case, it might not be helpful to divide the students into one large group (the center) and two small ones (the top and bottom achievers). The teacher may rather choose to divide the students into three groups of even size. While the distinctiveness criterion is violated, the proportionality criterion is optimized and the choice serves the teacher's purpose.

In other circumstances, it may be wise to place the cutlines at fixed percentiles, as is customary in the case of normal distributions; or the pragmatic choice may be to perform no subdivision at all. Pragmatism may even require non-contiguous subdivisions that (for example) exclude the middle and combine the extremes into one category: a conceptual choice that blatantly violates the convexity principle at the heart of CST.

As said, our approach, in line with CST and sensorimotor theory, assumes that conceptualization is based on experience data, which our model is limited to describing, for now, in terms of quantitative data matrices. There are three aspects of accumulating experience that such matrices may help to make clear.

First, an experience may add a new quality dimension of which the agent has had no previous awareness, manifest as a perceived distinction not accounted for by the existing conceptual hierarchy. A child may have been successful sorting objects according to their size and shape only to discover that she now needs to take account of weight as well. Second, the experience may add new instances of conceptual referents, be their objects, events, or properties. Third, these instances may be characterized in terms of values varying continuously or discretely along the relevant quality dimensions.

The relation of experience data to quantitative data matrices cannot be treated here in anything but the most preliminary way. Suffice to say that practical action validates the utility "hypothesis" made in the utility-optimization step determining whether a change is worth adopting into the established conceptual framework or further adjustment is needed. Consider again the infant who intentionally or unintentionally smiles by reflex, perhaps adding a gesture to evoke a smile response and so determine the "right" distinction to make in categorizing the person. Most of the time, the feedback she receives will direct her subsequent behavior appropriately. In unfortunate cases, it may not be: say, if the "other" smiles with intent to deceive.

3.5. Concept adoption

Practical action determines the utility of the chosen distinctor, whose selection determines a set of child concepts for the parent concept. These child concepts and the distinctor itself still need to be incorporated into the existing conceptual framework. Concepts that do not subsequently prove successful are submitted for refinement or rejection, which may trigger rejection of other, previously accepted concepts on up the conceptual hierarchy.

Consider the child who discovers that grownups do not always have one's best interests at heart or know anything more than the child about what they are doing. Consider the religiously devout person who, in mid-life crisis, confronts her lurking agnosticism – all because of one small piece of evidence that does not fit.

3.5.1. Grasping, nameability, and convexity

The success of a concept can be analyzed in terms of its contribution to the intended action via grasping, nameability, and convexity. *Grasping* is built into the etymology of "concept": "con-" means together, while "–cept" derives from the Latin *capere*: to capture or to catch (see Kaipainen & Hautamäki, 2019, p. 107). Grasping the concept means targeting the action to the intended referent(s) in the environment. Failure to grasp amounts either to false negatives: referents that were intended to be targeted but weren't; or false positives: referents not intended to be targeted that nevertheless were. False positives make the concept cluster unnecessarily "heavy" and so harder to "lift," false negatives make referents slip from conceptual grasp, leading to unnecessary repetition. The infant grasping the value of the distinctor SMILE-RESPONSIVENESS grasps in a very concrete physical way those she sorts into the new category of TRUSTABLE OTHERS even as she shies away from those who fall outside.

Nameability is the single-most crucial prelinguistic property of a concept. A concept is not itself a name nor need it ever have a name, but it is something that (at least much of the time) can be named and that affords the potential of naming: i.e., being in some fashion lexicalizable. We have tied conceptualization directly to action. So long as actions remain strictly individual, a concept's "name" need be nothing more than a personal, idiosyncratic label – a memory hook – that affords retrieval. As soon as an action is shared – as soon as it becomes part of collective action and intention – a need to communicate it emerges: through gesture (as some other species do), through a lexicon of simple vocal signals (as many other species do), or through human-style language. We suggest that names are most closely associated with the most recently applied distinctor(s) and that the iterative ISDM loop plays a crucial role in the emergence of human-style language.

We have touched on *convexity* as a core property of most concepts. Consider two points in similarity space *S*, representing concepts *x* and *y*, both of which fall within the super-category *S*: i.e., they are examples of *S* and, let us say, some sub-
concept of *S*: namely, *T*. If convexity holds, then all points located on a line between *x* and *y* must also be examples of *T* (and, of course, of *S*).[13] Some useful concepts (including OUTLIER and GENTILE) are not convex; and non-convex concepts, by their nature, afford little if any utility prediction. They are useful in another way, by saying what something is *not*: an outlier is not typical; a Gentile is anyone who is not Jewish. Of course, "bad" concepts can be intentionally non-predictive, the product of rank prejudice – consider the aforementioned MEXICAN RAPIST – or propagandistic manipulation. Simply put, the point of some concepts is to mislead (oneself or others). Most concepts that succeed, though, are successful to the extent that they are convex, facilitating predictions about concepts close to them in the similarity space.

3.5.2. The path beyond adoption or rejection

Conceptualization does not end at concept adoption/rejection. Top-down intentionality (steps 1 and 2) constantly meets bottom-up, perceptually driven forces (steps 3 and 4). Concept adoption contributes to subsequent intentional orientation as the conceptual hierarchy continues to be refined and experience continues to accumulate. We have described it as a circular process, but this is purely simplification for sake of explanation; the top-down and bottom-up forces are active simultaneously, not one after another; they truly are inseparable. Otherwise, the model would be subject to the same devastating objection Rodney Brooks (1995) makes to what he derides as the SMPA (sense-model-plan-act) model of perception.

4. Discussion

The model of conceptualization we have described assumes a continuous dynamic that molds concepts into each new context in intentional, mostly pre-intellectual "hands on" engagement with the environment. Both for individuals and societies, concepts emerge, get tested (or discarded!), adapt, mature, change into something quite different, or disappear altogether. PHLOGISTON is a concept that was useful but survives today mainly for historical purposes, known only by a few; while other concepts are entirely forgotten: either because the agents who entertained them never shared them, or because the shared usage or very purpose has been forgotten. The human species not only learns; it also forgets.

4.1. Concepts in motion

Whether at the individual or collective level (Parthemore, 2014),[14] concepts and conceptual frameworks only appear at first glance to be hyper-stable. Closer inspection shows constant movement throughout the hierarchy as individual concepts and entire subsections of the conceptual framework are evaluated and reevaluated based on their continued utility for action. Every change, however small, sends ripples throughout the entire system. Even such seemingly static concepts as PI evolve as new experiences bring new mathematical insights and number theory itself evolves: for pi and other transcendental numbers are understood in relation to the number theory of which they form a part. Typically, of course, the changes are unrecognizably small; but sometimes they are huge, as when number theory was expanded to include so-called *imaginary* and *complex numbers*.

Concepts appear most stable and representation-like when we stop and reflect on them; they are most changeable when we are simply getting on with using them, non-reflectively (Parthemore, 2011, pp. 37–38; 2019): still part of our conscious minds, to be sure, but not of our reflectively self-conscious awareness. Individually or collectively, the foundations of the evolving conceptual hierarchy lie in its oldest and most stable distinctions, built upon its protoconceptual foundations (Parthemore, 2014) and ultimately built into the species' collective memory and embodied cognition; while the tips of its branches show the most growth and change.

If we have not talked much about this essentially dynamic nature of concepts, that choice has been deliberate. We raised it in the introduction to set the context for the ISDM loop, but this paper is ultimately about that loop and not the way that, in our understanding, concepts are forever in motion: the ever-flowing product of action and themselves *in action*; like the elaboration of experience data, further discussion of those dynamics must await another paper.

4.2. Concepts in action

We consider concepts simultaneously as *instruments for action* and as *manufactured artifacts of action*, in intimate relation to human intentionality: concepts are what allow human beings to be intentional, to think and act with some degree of deliberateness and deliberation. In an important sense, concepts are no more natural than many other human achievements: which is to say that they, too, are artifacts and not mere affordances.[15]

In another sense, concepts may be taken to be perfectly natural: for it appears to be part of human nature, given the appropriate environment, to develop into conceptual agents. Douven and Gärdenfors (2018, p. 2) might be taken to imply that naturalness is a characteristic only of optimal concepts – in which case, we respectfully disagree. We have limited interest in describing optimal concepts. We have suggested here that concepts only need to be "good enough" in terms of the extent to which they provide utility at the moment.

We are in good company with all those for whom perception is conditioned by existing knowledge, which sets sharp boundaries on subsequent experience and action: as people go through life, they are more and more restricted to what they expect to encounter; it takes greater and greater force to break out of those patterns. Their conceptual frameworks simultaneously simplify and facilitate their interaction with the world and at the same time bind them in to one way of encountering it, to the exclusion of other possibilities: the conceptual equivalent of painting oneself into a corner.

Standing on the shoulders of such giants as Kant (Ben-Zeev, 1984), James (Stevens, 1974), Peirce (Ayer, 1982), Dewey (Prawat, 1995), Gregory (1980), and Ulrich Neisser (1976), this line of thought continues to evolve (see, e.g., Chen et al., 2018). So inspired, we trust that our model contributes to addressing the intricate puzzle of how knowledge of and prejudices about the environment accumulate, hand in hand – and how these, in turn, influence the conceptual agent's encounter with her environment – until something sweeps the structure away, so the conceptual agent must build again.

What we have not done in this paper is commit to any particular theory of action, which limits our ability to describe utility in anything approaching formal terms. Going forward, a theory of action is needed, drawing on the work of Elizabeth Anscombe (2000), Donald Davidson (1980) and Raimo Tuomela (2012) among others.

4.3. Concepts in perspective

From a mathematical point of view, our model applies an essentially quantitative set of algorithms to describe the qualitative experiences with which conceptualization is inextricably entangled – experiences that, by virtue of their complexity, ultimately lie beyond the possibility of formalization, in the realm of *terra incognita*.

One might well ask whether computational quantities are alien to experiential quality in the first place. Suffice to say that the distinction between the two is almost certainly conceptual rather than ontological. They relate to each other through embodied experience of action, where the world-in-motion is measured using the human body as its yardstick: an idea in harmony with contemporary thinking on the origins of mathematics (Dehaene & Brannon, 2011).

From a psychological point of view, our characterization of conceptualization bears close relation to discussions on memory. The dynamically evolving concepts we describe could be considered as means of abstracting, generalizing, and compacting *semantic memory* (Tulving, 1972) within a stream of accumulating experience data. Endel Tulving (1972, p. 385) contrasts semantic memory with episodic memory concerning "temporally dated episodes or events, and temporal-spatial relations among these events." Our approach requires no such sharp distinction, at least when it comes to concepts. Just as concepts may be seen to sit between *knowledge how* and *knowledge that* (Ryle, 1949), beholden to neither; so, too, may they be seen to sit between semantic and episodic memory[16]: push them in one direction, and they feel more like remembrances of facts about the world; push them the other, and they feel more like remembrances of the times and places in which they arose and were (re-)shaped.

Consider Tulving's (1985) distinction between anoetic consciousness, which is strictly limited to the present moment and context, purely "in the now" (independent of memory); noetic consciousness (tied to semantic memory), which is awareness of and ability to think about objects and events outside the present moment and context; and autonoetic consciousness (tied to episodic memory), which is awareness of and ability to think about one's personal subjective experience of objects and events outside the present moment and context. For all that their application is always in the present moment and context, concepts – by their ability to be applied systematically across unboundedly many contexts – are precisely what lift an agent out of anoetic consciousness (if it can be termed consciousness at all; we suggest not) to noetic and autonoetic levels. It is tempting to see this process in relation to, on the one hand, the distinction between protoconcepts, first-order concepts, and higher-order concepts (concepts of concepts); and, on the other, the construction of self on its different levels, from implicit recognition of self to explicit recognition to reflective awareness of self-as-myself (see Bruner, 1994; Parthemore, 2011, pp. 68–69)–topics, however, that once again fall outside the purview of this paper.

4.4. *Concepts in narrative*

Narrative may be understood as temporally extended representation of context. Although we have only been able to touch briefly on the idea of the ISDM loop constituting a kind of conceptual narrative, there are implications, worth touching on, for the emerging field of narrative psychology, which studies how "the story becomes an object of study, focusing on how individuals or groups make sense of events and actions in their lives" (Mitchell & Egudo, 2003, p. 2; see also Riessman, 1993; Vassilieva, 2016). If one accepts, for sake of argument, something like the ISDM loop, then it seems justified to interpret sequences of difference distinctions – individual cognitive events – as comprising a story or *autonarrative* (Gazzaniga, 1998) that the cognizer tells herself. It is the path she lays down in walking: the story created even as it is being told.

Long before verbalization begins, a personal narrative helps make sense of oneself-in-the-world by chaining embodied actions into quasi-causal explanations built on one's emerging conceptual framework. Future work might look for links between the relatively low-level sequences of different distinctions and things that are readily recognized as stories. Narrative priming by means of, e.g., cinematic immersion (Jääskeläinen et al., 2021; Tikka & Kaipainen, 2014) may open one path to empirical testing for models – like the ISDM – of dynamic, context-dependent concepts.[17] This could take the form of sorting experiments of the kind Kriegeskorte and Mur (2012) discuss, in which subjects are primed with similar vs. dissimilar contexts. The same experiments can be used to infer sorting criteria and map them to the optimization constraints we have discussed.

5. Conclusions

The main contribution of this paper is the way it relates concepts to action in terms of utility optimization on one hand and utility actualization on the other: i.e., the evaluation of concepts *in action*. Likewise, important is the way it elaborates the emergence and evolution, through trial-by-action, of those quality dimensions we call distinctors. While the present model primarily addresses concepts as constructs shaped by individual agents' experience, it suggests a pathway from those "personal" concepts to shared concepts via common action.

Concepts are the dynamic artifacts of an intentional mind trying to make sense of the world and acting on it. For some number of those agents who are also social agents, concepts serve as the requisite foundation for communication and – in the human case – language. Whether strictly "private" concepts (*pace* one common reading of Ludwig Wittgenstein, 2001, §256 *ff*.) arise in nonsocial species remains an open empirical question.

Our approach takes a pragmatic view, characterizing concepts through the ways they emerge and are applied in everyday behavior. We have extensively elaborated the original ISDM model described by Kaipainen and Hautamäki (2019), setting out a four-step cycle with emphasis on utility optimization. We have argued for three constraints on that optimization: non-redundancy ("do not reinvent the wheel"), distinctiveness ("avoid differences that do not make a difference") and proportionality ("be measured"), while leaving the door open to inevitable idiosyncratic constraints.

Non-redundancy relates to one of the key issues facing cognitive systems: how to deal with overlap among the pre-established qualities a conceptual agent must choose from in making sense of her experience. One quality should serve the agent's present intentions better than all the rest (*cf*. the extensive literature on pattern separation); it becomes what we call the distinctor. Selecting the best distinctor need not be interpreted as rejecting the competitors; the "winning" distinctor may represent a bundle of related dimensions reflecting one and the same underlying phenomenon. Distinctiveness and proportionality, meanwhile, are handy but far from universally applicable tools: some distributions are effectively continuous and so do not afford breaking points – in which case, the conceptual-subdivision buck stops here.

One of the many things lacking at the moment is a fully worked out model of re-conceptualization within the ISDM. That is an accident of the ISDM's development to date. When a concept is rejected for adoption and when this in turn forces reconsideration of higher-level (more abstract) levels of the concept hierarchy, the critical question is: how much structure is removed, and where does one stop? Remove *too* much, after all, and the agent will no longer be able to function effectively; remove too little, and while the immediate symptoms are addressed, the underlying roots of the problem are not.

Our ultimate goal is an integrated mathematical model of the difference-distinction cycle and a new logic for specifying concepts, one that allows, e.g., statements of the form:

$$p^a \wedge \neg p^b$$

... To be evaluated as the conjunction of two true statements implying a true statement where p is true in context a and not true in context b: i.e., the logic should be context sensitive, like Hautamäki's (2022) viewpoint logic. The more clearly detailed the algorithm at the heart of the ISDM, the clearer the emerging logic should be: that is, the former should develop naturally into the latter.

We have deliberately refrained from speculating on the neural mechanisms underlying the ISDM. The relationship between neural mechanisms and high-level conceptual cognition is (at the least) staggeringly complex, naysaying any simple reductionist approach. More likely, the full connections between the two levels outstrip the power of human cognition to know itself in any complete and consistent way. Of course, there are important things to be said about neural plasticity and the relationship (if any) of artificial neural networks (ANNs) to actual brains; but that, again, remains the subject for another paper.

Notes

1. Strong empirical evidence for similarity's dependence on context comes, e.g., from Yearsley et al. (2022), which shows how distractors can be used both to increase and decrease judgments of similarity. For an overview of empirical evidence for the general context sensitivity of meaning, see Hampton (2017).
2. That could change, of course, if evidence accumulates that concepts are physical symbols in the brain *per* Allen Newell and Herbert Simon (1976) and Fodor (1998); or reliably identifiable patterns of neural activation, particularly if those patterns prove to be consistent across individuals. At this junction, the Physical Symbol System Hypothesis has largely fallen out of favor, while no evidence exists to date that specific concepts can be identified by neural activation: i.e., we know this person is thinking about *X*.
3. Despite one reviewer's claim to the contrary, the two are *not* equivalent: many important case studies and other forms of exploratory empirical research are done without making falsifiable hypotheses.
4. Our only commitment in using the term "qualia" is to suggest that experience is somehow systematically structured. Nothing further should be taken as implied.
5. That concepts are intimately bound to perception (the source of all experience data) we take as a starting assumption. That the raw qualitative data created through perception can be quantified we take as a useful simplifying assumption for purposes of elaborating CST and the ISDM. However, the precise manner by which perception becomes qualitative experience data or qualitative experience data can best be quantified we take to be outside the scope of this paper.
6. For an introduction to uses of this now-popular phrase, see Ward and Stapleton (2012).
7. Note that Gärdenfors (2014, pp. 23–25) attempts to make a sharp distinction between concepts and properties.
8. Gärdenfors suggests (Gärdenfors, 2014, p. 25)—without elaboration – that not all conceptual domains have a metric; our assumption is that they do.
9. The hierarchy is an exceedingly tangled one with recurrent feedback possible between all levels, akin to what Douglas Hofstadter describes as a *tangled hierarchy* (Hofstadter, 2000) or *strange loop* (Hofstadter, 2007).
10. Though attention determines the choice of focal concept, an account of focused attention lies beyond the scope of this paper.

11. Be reminded that, while we are forced to use linguistic labels to refer to concepts, we in no way wish to suggest that concepts just are lexical concepts, as *per* Fodor (1998) and as many other researchers take them to be. *Pace* Fodor, the mapping of concepts to words of a language is a far from straightforward process. Agents may hold several near-synonymous concepts that, for most purposes, have the same extension and only diverge under circumstances that do not presently apply.
12. See also Clark's work on *predictive processing*: e.g. (Clark, 2015).
13. Consider two points in the COLOR space, as defined by the quality dimensions HUE, SATURATION, and BRIGHTNESS: if both points belong to RED (a child concept of COLOR), then any points between them must also belong to RED; they are all shades of red because RED is a convex region within the COLOR space.
14. Gärdenfors (2014, pp. 23,23) makes a sharp distinction here between what he sees as two fundamentally entities: what he calls "cognitive" (individual) and "scientific" (collective) concepts. We do not. Indeed, the argument in (Parthemore, Joel. Parthemore, 2014) is that they represent the same entity operating on different time scales.
15. . . . Albeit a class of artifacts humans appear to share with a range of other species; there is now a wealth of literature on non-human animal concepts. For a good starting point, see Newen and Bartels (2007).
16. Note that Tulving's semantic/episodic memory distinction, in which semantic memory is more basic and primary, crosscuts in a curious way with Ryle's knowledge how/knowledge that distinction, in which knowledge how is more basic and primary.
17. For further consideration of empirical testing of this and related versions of CST, see (Parthemore, 2015).

Disclosure statement

The authors confirm that they have no potential conflicts of interest to report.

References

Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Blackwell.

Anscombe, E. (2000). *Intention* (2nd ed.). Harvard University Press.

Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman & Hall.

Ayer, A. J. (1982). *The origins of pragmatism: Studies in the philosophy of Charles Sanders Peirce and William James*. Macmillan.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 101–140). Cambridge University Press.

Barsalou, L. W., & Goldstone, R. L. (1998). Reuniting perception and conception. *Cognition*, *65*(2–3), 231–262. https://doi.org/10.1016/S0010-0277(97)00047-4

Ben-Zeev, A. (1984). The Kantian revolution in perception. *Journal for the Theory of Social Behaviour*, *14*(1), 69–84. https://doi.org/10.1111/j.1468-5914.1984.tb00488.x

Bishop, M., & Martin, A. O. (Eds.). (2014). *Contemporary sensorimotor theory: Studies in applied philosophy, epistemology, and rational ethics* (Vol. 15). Springer.

Brentano, F. (2015). *Psychology from an empirical standpoint*. Taylor & Francis.

Brooks, R. (1995). Intelligence without reason. In L. Steels & R. Brooks (Eds.), *The artificial life route to artificial intelligence* (pp. 13–69). Taylor & Francis. https://doi.org/10.4324/9781351001885

Bruner, J. (1994). The "remembered" self. In U. Neisser & R. Fivush (Eds.), *The remembering self: Construction and accuracy in the self-narrative*(p. 41–54). Cambridge University Press.

Chance, S. A. (2014). The cortical microstructural basis of lateralized cognition: A review. *Frontiers in Psychology*, *5*, 820. https://doi.org/10.3389/fpsyg.2014.00820

Chen, H., Ning, X., Wang, L., & Yang, J. (2018). Acquiring new factual information: Effect of prior knowledge. *Frontiers in Psychology*, *9*, 1734. https://doi.org/10.3389/fpsyg.2018.01734

Child, D. (2006). *The essentials of factor analysis* (3rd ed.). Continuum.

Clark, A. (2008). *Supersizing the mind*. Oxford University Press.

Clark, A. (2015). Radical predictive processing. *The Southern Journal of Philosophy*, *53*(S1), 3–27. https://doi.org/10.1111/sjp.12120

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, *58*(1), 7–19. https://doi.org/10.1093/analys/58.1.7

Davidson, D. (1980). *Essays on actions and events*. Oxford University Press.

Dehaene, S., & Brannon, E. (2011). *Space, time and number in the brain: Searching for the foundations of mathematical thought*. Academic Press.

Dewey, J. (1925). *Experience and nature*. Open Court Publishing.

Douven, I., & Gärdenfors, P. (2018). What are natural concepts? A design perspective. *Mind & Language*, *35*(3), 313–334. https://doi.org/10.1111/mila.12240

Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford University Press.

Fox, M. A. (2005). *The accessible Hegel*. Prometheus Books.

Gärdenfors, P. (1995). *Language and the evolution of cognition* [Unpublished manuscript]. Department of Philosophy, Lund University. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.58.380&rep=rep1&type=pdf

Gärdenfors, P. (2003). *How homo became sapiens: On the evolution of thinking*. Oxford University Press.

Gärdenfors, P. (2004). *Conceptual spaces: On the geometry of thought*. MIT Press.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press.

Gazzaniga, M. S. (1998). *The mind's past*. University of California Press.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.

Giere, R. N. (2006). *Scientific perspectivism*. University of Chicago Press.

Goodman, N. (1972). Seven strictures on similarity. In N. Goodman (Ed.), *Problems and projects* (pp. 437–446). Bobbs-Merrill.

Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *290*(1038), 181–197. https://doi.org/10.1098/rstb.1980.0090

Hampton, J. A. (2017). Compositionality and concepts. In J. A. Hampton & Y. Winter (Eds.), *Compositionality and concepts in linguistics and psychology* (pp. 95–121). Springer.

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 1–52). Cambridge University Press.

Hautamäki, A. (1986). *Points of view and their logical analysis* [Doctoral dissertation]. Societas Philosophica Fennica.

Hautamäki, A. (2016). Points of view: A conceptual space approach. *Foundations of Science*, *21*(3), 493–510. https://doi.org/10.1007/s10699-015-9422-2

Hautamäki, A. (2020). *Viewpoint relativism: A new approach to epistemological relativism based on the concept of points of view (Synthese Library 419)*. Springer.
Hautamäki, A. (2022). Disagreement, points of view, and truth-relativism. *Acta Analytica*, *37*(4), 531–550. https://doi.org/10.1007/s12136-021-00504
Hofstadter, D. R. (2000). *Gödel, Escher, Bach: An eternal golden braid* (20th-anniversary ed.). Penguin Books.
Hofstadter, D. R. (2007). *I am a strange loop*. Basic Books.
Howlin, C. P., & Dziuban, C. D. (2019). Detecting outlier behaviors in student progress trajectories using a repeated fuzzy clustering approach [Conference presentation]. In *International Conference on Education Data Mining*, Montreal, 2-5 July 2019. http://files.eric.ed.gov/fulltext/ED599200.pdf
Ihde, D., & Malafouris, L. (2019). Homo faber revisited: Postphenomenology and material engagement theory. *Philosophy & Technology*, *32*(2), 195–214. https://doi.org/10.1007/s13347-018-0321-7
Jääskeläinen, I. P., Sams, M., Glerean, E., & Ahveninen, J. (2021). Movies and narratives as naturalistic stimuli in neuroimaging. Elsevier. *NeuroImage*, *224*, 117445. https://doi.org/10.1016/j.neuroimage.2020.117445
Jäger, G. (2010). Natural color categories are convex sets. In M. Aloni, H. Bastiaanse, T. de Jager, & K. Schulz (Eds.), *Logic, language and meaning: Lecture notes in computer science* (pp. 11–20). Springer.
James, W. (1890). *The principles of psychology* (Vol. 1). Holt.
Kaipainen, M., & Hautamäki, A. (2015). A perspectivist approach to conceptual spaces. In F. Zenker & P. Gärdenfors (Eds.), *Applications of conceptual spaces: The case for geometric knowledge representation* (pp. 245–258). Springer.
Kaipainen, M., & Hautamäki, A. (2019). Seeking for the grasp: Iterative subdivision model of conceptualization. In M. Kaipainen, F. Zenker, A. Hautamäki, & P. Gärdenfors (Eds.), *Concepts and their applications* (pp. 103–126). Springer.
Kriegeskorte, N., & Mur, M. (2012). Inverse MDS: Inferring dissimilarity structure from multiple item arrangements. *Frontiers in Psychology*, *3*, 245. https://doi.org/10.3389/fpsyg.2012.00245
Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.
Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
LeDoux, J. (2020). *The deep history of ourselves: The four-billion-year story of how we got conscious brains*. Penguin Books.
Linnaeus, C. (1735). *Systema naturae* (Vol. 1). Laurentii Salvii.
Luce, R. D., & Raiffa, H. (1956). *Games and decisions: Introduction and critical survey*. John Wiley and Sons.
Marx, K. (1941). *Grundrisse der Kritik der politischen œkonomie: Anhang; 1850-1859* [Outlines of the critique of political economy]. Verlag für fremdsprachige Literatur.
Maturana, H. R., & Varela, F. J. (1992). *The tree of life: The biological roots of human understanding*. Shambhala.
Mitchell, M. C., & Egudo, M. (2003). *A review of narrative methodology*. DSTO Systems Sciences Laboratory. https://apps.dtic.mil/sti/pdfs/ADA421725.pdf
Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. W. H. Freeman.
Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, *19*(3), 113–126. https://doi.org/10.1145/360018.360022
Newen, A., & Bartels, A. (2007). Animal minds and the possession of concepts. *Philosophical Psychology*, *20*(3), 283–308. https://doi.org/10.1080/09515080701358096

Nnamoko, N., & Korkontzelos, I. (2020). Efficient treatment of outliers and class imbalance for diabetes prediction. *Artificial intelligence in medicine*, *104*, 101815. https://doi.org/10.1016/j.artmed.2020.101815

Omicini, A., Ricci, A., & Viroli, M. (2006). Agens Faber: Toward a theory of artifacts for MAS. *Electronic Notes in Theoretical Computer Science*, *150*(3), 21–36. https://doi.org/10.1016/j.entcs.2006.03.003

Parthemore, J. (2011). *Concepts enacted: Confronting the obstacles and paradoxes inherent in pursuing a scientific understanding of the building blocks of human thought* [Doctoral dissertation]. University of Sussex, UK Sussex Research Online. https://core.ac.uk/download/pdf/2710203.pdf

Parthemore, J. (2013a). The "final frontier" as metaphor for mind: Opportunities to re-conceptualize what it means to be human. In D. Duner, J. Parthemore, E. Persson, & G. Holmberg (Eds.), *The history and philosophy of astrobiology: Perspectives on the human mind and extraterrestrial life* (pp. 67–92). Cambridge Scholars Publishing.

Parthemore, J. (2013b). The unified conceptual space theory: An enactive theory of concepts. *Adaptive Behavior*, *21*(3), 168–177. https://doi.org/10.1177/1059712313482803

Parthemore, J. (2014). The case for protoconcepts: Why concepts, language, and proto-language all need protoconcepts. *Theoria Et Historia Scientiarum*, *11*, 159–178. https://doi.org/10.12775/ths-2014-009

Parthemore, J. (2014). Conceptual change and development on multiple time scales. *Sign Systems Studies*, *42*(2–3), 193–218. https://doi.org/10.12697/SSS.2014.42.2-3.03

Parthemore, J. (2014c). From a sensorimotor to a sensorimotor++ account. In M. J. Bishop & A. O. Martin (Eds.), *Contemporary sensorimotor theory: Studies in applied philosophy, epistemology and rational ethics* (pp. 137–158). Springer.

Parthemore, J. (2015). Specification of the unified conceptual space, for purposes of empirical investigation. In F. Zenker & P. Gärdenfors (Eds.), *Applications of conceptual spaces: The case for geometric knowledge representation* (pp. 223–244). Springer.

Parthemore, J. (2017). Consciousness, semiosis, and the unbinding problem. *Language & Communication*, *54*, 36–46. https://doi.org/10.1016/j.langcom.2016.10.004

Parthemore, J. (2019). On the essentially dynamic nature of concepts: Constant if incremental motion in conceptual spaces. In M. Kaipainen, F. Zenker, A. Hautamäki, & P. Gärdenfors (Eds.), *Concepts and their applications* (pp. 83–102). Springer.

Parthemore, J., & Morse, A. (2010). Representations reclaimed: Accounting for the co-emergence of concepts and experience. *Pragmatics and Cognition*, *18*(2), 273–312. https://doi.org/10.1075/pc.18.2.03par

Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, *58*(347–352), 240–242.

Pearson, K. (1901). LIII: On lines and planes of closest fit to systems of points in space. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572. https://doi.org/10.1080/14786440109462720

Piaget, J. (1972). *The psychology of intelligence*. Littlefield.

Popper, K. (1934). *The logic of scientific discovery*. Routledge.

Prawat, R. S. (1995). Misreading Dewey: Reform, projects, and the language game. *Educational Researcher*, *24*(7), 13–22. https://doi.org/10.3102/0013189X024007013

Richards, R. J. (2008). *The tragic sense of life: Ernst Haeckel and the struggle over evolutionary thought*. University of Chicago Press.

Riessman, C. K. (1993). *Narrative analysis*. SAGE.

Robinson, D. (2013). *Feeling extended: Sociality as extended body-becoming-mind*. MIT Press.

Rosch, E. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328–350. https://doi.org/10.1016/0010-0285(73)90017-0

Rosling, H., Rosling, O., & Rönnlund, A. R. (2018). *Factfulness: undefined reasons we're wrong about the world - and why things are better than you think*. Sceptre.

Rupert, R. D. (2009). *Cognitive systems and the extended mind*. Oxford University Press.

Ryle, G. (1949). *The concept of mind*. Penguin.

Salzen, E. A. (1963). Visual stimuli eliciting the smiling response in the human infant. *The Journal of Genetic Psychology*, *102*(1), 51–54. https://doi.org/10.1080/00221325.1963.10532722

Searle, J. R. (1980). Minds, brains, and programs. *The Behavioral and Brain Sciences*, *3*(3), 417–457. https://doi.org/10.1017/S0140525X00005756

Simpson, E. A., Murray, L., Paukner, A., & Ferrari, P. F. (2014). The mirror neuron system as revealed through neonatal imitation: Presence from birth, predictive power and evidence of plasticity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1644), 20130289. https://doi.org/10.1098/rstb.2013.0289

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In B. Burns (Ed.), *Percepts, concepts, and categories: The representation and processing of information* (pp. 233–274). Elsevier.

Stevens, R. (1974). Sensation, perception, conception. In R. Stevens (Ed.), *James and Husserl: The foundations of meaning* (pp. 24–46). Springer.

Stewart, N., & Morin, C. (2007). Dissimilarity is used as evidence of category membership in multidimensional perceptual categorization: A test of the similarity-dissimilarity generalized context model. *The Quarterly Journal of Experimental Psychology*, *60*(10), 1337–1346. https://doi.org/10.1080/17470210701480444

Taylor, K. I., Brugger, P., Weniger, D., & Regard, M. (1999). Qualitative hemispheric differences in semantic category matching. *Brain and Language*, *70*(1), 119–131. https://doi.org/10.1006/brln.1999.2148

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. The Belknap Press of Harvard University Press.

Tikka, P., & Kaipainen, M. (2014). From naturalistic neuroscience to modeling radical embodiment with narrative enactive systems. *Frontiers in human neuroscience*, *8*, 794. https://doi.org/10.3389/fnhum.2014.00794

Trigoni, T. (2015). Corporeal cognition: Pragmatist aesthetics in William James. In A. Scarinzi (Ed.), *Aesthetics and the embodied mind: Beyond art theory and the cartesian mind-body dichotomy* (pp. 55–69). Springer.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). Academic Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie Canadienne*, *26*(1), 1–12. https://doi.org/10.1037/h0080017

Tuomela, R. (2012). *A theory of social action* (Vol. 171). Springer.

Varela, F., Thompson, E., & Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*. MIT Press.

Vassilieva, J. (2016). *Narrative psychology: Identity, transformation and ethics*. Springer.

Versace, R., Vallet, G. T., Riou, B., Lesourd, M., Labeye, É., & Brunel, L. (2014). Act-In: An integrated view of memory mechanisms. *Journal of Cognitive Psychology*, *26*(3), 280–306. https://doi.org/10.1080/20445911.2014.892113

von Bertalanffy, L. (1973). *General systems theory: Foundations, development, applications*. Braziller.

von Wright, G. H. (1971). *Explanation and understanding*. Routledge.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4*(1), 45–69. https://doi.org/10.1016/0959-4752(94)90018-3

Ward, D., & Stapleton, M. (2012). Es are good: Cognition as enacted, embodied, affective and extended. In F. Paglieri (Ed.), *Consciousness in interaction: The role of the natural and social context in shaping consciousness* (pp. 89–106). John Benjamins.

Wittgenstein, L. (2001). *Philosophical investigations: The German text, with a revised English translation* (G.E.M. Anscombe, trans.). Blackwell.

Yearsley, J. M., Pothos, E., Barque Duran, A., Trueblood, J., & Hampton, J. A. (2022). Context effects in similarity judgments. *Journal of Experimental Psychology. General*, *151* (3), 711–717.