

A Generalized Patchwork Approach to Scientific Concepts

Philipp Haueis

ABSTRACT: Polysemous concepts with multiple related meanings pervade natural languages, yet some philosophers argue that we should eliminate them to avoid miscommunication and pointless debates in scientific discourse. This paper defends the legitimacy of polysemous concepts in science against this eliminativist challenge. My approach analyses such concepts as patchworks with multiple scale-dependent, technique-involving, domain-specific and property-targeting uses (patches). I demonstrate the generality of my approach by applying it to "hardness" in materials science, "homology" in evolutionary biology, "gold" in chemistry and "cortical column" in neuroscience. Such patchwork concepts are legitimate if the techniques used to apply them produce reliable results, the domains to which they are applied are homogenous, and the properties they refer to are significant to describe, classify or explain the behavior of entities in the extension of the concept. By following these normative constraints, researchers can avoid miscommunication and pointless debates without having to eliminate polysemous patchwork concepts in scientific discourse.

Keywords: Concepts, Polysemy, Patchwork Concepts, Eliminativism, Pluralism

1. *Introduction*
2. *The Patchwork Structure of Scientific Concepts*
 - 2.1 *Patchwork concepts and polysemy*
 - 2.2 *A general format for representing patchwork structure*
3. *How Patches Can Break: Three Kinds of Failure*
 - 3.1 *Unreliable techniques*
 - 3.2 *Heterogenous domains*
 - 3.3 *Insignificant properties*
 - 3.4 *When do patch failures necessitate novel concepts?*
4. *The Legitimacy of Patchwork Concepts*
 - 4.1 *Normative constraints on using polysemous scientific concepts*
 - 4.2 *Meeting the eliminativist challenge to polysemous concepts*
5. *Conclusion*
6. *References*

1 Introduction

There is a growing consensus in philosophy (of science) that we should not expect that (scientific) concepts can be analysed as a set of properties that are individually necessary and jointly sufficient conditions of application (Taylor and Vickers [2017]). We learned this lesson in part because during scientific growth and development, scientists routinely migrate concepts and allow them to change, often imperceptibly at first (Wilson [2006]). An adequate account of scientific concepts should recognize that there is a necessary kind of flexibility and looseness in how concepts are deployed and develop in scientific thought. Yet scientific concepts with too much slack cease to function as such: they invite equivocation where we wanted clarity of understanding.

This paper focuses on polysemy as a legitimate and indeed necessary kind of looseness in scientific language and provides a general approach to explain how scientists can use polysemous concepts rigorously and without equivocation. Polysemous concepts are words which have multiple meanings that are related to each other (Vicente [2017]). Polysemy is a pervasive feature of natural languages, yet linguists observe that it typically does not pose problems in everyday communication (Taylor [2003], Falkum [2015]). What is rarely discussed in linguistics are cases of scientific polysemy like the following:

- (1) The temperature of a gas increases when its volume is compressed.
- (2) The temperature of a rubber band increases when its diameter is expanded.

For (1), equilibrium thermodynamics defines “temperature” as “mean molecular kinetic energy”. This definition cannot be extended to (2), however, because rubber bands are not in molecular thermodynamic equilibrium. They reach this state only after dissolving into molecular polymer soup over hundreds of years. Yet rubber bands display the higher-scale property of frozen order of polymer chains, which determines their ability to move freely. When the

rubber band is expanded, the polymer chains have fewer degrees of freedom to wiggle around than in the relaxed state, which increases the temperature of the rubber band.

So-called ‘patchwork approaches’ account for this kind of polysemy by claiming that scientists often systematically employ the same term in different ways, each of which depends on the length scale of investigation, what technique is involved, the specific domain of application and which property scientists target. For example: in the domain of gases, “temperature” involves equations from kinetic gas theory and targets the property *mean kinetic energy* at the molecular scale, whereas in the domain of solids, “temperature” involves restricted ensemble techniques and targets *frozen order* at the polymer scale (Wilson [2017], p. 188). Following pluralists, patchwork approaches assume that words with different scale-dependent, technique-involving, domain-specific and property-targeting uses (‘patches’) have multiple, equally legitimate meanings (Stanford and Kitcher [2000], Brigandt [2010], Hochstein [2016]).

Patchwork approaches combine elements of internalist and externalist semantics to describe how polysemy arises when researchers extend their concepts to related yet subtly different phenomena. Researchers often use the same ‘general reasoning strategy’ when extending a concept to novel cases (internalist element, section 2.1). Such extensions can lead to initially unnoticed changes in how the concept is used, because the hidden hand of nature silently shifted its reference to a novel property (externalist element, sections 2.2 and 3.3). This reference shift requires scientists to update their ‘semantic picture’, a generic representation of how a representational device (linguistic term, diagram, equation etc.) correlates with the world (Wilson [2006], 516, section 4.1). Patchwork approaches are defended via detailed case studies, for example “hardness” and “force” in physics (Wilson [2006], [2017]), “homology” and “species” in biology (Novick [2018], Novick and Doolittle [2021]), “gold” in chemistry (Bursten [2018]) and “neural function”, “cortical column” and “hierarchy” in neuroscience (Haueis [2018],

[2020], Burnston and Haueis [2021]).¹ Echoing the linguistics literature, these case studies suggest that polysemous concepts are a pervasive feature of scientific language which enables successful communication.

In contrast to patchwork approaches, eliminativists call the legitimacy of polysemous scientific concepts into question (Ereshefsky [1992], Machery [2010]). Eliminativists pose a normative challenge to patchwork approaches: even if concepts with multiple related meanings exist in scientific practice, we should eliminate them if they lead to miscommunication and pointless debate in scientific (or philosophical) discourse (Taylor and Vickers [2017]). Existing case studies only justify why scientists do not eliminate a particular concept after it has developed into a patchwork. But the case studies do not provide a *general* answer why legitimate patchwork concepts should not be eliminated. To move beyond case-by-case justification, a general answer to the eliminativist challenge requires normative constraints for evaluating the legitimacy of patchwork concepts. When have scientists extended a concept legitimately because it characterizes phenomena that are related in epistemically significant ways? And when does extending an existing concept to a novel phenomenon fail to foster our descriptive, classificatory or explanatory goals, such that the formation of a new concept becomes necessary?

I answer these questions by developing a generalized patchwork approach which provides normative constraints for legitimate patchwork concepts. I first motivate why we need a separate approach for polysemy in the sciences (section 2.1). I then suggest a general format to model multiple meanings of a scientific term as patches of a patchwork concept, and apply this format to “hardness”, “homology”, “cortical column” and “gold” (section 2.2). The general format reveals three ways in which extending a patch to novel cases can fail: when the technique used to apply it becomes unreliable (section 3.1), when the domain to which it is applied

¹ Further cases—though not explicitly couched in patchwork terms so far—include “acid” (Kitcher and Stanford [2000]), “gene” (Brigandt [2010]) and “water” (Havstadt [2018]). See also the list in Taylor and Vickers ([2017]).

becomes heterogenous (section 3.2), or when the property it targets becomes insignificant for reaching epistemic goals (section 3.3). If such failures persist and accumulate over time, extending a patch becomes illegitimate and scientists should form a novel concept instead (section 3.4). Putting the failures together reveals normative constraints that scientists follow to use individual patches rigorously, and to relate patches to each other without equivocation (section 4.1). I use these normative constraints to defend the legitimacy of scientific patchwork concepts against the eliminativist challenge (section 4.2). My defense reveals that legitimate patchwork concepts have a pragmatic form of unity overlooked by eliminativist accounts. Section 5 concludes.

2 The Patchwork Structure of Scientific Concepts

2.1 Patchwork concepts and polysemy

Polysemous words acquire multiple related meanings when speakers extend them to novel cases. This process is a feature of the flexibility of language, not a bug of defective concepts. Polysemy rarely poses a problem for successful everyday communication since speakers are generally able to decode the relevant meaning from a conversational context (Falkum [2015]). Similarly, I argue that the patchwork structure of scientific concepts rarely poses a problem for scientific communication. The specific investigative context usually allows researchers to select the relevant subset of information from the overall content of the polysemous concept (Wilson [2017], ch. 1; section 4.2).

There are two important differences between everyday polysemy and scientific patchwork concepts. First, everyday polysemy usually occurs at the level of lexical word meaning. Consequently, linguists model how speakers decode the contextually appropriate meaning from a single sentence (Falkum [2015], Vicente [2017]). By contrast, patchwork structures arise when researchers extend a general reasoning strategy to novel cases (Wilson [2006], p. 288, p.

356; [2017], p. 25, p. 226). A general reasoning strategy consists of stepwise instructions which can be realized by different experimental or modeling techniques (section 2.2), and which provide a pragmatic unity to the different uses of a patchwork concept (section 4.2). Patchwork approaches describe how extending a reasoning strategy changes the meaning of concept because it *encodes* a novel subset of information about the behavior of a class of entities. For example: the thermodynamic reasoning strategy of “temperature” instructs researchers to (a) find the energetic equilibrium state of a system and (b) calculate how increases in kinetic energy relate to decreases in degrees of freedom. Extending this strategy from gases to solids shifts the meaning of “temperature”: it encodes information about mean kinetic energy of molecules in the former and information about the frozen order of polymer chains in the latter.

Second, many polysemous words in everyday language possess a *core meaning*. For example: the core meaning of “mouth” is “oral cavity of animals”, and all its other meanings are related to this meaning (Vicente [2017], p. 19). By contrast, patchwork approaches hold that many scientific concepts have a semantic structure that is analogous to the material structure of a quilt. Each patch in a quilt is stitched to neighboring patches, but there is no central piece to which all patches are connected. Analogously, patchwork concepts have different scale-dependent, technique-involving, domain-specific and property-targeting uses (patches) that are related locally to one another. But at least in all cases examined so far, patchwork concepts lack a core meaning to which all uses are related.

This point holds no matter which linguistic model of polysemy we adopt (see Vicente [2017], p. 6). Patchwork concepts lack a privileged meaning from which all others can be derived (literalist model), a total meaning from which word uses recruit a limited set of features (overspecification model) or an abstract and general summary representation, which encompasses and gives access to all its meanings (underspecification model). Consider first literalism. Sometimes physicists suggest that “mean molecular kinetic energy” represents the privileged

meaning of “temperature” (Tadmor and Miller [2012], p. 554). But that would misidentify the meaning which has been diagnosed initially as the most significant one (Wilson [2017], p. 192). Physicists now recognize that “mean molecular kinetic energy” is most informative for gases, whereas “frozen order of polymer chains” is most informative for solids. Neither domain-specific use represents the privileged meaning of “temperature”.

Concerning the overspecification model, consider how researchers use different tests to apply “hardness” to different classes of material:

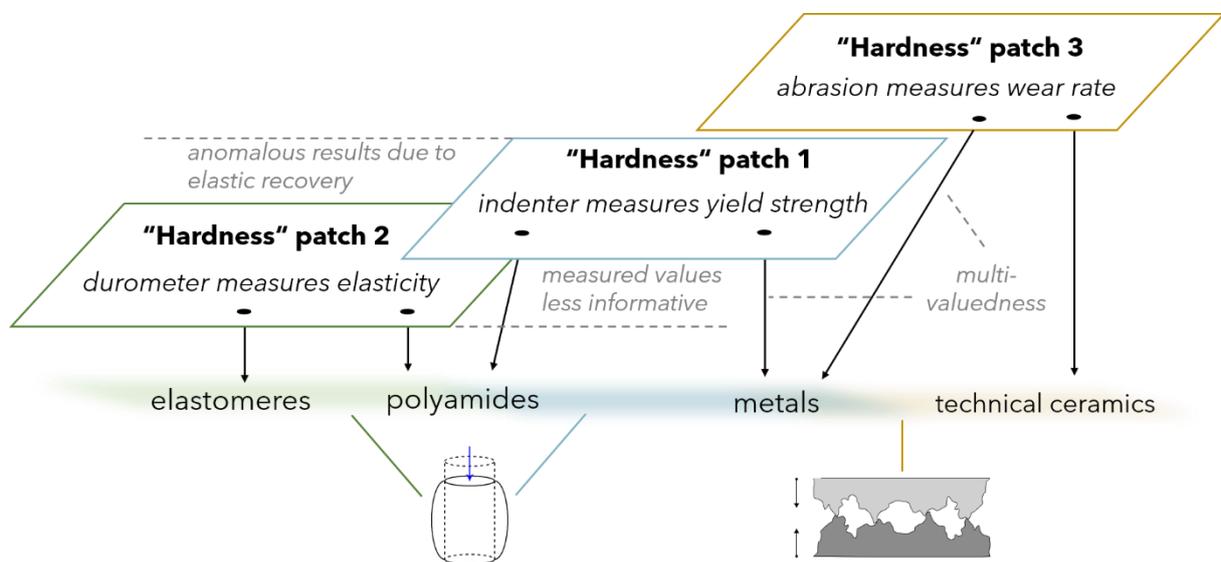


Fig. 1. Simplified model of the patchwork structure of “hardness”, based on Wilson ([2006], p. 338).

For metals, indenter tests measure the physical quantity ‘yield strength’ (patch 1, Fig. 1), whereas for elastomers, durometer tests measure ‘Young’s modulus of elasticity’ (patch 2, Fig. 1). The meaning of “hardness” is derived from such tests because there is no theory that unifies all uses of “hardness”, and the theoretical accounts which do exist are often unreliable to calculate trustworthy hardness values in engineering contexts (Wilson [2006], p. 341; Arnell [2010], p. 60). Nevertheless, indenter and durometer tests suggest that the total meaning of hardness is “resistance to deformation of three-dimensional shape”. Each test recruits a limited set of features from this meaning: resistance to plastic or elastic deformation respectively (blue

and green lines below patch 1 and 2). However, this is only a local relation, because other tests measure physical quantities that resemble, but are not identical to three-dimensional shape deformation. In metals and technical ceramics, abrasion tests measure asperity, i.e. surface roughness (yellow line below patch 3). Asperity determines the ‘wear rate’ of a material when another material is slid across its surface. Because the same material can behave differently when abraded or indented, multiple conflicting hardness values arise from different tests (“multi-valuedness” in Fig. 1, see Wilson [2006], p. 338; Pérez Carballo [2020], p. 22f.). It is partly because of these conflicts that material scientists deny that hardness is a fundamental property (Newey and Weaver [2013], p. 13). Yield strength, elasticity and wear rate are three distinct properties which all objects in the extension of “hardness” possess. Each property is primarily significant for generating knowledge about the behavior of restricted classes of material (with important overlaps, section 4.1). But they are not features of a common property which specifies the total meaning of “hardness”.

Because of the domain-specific epistemic significance of properties, patchwork concepts lack a privileged or a total meaning to which all uses are related. Yet, the general reasoning strategy of a patchwork concept seems to provide an underspecified core meaning. For example: the general reasoning strategy of “hardness” instructs researchers to (a) perform a mechanical intervention on a material and (b) find a quantity which describes how the material resists that intervention. This seems to suggest that “resistance to mechanical intervention” is a general and abstract summary representation which encompasses and gives access to all meanings of “hardness”. The underspecification model certainly comes closest to describe polysemous concepts in science. Yet it fails to capture that the primary role of general reasoning strategies is *instructive*: they tell researchers what operations they need to undertake to reach an epistemic goal. Accordingly, the general reasoning strategy of “hardness” does not determine what this concept refers to. Rather it instructs researchers how to determine properties that are significant

to describe material behavior for engineering purposes. General reasoning strategies thus resemble internalist semantics in which meanings are instructions on how to assemble concepts (Pietroski [2018], pp. 24–9). Analogously, a general reasoning strategy provides researchers instructions on how to assemble patches of a patchwork concept. These instructions are not abstract summary representations to which researchers add features to arrive at the different meanings of a polysemous term. The instructions rather constrain which techniques researchers can use to discover properties that are epistemically significant, i.e. properties that can be used to explain, classify or explain the behavior of entities. Linguistic and semantic theories of polysemy usually do not highlight this pragmatic and instructive role of general reasoning strategies.

Another difference to underspecified core meanings is that general reasoning strategies constrain but do not encompass all technique-involving uses of the same word. Consider “homology”:

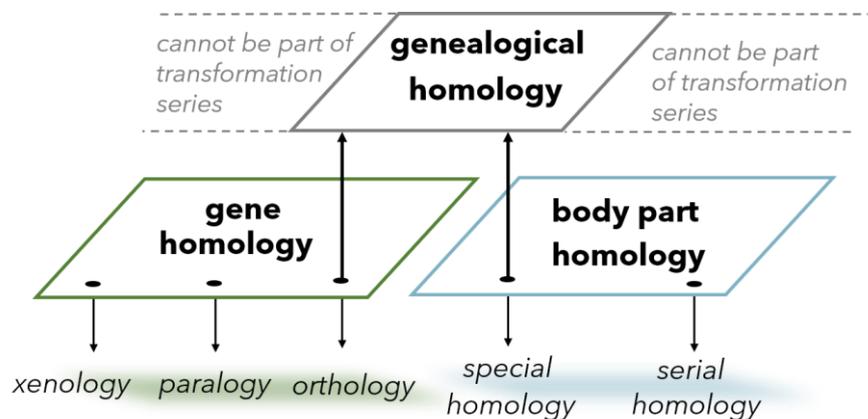


Fig. 2. A simplified representation of the patchwork structure of “homology”, based on Novick ([2018], Fig. 1).

Fig. 2 shows the relations between the general reasoning strategy of ‘genealogical homology’ (grey rectangle) and scale-dependent, technique-involving, domain-specific and property-targeting uses of “homology” (green and blue rectangles). Genealogical homology instructs researchers to search for biological entities which (a) descended from the same entity in a

common ancestor and (b) form a transformation series during evolution (Novick [2018]). By devising techniques which realize this strategy, biologists discovered that two genes are genealogically homologous if they are produced by the gene replication mechanism and are passed via vertical descent between two species (“orthology”, Fig. 2). Two body parts are genealogically homologous if—according to Wagner ([2014])—they are produced by the same gene regulatory network in two species and transform into one another during evolution (“special homology”, Fig. 2). The genealogical reasoning strategy integrates two domain-specific uses of “homology” (black upwards arrows, Fig. 2). It does not, however, abstractly represent all sameness relations that biologists classify as homologies. Extending the very same techniques to novel cases, biologists also discovered genes and body parts that cannot be part of a transformation series (“paralogy”, “xenology”, “serial homology”, Fig. 2). This example shows that while general instructions provide a common thread to some patches, patchwork concepts also have *local* relations which connect these uses to cases which fall outside the general reasoning strategy.

In sum, polysemy in science arises from extending general reasoning strategies rather than single words to novel domains, and the resulting patchwork concepts lack a core meaning. No single use expresses the privileged meaning of a patchwork concept, even though some are historically more prominent than others. There is no total meaning, since frequently there is no common property from which techniques measure different features, only related properties whose epistemic significance varies by domain. And while reasoning strategies do not abstractly represent all meanings, they integrate several uses, which together with local relations provides a pragmatic unity to patchwork concept (section 4.2).

2.2 A general format for representing patchwork structure

One may question the legitimacy of polysemous scientific concepts because without a core meaning, it is unclear what unites different uses of the same word, and how scientists avoid equivocation when a polysemous term is applied to novel cases. The generalized patchwork approach answers these worries by providing norms on the level of a patch, which is a scale-dependent, technique-involving, domain-specific and property-targeting way of using a word. Patchwork concepts have multiple meanings because they have multiple patches. I distinguish four elements which determine the meaning of a patch:

$$“P”^s < t, d(\theta) >$$

The first element of a patch “ P ” is the scale s . Many scientific concepts describe entities that display different properties or behaviors at characteristic length scales. Bursten ([2018]) calls such properties or behaviors ‘scale-dependent’ because they are only instantiated by entities at a characteristic length scale, and because they are described in scale-specific models of a multiscale modeling schema (Batterman [2013], Wilson [2017], ch 5., see section 4.1). I will not use this element separately to evaluate patches because it is unclear in the above writings *when* a change of scale generates a novel meaning. In the cases below all scale changes co-occur with changes of properties (θ). For now, I define s as the minimal spatial, temporal or energetic interval at which an entity instantiates a scale-dependent property or behavior.

The second, operationalist element t specifies an experimental or modeling ‘technique’. An experimental technique instructs researchers how to use a measurement or intervention device to produce a specific experimental result (Colaço [2018]). Modeling techniques instruct researchers how to use a mathematical equation to produce a specific kind of numerical result. Techniques can realize the general reasoning strategy of a patchwork concept because they consist of experimental or mathematical operations that implement each step of the instructions

the strategy provides. Hence scientists attempt to extend a patch to novel cases by applying a specific technique which realizes the strategy. Such extensions fail when researchers realize that the technique produces unreliable results (section 3.1). While each patch has a technique specific to it, a patch is not defined by that technique. In practice researchers use multiple techniques without switching patches (section 4.1) or add novel techniques to the same patch over time.

The third element (domain d) and the fourth element (property θ) resemble externalist semantics (see section 3.4). Each patch specifies a class of entities (domain d) and targets a property θ , which assigns members of that class to the extension of the concept. For example: yield strength (θ) is used to assign members of the class *metal* to the extension of “hardness”. By contrast, Young’s modulus of elasticity (ψ) is used to assign members of the class *elastomer* to “hardness”.² Each domain is thus akin to a partial extension of a patchwork concept (Field [1973], p. 477; Wilson [1982], pp. 551). The number and scope of domains must be discovered empirically (Novick [2018], p. 6). Finally, θ can be a physical quantity, a disposition, a mechanism, or other objective features of scientific interest. The common epistemic function of properties is that they can be ‘significant’. A property is significant if researchers can use it to generate scientific knowledge about the behavior of entities which have that property (section 3.3).

Let us now apply the format to specify the multiple meanings of “hardness”:

- (1) “Hardness”^{Macro} < *indenter, metals* (yield strength) >
- (2) “Hardness”^{Macro} < *durometer, elastomers* (Young’s modulus of elasticity) >
- (3) “Hardness”^{Macro} < *dry sand wear test, technical ceramics* (wear rate) >

These patch-specific meanings are distinguished by the technique used to pick out a physical quantity that is significant to describe the behavior of a class of materials. The patches are

² Throughout the paper I use symbols θ and ψ when talking about two different kinds of properties.

related because their domains partially overlap and their techniques realize the same general reasoning strategy (section 4.1).

For “homology (Fig. 2), we get five patch-specific meanings:

- (1) “Homology”^{gene/phylogenetic} < *phylogenetic tree comparison, orthology* (gene replication) >
- (2) “Homology”^{gene} < *phylogenetic tree comparison, paralogy* (gene duplication) >
- (3) “Homology”^{gene} < *phylogenetic tree comparison, xenology* (horizontal gene transfer) >
- (4) “Homology”^{body part/phylogenetic} < *gene knockout/enhancement, special homology* (conserved GRN)
- (5) “Homology”^{body part} < *gene knockout/enhancement, serial homology* (re-expressed GRN)>

“Homology” patches (1)–(3) involve phylogenetic tree comparison, which instructs biologists to reconcile a phylogenetic species tree with a gene family tree (Altenhoff, Glover and Desimoz [2019]). Gene knockout/enhancement techniques (patches (5)–(6)) instruct researchers to inhibit or excite multiple genes to identify the gene regulatory network which produces a body part (Wagner [2014]). Each technique picks out genealogical homologs when applied to ortholog genes or special homolog body parts, respectively. These entities mark the evolution of novel species at a phylogenetic timescale (included in *s* of patches (1) and (4)). Both techniques can be extended to non-genealogical homologs (patches (2), (3) and (5)).

Let’s now apply the format to two further examples. Consider first “cortical column”, whose reasoning strategy instructs neuroscientists to (a) search for vertical structures and (b) to determine whether neurons within these structures have similar functional properties:

- (1) “Hypercolumn”^{2-3mm} < *tangential recordings, VI, V2, MT* (sequence regularity) >
- (2) “Column”^{0.5mm} < *vertical/tangential recordings, primary sensory areas* (uniform responses) >
- (3) “Minicolumn”^{30-80μm} < *golgi stain, neocortex* (vertical connections > horizontal connections) >

At the microscale “cortical column” picks out minicolumns, cell bands whose vertical, internal connections are stronger than their horizontal, external connections (Mountcastle [1997]). Identifying this property involves anatomical staining techniques, for example golgi staining. At the

mesoscale “cortical column” refers to circuits which uniformly respond to the same sensory stimulus. At the macroscale, many columns form a ‘hypercolumn’ if their uniform responses are sequence regular, they progress in an orderly manner across the cortex (Hubel and Wiesel [1977]). Identifying these functional properties involves electrophysiological recording techniques. These scale-dependent and technique-involving uses of “cortical column” are domain-specific because each of the three properties targeted is found in a restricted set of brain areas.

Finally, consider the chemical kind term “gold” (Bursten [2018]):

- (1) “Gold”^{macro} < *acid test, macroscopic metal objects* (chemically inert) >
- (2) “Gold”^{nano} < *spatial modulation spectroscopy, metal nanoparticles (LSPR)* >
- (3) “Gold”^{atomic} < *atomic orbital model, atoms* (electronic symmetry) >

The analysis highlights that scientists classify instances of “gold” by properties that are characteristic for different length scales (Bursten [2018]). Macroscopic gold lumps are chemically inert, they will react with almost no other chemical. By contrast, a gold nanoparticle will be highly reactive because most atoms lie on the material surface, which enhances the electromagnetic field around the nanoparticle. This scale-dependent property is called localized surface plasmon resonance (LSPR). A gold atom isolated in a vacuum will possess none of these higher-scale properties but only its characteristic atomic properties. Identifying each of these scale-dependent properties involves specific experimental or modeling techniques (Ebbs [2003], Muskens *et al.* [2008], Milton *et al.* [2011]).

The case of “gold” raises the question whether all the examples of patchwork concepts are analogous. For instance, different technique-involving uses of “homology” do not presuppose the same kind of relation between ancestors at different scales of biological organization. By contrast, different technique-involving uses of “gold” presuppose that the samples to which

the techniques are applied belong to the same kind of chemical substance.³ According to natural kind essentialism, what unites the different uses of “gold” is that all samples of gold possess the same atomic property necessarily (Hendry [2006]). If natural kind essentialism about gold is true, then “gold” is not analogous to the other examples. If it is false, then “gold” is analogous because no essential property unifies the different scale-dependent, technique-involving, domain-specific and property-targeting uses of “gold”. Although answering this question is tangential to the discussion below, I do think it depends on there being a property that is equally significant to explain the behavior of macroscopic lumps, nanoparticles and isolated atoms of gold (sections 3.3 and 5). In any case, the general format reveals a conceptually copious use of “gold” in chemistry that is similar to other chemical kind concepts, like “acid” (Stanford and Kitcher [2000]) or “water” (Wilson [2006], p. 428; Havstadt [2018], pp. 725–8).

In sum, the general format introduces four elements that determine the patch-specific meanings of a patchwork concept: scale, technique, domain and property. Patchwork concepts have multiple meanings because each patch involves a different technique, which scientists apply to a specific domain of entities to target a particular scale-dependent property. Taylor and Vickers ([2017]) speak of ‘fragments’ and ‘fragmentation’, highlighting that different meanings of a polysemous concept play different theoretical roles. This terminology conceals, however, that polysemous concepts often only fulfil these roles because their meanings are systematically related to each other (Pérez Carballo [2020], p. 24). I highlight this fact by defining the patchwork structure of a concept as the sum of its patches plus their relations to one another. In a legitimate patchwork concept, these relations help scientists to achieve descriptive, classificatory explanatory goals (section 4.1). To determine the legitimacy of patchwork concepts, I now use the general format to examine various failures which can occur when researchers extend patches beyond their proper boundaries.

³ Thanks to an anonymous referee for suggesting this formulation.

3 How Patches Can Break: Three Kinds of Failure

3.1 Unreliable techniques

The first way in which extending a patch can fail is when the technique it involves permanently and systematically produces unreliable results. Techniques can be reliable or unreliable because scientists frequently use techniques without knowing exactly which property they characterize (Feest [2010], [2011]). A technique t is reliable if it can detect the same property θ in entity e_1 and entity e_2 . Consequently, t is unreliable if using it fails to detect the same θ in e_1 and e_2 . This failure must occur not only temporarily but permanently: it occurs across multiple repeated uses of t . Failure also must be systematic: it is caused by the design of t and not by random intervening factors. One can similarly model the reliability of modeling techniques by substituting “detect” with “calculate” and explicate systematic failure as a failure of the modeling assumptions being met by the target system (for examples, see Wilson [2006] pp. 213–22, pp. 320–1; [2017]), pp. 215–9). Such reliability failures are crucial to determine the proper application range of a technique.

Consider first the reliability of the indenter hardness test. When applied to a steel bar and a brass plate, the test reliably detects yield strength because both deform plastically. The same amount of deformation requires applying a larger force to steel than to brass, from which scientists correctly infer that steel is harder than brass. By contrast, indenters fail to reliably detect yield strength in a rubber ball because its surface will not plastically deform but recover elastically. This failure is permanent because elastic behavior persists across trials, and it is systematic because indenter tests are not designed for elastic surfaces. Materials which are too elastic (or too brittle) lie outside the proper application range of indenter tests. Applying indenter tests to such materials presents an inadequate stretch of patch (1) of “hardness”. Such an extension would mistakenly suggest that rubber balls are harder than steel bars (because they show less

plastic deformation), whereas the proper conclusion should be that rubbers are “hard” in a related but different sense.

Consider next gene knockout/enhancement techniques. According to Wagner ([2014]), they reliably identify body part homology if they detect the same ‘character identity network’. If knockout/enhancement of network genes inhibits/activates the production of a body part, then biologists can infer a character identity network which is either conserved to produce the same body part across species (Oppenheimer *et al.* [1999]) or re-expressed in the same organism to produce serial homology (Hu *et al.* [2019]). Now, consider extending this technique from whole gene networks to single ortholog genes, for example *ey* in drosophila *Pax6* in vertebrates. Both genes are necessary for normal eye development, and *Pax6* and *ey* knockout organisms fail to develop eyes (Wagner [2014], 107f.). This suggests that vertebrate and insect eyes are special homologues. A complex trait like an eye, however, is not produced by a single gene. To assess the functional role of ortholog genes we must situate them in a larger causal context (see Craver [2007], 257). That context is the network of genes and proteins with which *Pax6* and *ey* interact, and those networks are not identical across species (Wagner [2014], 109). Knocking out/enhancing *ey* and *Pax6* fails to reliably identify special homologues because they do not belong to a conserved character identity network. This failure is systematic because knockout/enhancement techniques detect the effect of genes on an entire gene network mechanism. It is permanent because species differences between gene networks will occur repeatedly across experiments. The failed application to *ey/Pax6* shows biologists that enhancement/knockout techniques should be restricted to gene networks underlying body part homology, and that other techniques are required to study gene homology (Novick [2018]).

The two examples illustrate that researchers delineate the boundaries of a patch in part by trying to apply a technique to novel cases. Reliability failures reveal that they should not apply this technique because the case lies outside its proper application range. In a legitimate

patchwork structure, researchers thus use multiple techniques, each being restricted to cases where it produces reliable results (section 4.1).

3.2 Heterogenous domains

The second way in which extending a patch can fail is when its domain contains a heterogenous class of entities. A domain d is homogenous if scientists can use the same property θ to assign each member in d to the extension of the concept. Which type of property is used to determine homogeneity depends on the technique scientists use to study entities in d . Once the property type is determined, homogenous domains correspond to real groupings of properties which are reflected in the patchwork structure of concepts (Bursten [2018], section 4.1). A domain d is heterogenous if for some members of d , researchers use ψ but not θ to assign them to the concept's extension. Entities are not assigned to d simply because they have θ , but rather because θ is significant, i.e. it can be used to achieve an epistemic goal associated with the concept (sections 3.3 and 3.4). This epistemic element makes domains a hybrid ontic/epistemic category: each individual patch describes a class of entities based on shared properties (ontic dimension). But the homogeneity/heterogeneity of these domains depends on which properties best inform our descriptions, classification systems or explanatory models (epistemic dimension). This hybrid nature explains why the domains of two patches can legitimately overlap without being heterogenous (section 4.1).

Homogeneity failures often co-occur with technique failures. Consider extending the domain of “gold” patch (1) to clusters with less than 10.000 atoms (<10nm in size). Above this scale, the acid test reliably detects chemical inertness in metal samples made of gold. Below this scale, however, the test fails because gold nanoparticles are catalytic, i.e. they start to bind oxygen molecules on their surface. This change in chemical behavior can be explained by nanoscale geometry rather than atomic orbital symmetry. At the corners of nanoparticles, the number of gold atom neighbors drops from nine to three or four (Hvolbæk *et al.* [2007], p. 16).

A Generalized Patchwork Approach

The decrease increases the likelihood of binding of oxygen and other molecules, which makes gold nanoparticle catalytic. Extending the domain of “gold” patch (1) to nanoparticles would be inadequate because researchers cannot use ‘being inert’ but instead use ‘being catalytic’ as the property to assign nanoparticles to the extension of “gold”. This homogeneity failure co-occurs with reliability failures because neither the acid test nor atomic models adequately capture the catalytic behavior of gold nanoparticles.

While often co-occurring, reliability and homogeneity failures are distinct because they can come apart. This happens when a technique reliably measures property θ in two entities but θ is only significant to describe, classify or explain the behavior of e_1 but not e_2 . For patch (2) of “cortical column”, neuroscientists use uniform functional responses to describe primary sensory areas as being composed of cortical columns. In the 1970s, researchers argued to extend this domain to cortical barrels in rodents. Barrels are vertical structures with uniform responses because they isomorphically represent the cylindrical shape of the whisker hair (Horton and Adams [2005]). In columns, researchers can use uniform responses to describe columnar functional architecture—how the vertical structures process stimulus features other than its topography (Hubel and Wiesel [1977], p. 2). In barrels, uniform responses do not describe columnar functional architecture because barrels only represent sensory topography (whisker hair location). Extending “cortical column” patch (2) to barrels would create a heterogeneous domain. Researchers would describe topography and not columnar functional architecture, even though vertical recordings reliably detect uniform responses in these structures. This example illustrates that the homogeneity of a domain not only depends on which entities share similar properties. It also depends on the properties being significant to achieve epistemic goals.

3.3 Insignificant properties

The third way in which extending a patch can fail is when property θ becomes insignificant once another property ψ directs the application of a patch. Shifts of significance remain initially unrecognized and are ultimately beyond the control of speakers (Wilson [2006], pp. 102–4, pp. 136–7). Consequently, scientists cannot predict *a priori* how extending a concept changes which property is significant, they can only update their semantic picture after significance shifts have occurred (sections 3.4 and 4.1). The failures discussed above are indicators to evaluate the significance of a property. A property θ is likely significant if (S1) it is the property detected by reliable instances of t , (S2) it can be used to assign entities in d to the concept's extension and additionally, (S3) it contributes to a descriptive, classificatory or explanatory goal associated with the concept.⁴ Consequently θ is likely insignificant if it fails to meet one or more of (S1)–(S3). A patch targeting an insignificant property fails to generate scientific knowledge because scientists cannot use the property to describe, classify or explain the behavior of entities falling under the concept.

There are synchronic failures of significance: at the same stage of conceptual development θ is significant in d_1 but not in d_2 . For “hardness”, yield strength is significant because (S1) it is detected by reliable instances of indenter tests (in metals), (S2) it is used to assign members of the class ‘metal’ to the extension of “hardness” and (S3) it describes when metals endure or break under an applied load. But yield strength becomes insignificant in elastomers because (S1) indenter tests produce unreliable results (section 3.1), (S2) yield strength cannot be used to assign elastomers to the extension of “hardness” and (S3) yield strength contributes little to the description of elastomer behavior (Fig. 3, section 4.1).

⁴ (S3) generalizes Brigandt's ([2010]) proposal to include the epistemic goal into the content of a concept. Brigandt focuses on explanatory goals and selects one goal to justify conceptual change as rational. I emphasize that patchwork concepts serve multiple kinds of goals (description, classification and explanation) at any point in time.

Synchronic significance provides an additional constraint that goes beyond the reliability of techniques. Scientists often use multiple techniques simultaneously to generate empirical knowledge about the entities they investigate (see Feest [2011], 398). But using the same technique t can hinder knowledge generation if t reliably detects a property that is only significant to characterize the behavior of some, but not other entities (section 3.2). Determining synchronic significance restricts reliable uses of t to those entities where scientists can use θ to achieve their epistemic goals. In a legitimate patchwork concept, multiple techniques reliably detect different properties which are synchronically significant for the behavior of different classes of entities (section 4.1).

There are also diachronic failures of significance, i.e. θ is significant in d_I at an earlier but not at a later stage of conceptual development. An example is the property ‘vertical connections > horizontal connections’ from “cortical column” patch (3). This property was initially evaluated as significant: it was (S1) detected by reliable instances of the best anatomical techniques, and thus (S2) used to assign the class ‘vertical cell band’ to the extension of “cortical column” (Mountcastle [1997]). Furthermore, (S3) the property was used to explain why cells in a minicolumn have the same stimulus preference (Haueis [2021]). Newer quantitative analyses, however, suggest that vertical connections are equal to or weaker than horizontal connections (da Costa and Martin [2010], pp. 6–7). The property ‘vertical connections > horizontal connections’ is thus diachronically insignificant. The actual ratio of vertical and horizontal connections (S1) was not reliably detected by earlier anatomical techniques. Thus, (S2) the property could no longer be used to assign the class ‘vertical cell band’ to the extension of “cortical column”. Consequently, neuroscientists no longer refer to ‘vertical connections > horizontal connections’ but to other connectivity properties when attempting to (S3) explain how stimulus-selectivity is organized at the microscale (da Costa and Martin [2010], pp. 8–9). Diachronic failures of significance highlight that scientific concepts can silently change their reference

when scientists attempt to meet the epistemic goal associated with the concept to a higher degree, or when they introduce novel epistemic goals (Brigandt [2010], section 3.4).

Synchronic and diachronic failures of significance can co-occur in the same patchwork concept. For example: LSPR is used to classify gold nanoparticles (Muskins *et al.* [2008]) and to explain their reddish-brown colour. But LSPR fails to be synchronically significant to classify macroscopic gold lumps because it only occurs when the diameter of the particle is shorter than the wavelength of visible light (Bursten [2018]). In addition, atomic weight failed to be diachronically significant and was superseded—at least in the official reading—by nuclear charge as the atomic property that is used to classify gold and other elements in the periodic table (Hendry [2006]).

3.4 When do patch failures necessitate novel concepts?

Unreliable techniques, heterogeneous domains and insignificant properties indicate that scientists have extended a patch beyond its proper boundaries. When do these failures require novel concept, rather extending a patch to a novel discovery? In this section I propose that scientists should form a novel concept if patch failures accumulate over time. A new concept is necessary if (a) a failure at time point t_1 persists at t_2 and (b) more failures occur at t_2 than at t_1 . Condition (a) implies that individual failures are insufficient to necessitate a novel concept. In practice, scientists often “repair” an initially faulty patch extension, for instance by modifying an unreliable technique, re-carving a heterogeneous domain, or continuing their search for significant properties. A patch extension becomes illegitimate only if the failure persists despite scientists’ repair efforts. Condition (b) implies that it is not necessary that all three failures co-occur to justify introducing a new concept. Researchers often judge a patch extension as faulty if further failures accrue, and despite the fact that particular norm remains fulfilled.

Consider the following cases of how cumulative patch failures necessitate novel concepts:

A Generalized Patchwork Approach

- (A) In 1967, molecular biologists proposed a novel kind of homology—“sequence homology”—based on the degree of DNA similarity between chromosomes (Neurath *et al.* ([1967]).
- (B) In the 1970s, some neuroscientists used physiological and anatomical similarities to previous cases to assign cortical barrels to the extension of “cortical column” (section 3.2).
- (C) In 1750s, some chemists used similarities between gold and platinum to call the latter “white gold” (Ebbs [2003], 254).

In (A), researchers initially favored extending “homology” to sequence similarity because it was significant to explain Mendelian inheritance patterns (Egel [2000]). In 1987, however, a number of prominent biologists rejected “sequence homology” (Reeck *et al.* [1987]). This extension (a) involved a technique which persistently failed to distinguish genealogical from nongenealogical forms of homology (Saap [2009], p. 227). Also, (b) DNA similarity failed to be significant to reconstruct the phylogeny of organisms (Reeck *et al.* [1987]). Persistent reliability failure, combined with significance failure, thus led biologists to distinguish DNA similarity from “homology”.⁵

In (B), neuroscientists extended “cortical column” to cortical barrels because their discrete anatomical boundaries were significant to achieve the goal of describing discrete building blocks in the neocortex. Researchers accepted this extension despite domain heterogeneity (section 3.2). However, neuroscientists subsequently formed the concept “cortical isomorph” to describe structures which represent the topography of a body part (Catania [2002]). This concept is necessary because (a) domain heterogeneity persisted as neuroscientists discovered more cortical isomorphs without columnar functional architecture (Horton and Adams [2005]). Also (b) discrete barrel boundaries turned out diachronically insignificant, as neuroscientists no longer used “cortical column” to describe anatomically discrete building blocks (Haeuis [2021]). Persistent homogeneity failure, together with significance failure explains why neuroscientists rejected the barrel extension as illegitimate.

⁵ Thanks to Aaron Novick for drawing my attention to this example.

Finally, (C) conceptualizing platinum as “white gold” was reasonable at first because the acid test (“gold” patch (1)) reliably dissolves samples of gold and platinum (Ebbs [2003], p. 254). The subsequent discovery of atomic properties by analytic chemistry, however, revealed that “white gold” (a) involves a technique which does not reliably distinguish atomic properties of different noble metals. It also (b) creates a heterogenous domain of chemical entities, which fails to achieve the goal of explaining differences in macroscale chemical behavior in terms of differences in lower-scale properties (Bursten [2018]).⁶ In response to this cumulative failure, chemists formed the element name “platinum” and abandoned “white gold”.

The examples above show that cumulative failure occurs over extended periods of time. This explains why in ongoing practice, it is partially indeterminate whether a patch extension is legitimate or whether a novel concept is necessary. In (A)–(C), scientists initially had good reasons to extend existing patches: at the time of discovery, they involved reliable techniques (white gold) or picked out significant properties (cortical barrels, sequence homology). But the legitimacy of these extensions depends on future research. Further experimentation and theorizing determines whether individual failures persist, or if the fulfilment of norms needs to be re-evaluated. Future research also determines if patch extensions meet or fail norms whose fulfilment was unknown to scientists at the time of discovery. For example: researchers initially did not know if “sequence homology” or “white gold” results in domain heterogeneity or not. In ongoing research, it is partially indeterminate whether novel discoveries require broadening or narrowing the extension of a patchwork concept (Ebbs [2003], p. 252; Feest [2010], p. 183).

This partial indeterminacy distinguishes the generalized patchwork approach from externalist semantics. Externalists assume that the extension of kind terms is determined by an ostensive definition and the property present when speakers introduced the term via the definition

⁶ In this case reliability failure is retroactively determined as persistent after researchers re-evaluated the technique after discovering atomic differences between platinum and gold. This supports the claim that cumulative patch failure depends on future research (see below).

(Ebbs [2003], p. 253). This property controls concept application (Nimtz [2018], p. 8). By contrast, the general reasoning strategy of a patchwork concept instructs researchers how to use techniques to search for properties in a domain of entities. But these strategies only describe how researchers extend patchwork concepts. They do not tell us which extension is legitimate. For this, researchers need to ask if the techniques used to realize the strategy are reliable, the domain to which the strategy is extended is homogenous, and whether the properties targeted are epistemically significant (section 4.1). Contra externalism, properties alone do not control the application of a patchwork concept since significance changes both synchronically and diachronically (section 3.3). We can thus only retrospectively, but not prospectively determine if extending a patch to a novel case is legitimate or whether it will accrue failures, which requires a novel concept.

More can be said about the relation between patchwork concepts, externalism and normativity (see Rouse [2014]). For now, I assume that cumulative failures can be used to retrospectively evaluate when extending a patch is illegitimate. This motivates the proposal that legitimate uses of a patchwork concept avoid these failures.

4. The legitimacy of patchwork concepts

4.1 Normative constraints on using polysemous scientific concepts

Combining the three kinds of patch failure reveals normative constraints which scientists follow to use terms with multiple related meanings rigorously and without equivocation. First, scientists switch techniques to keep producing reliable results when extending the same word to novel cases. When extending “hardness” from metals to elastic or brittle materials, material scientists switch from indenter to durometer or abrasion tests (Fig. 1). Similarly, when extending “homology” from homolog body parts to homologous genes, biologists switch from gene

enhancement/knockout techniques to phylogenetic tree comparison (section 3.1). Neuroscientists and chemists similarly switch techniques to reliably detect scale-dependent properties grouped under “cortical column” and “gold”. This reliability constraint keeps scientists from stretching techniques beyond their proper application range.

Second, scientists split the application domain of a patchwork concept into multiple homogenous domains based on which property θ is most informative to describe, classify or explain the behavior of a class of entities. Material scientists split the application domain of “hardness” into the domains of elastomers, metals and technical ceramics because different properties are significant to describe their behavior. Similarly, biologists split “homology” into different domains based on which particular mechanism is significant to explain the sameness of genes or of body parts. To describe the functional architecture of the cortex, neuroscientists split the domain of “cortical column” into columnar, hypercolumnar and minicolumnar domains. And to classify scale-dependent properties of gold, chemists split the application domain of “gold” into the homogenous domains of chemically inert macroscale lumps, catalytic nanoparticles and isolated gold atoms. This homogeneity constraint is crucial to avoid lumping multiple distinct phenomena into one domain and splitting instances of the same phenomenon across multiple domains (Craver [2007], pp. 123–4).

Third, scientists update their semantic picture to capture which properties different patches refer to, and why these properties are significant to reach an epistemic goal. For example: material scientists tried and failed to develop a notion of “absolute hardness” which identifies hardness with a single physical quantity like yield strength (Williams [1942]). In contrast, what has proven immensely successful are material property charts, which describe the relation between multiple different quantities of “hardness”:

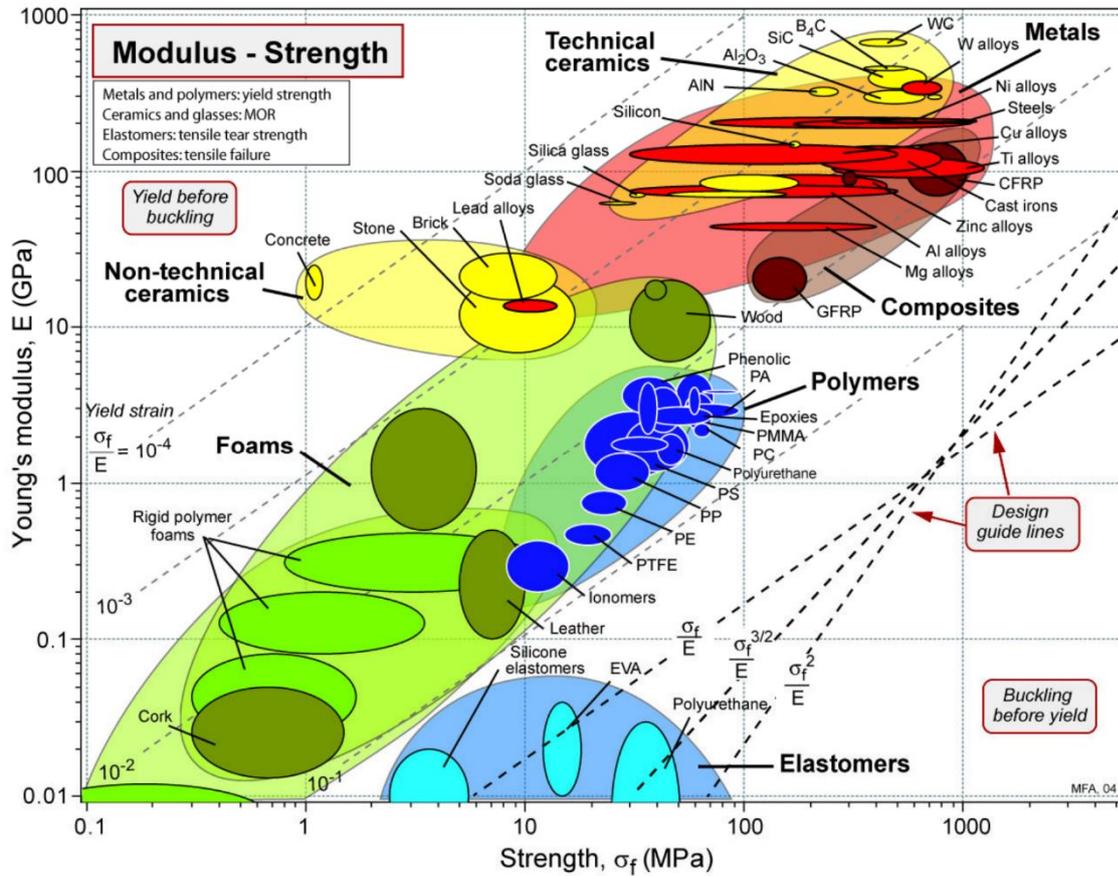


Fig. 3 Relationship of Young’s modulus of elasticity measured in Gigapascal and strength measured in Megapascal for different kinds of material (coloured circles). Chart from CES EduPack 2010, Granta Design Limited, Cambridge, UK, 2010.

Fig. 3 represents parts of the semantic picture of “hardness”. It displays how strength and elasticity are distributed in different classes of material. This semantic picture of “hardness” accounts for how material scientists use “hardness” when designing safe steel bridges, flexible strings or anti-scratch glass. Rather than univocally referring to the same property of *hardness* in all materials, “hardness” partially refers to more than one physical quantity (Field 1973, p. 475; Stanford and Kitcher [2000], p. 119). These quantities best account for the behavior of restricted classes of material that matter for engineering purposes. The significance constraint thus ensures that scientists avoid semantic pictures which correlate concepts with properties that fail to adequately describe, classify or explain the behavior of these entities.

Legitimate patchwork concepts acquire multiple meanings because scientists switch techniques, split domains and update semantic pictures in their attempt to meet the reliability, homogeneity and significance constraints on individual patches. These normative constraints also explain when scientists can legitimately relate multiple patches to one another to achieve their epistemic goals in scientific practice. This epistemic role of legitimate relations is part of the pragmatic unity of patchwork concepts overlooked by eliminativists (section 4.2). First, the reliability constraint allows scientists to reuse a technique across patches. Biologists reuse phylogenetic methods across “homology” patches (1), (2) and (3), because they can be used to reliably detect orthologs and paralogs, and under certain conditions also xenologs (Altenhoff, Glover and Dessimoz [2019], Ravenhall *et al.* [2015]). Neuroscientists reuse tangential recordings across “cortical column” patches (1) and (2). In some areas, they reliably detect functional boundaries between columns (by showing where neuronal responses shift across the cortex) and whether uniform responses of multiple columns are sequence regular (by showing that they shift in an orderly manner, Hubel and Wiesel [1977], pp. 24–5). By reusing techniques across patches, researchers connect multiple domain-specific, scale-dependent and property-targeting uses of a concept into a larger patchwork structure.

Second, the homogeneity constraint allows that two patches have overlapping domains. An overlap region O is legitimate if θ from d_1 or ψ from d_2 can be used equivalently to assign entities in O to the extension of a concept. For example: the overlap region of “hardness” patches (1) and (2) is the class of polyamides, which tells material scientists that indenter and durometer tests produce convergent results for materials like nylon. This convergence allows them to use yield strength or Young’s modulus of elasticity equivalently to assign nylon to the extension of “hardness”. It also tells them something about the world, i.e. that polyamides combine the elasticity of rubbers with the strength of metals (Fig. 3). Similarly, the overlap region of “homology” patches (1) and (4) is the class of genealogical homologs, which tells biologists

that certain kinds of genes and body parts transform into one another on an evolutionary time-scale. This fact is important to integrate knowledge about homology from phylogenetics, comparative anatomy and developmental genetics (McCune and Shimenti [2012]). By letting patches intersect where domains legitimately overlap, scientists create patchwork structures which are not only of theoretical and experimental importance, but also informative about the structure of the world.

Finally, the significance constraint allows that scientists combine properties from several patches to achieve epistemic goals of a concept. One pertinent example is multiscale modeling, where researchers link multiple scale-dependent uses of a concept together to describe, classify or explain the behavior of an entity (Batterman [2013], Wilson [2017], ch. 5). Consider the famous ‘ice-cube model’ (Hubel and Wiesel [1977]):

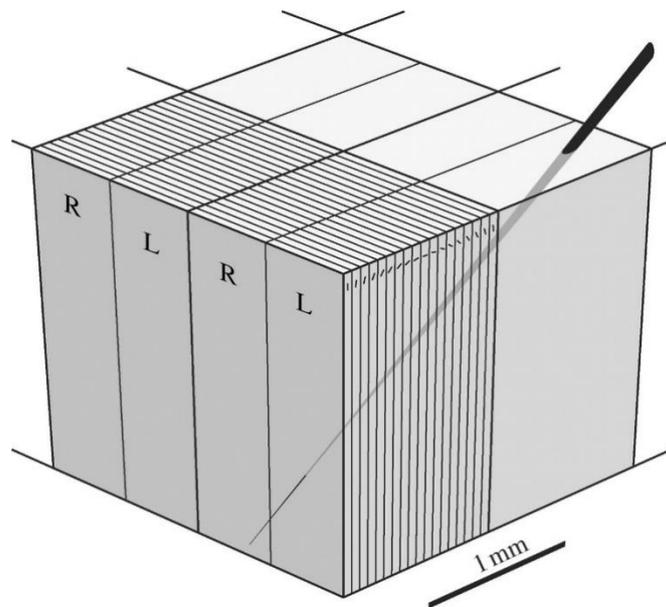


Fig. 4: Ice cube model of the primary visual cortex. Adopted from Horton and Adams ([2005]).

This model combines “cortical column” patches (1)–(3) to describe primary visual cortex as composed of orientation columns (minicolumn scale) that are sequence regular (hypercolumn scale), and which orthogonally intersect with two ocular dominance columns (column scale). The ice-cube model was taken to adequately describe the functional architecture in the

neocortex until researchers re-evaluated the significance of minicolumnar properties and updated their semantic picture of “cortical column” (sections 3.3 and 4.2). Another instance of combining scale-dependent properties are scale-bridging explanations (Bursten [2018]). For example, electronic symmetry (atomic scale) is significant to explain the chemical inertness of macroscale gold lumps (Hvolbæk *et al.* [2007], p. 15). When explaining the reactivity of nanoparticles, however, researchers emphasize molecular geometry rather than electronic symmetry. This example suggests that a conceptual patchwork structure provides multiple ways to combine properties in scale-bridging explanations, without any single property being equally significant in all explanatory contexts.

In sum, we can evaluate the legitimacy of a patchwork concept by asking whether (i) the techniques, domains and properties of individual patches fulfil the reliability, homogeneity and significance constraint and whether (ii) multiple patches are related via the reuse of reliable techniques, the intersection of legitimately overlapping domains and the combination of significant properties in multiscale modeling or scale-bridging explanations.

4.2 Meeting the eliminativist challenge to polysemy

The normative constraints on legitimate patches and their relations now allow us to address the eliminativist challenge to polysemy in science. I focus on Taylor and Vickers ([2017]) because they agree with previous case studies that the formation of patchwork concepts is a ubiquitous and perhaps inevitable feature of scientific development. Yet, they deny the normative conclusion that we “should continue to use the original term in our investigations” (Taylor and Vickers, p. 20). Under certain circumstances, scientists should get rid of the superordinate patchwork concept and only use the precisely defined patches. Patchwork concepts like “hardness” should be selectively eliminated:

A Generalized Patchwork Approach

In a context in materials science where the differences between the [patches] really matter, there is just no point to using the word ‘hard’—it does not communicate *enough* to help. And there is a danger that different interlocutors will interpret ‘hard’ in different ways and miscommunicate. Further, one just does not *need* the word ‘hard’ to communicate about materials – quite clearly, one can say everything one wants to say about materials using other terms such as [“yield strength”, “Young’s modulus” or “wear rate”] (Taylor and Vickers [2017], p. 33).

The generalized patchwork approach agrees with selective eliminativism that the word ‘hard’ does not itself encode specific information about material behavior. That work is done by patches specifying which technique is reliable to apply “hardness” to a homogenous domain of entities with the same significant property. So, my first response is if each patch is properly normatively constrained, then there is no issue with using the polysemous word in scientific practice. This seems to be the case for “hardness” and also—to a large extent—for the other examples.

Eliminativists will reply that there are *risks* associated with keeping the superordinate polysemous term (Taylor and Vickers [2017], 28f.). First, the more theoretical roles (i.e. epistemic goals) a term is used for, the more different meanings it acquires and thus the greater the risk of miscommunication becomes. Second, if the concept is pivotal to reach the conclusion of an argument, there is a risk of pointless debate that rests on different understandings of the term.

Regarding miscommunication, Taylor and Vickers ([2017], p. 32) do point to linguistic work on polysemy, but do not mention the observation that polysemous words rarely pose a problem for successful everyday communication (Taylor [2003], p. 647; Falkum [2015], p. 84). Similarly, polysemous patchwork concepts need not increase the risk of miscommunication among scientists trained to use them in a context-sensitive manner. For “hardness”, researchers learn how to use hardness tests within their proper application range (reliability constraint), and how to use material property charts to describe which properties are significant to describe how

classes of material behave (homogeneity and significance constraint). These charts illustrate why polysemy is not eliminated by talking directly about patches. The general quantity of strength, for example, refers to one of four specific quantities depending on the material (Fig. 3, upper left corner). Researchers retain superordinate terms like “strength” or “hardness” to convey that these quantities are related to one another. These local relations are crucial to achieve the epistemic goals of a concept.

Besides local relations, there is a general reasoning strategy for each patchwork concept. Together, the relations and the reasoning strategy integrate different patches, which gives patchwork concepts a pragmatic unity. Two patches “ P_1 ” $\langle t_1, d_1(\theta) \rangle$ and “ P_2 ” $\langle t_2, d_2(\psi) \rangle$ are integrated if

- (i) t_1 and t_2 realize the same general reasoning strategy
- (ii) t_1 can be re-used to reliably detect some instances of ψ or t_2 can be re-used to reliably detect some instances of θ
- (iii) there is a legitimate overlap region O between d_1 and d_2
- (iv) θ and ψ can be combined to achieve an epistemic goal associated with the patchwork concept

Condition (i) ensures that two technique-involving uses are patches of the same patchwork concept because they implement the stepwise instructions of the same general reasoning strategy. Conditions (ii)–(iv) specify the extent of integration by ensuring that the local relations between patches respect the normative constraints on each related patch. For example: “hardness” patches (1)–(3) are integrated. Each hardness test (i) realizes the same reasoning strategy because it (a) mechanically intervenes on a material and (b) quantifies the material’s resistance to this intervention (section 2.1). Additionally, (iii) indenter and durometer patches legitimately overlap for polyamides, and (iv) material property charts combine multiple hardness quantities to describe the behavior of different classes of material. Selective eliminativism overlooks the extent to which multiple patches of a polysemous patchwork concept are integrated. The strength of the generalized patchwork approach is that it explains how this pragmatic unity emerges from the epistemic practices of scientists being responsive to the world. It does so

parsimoniously, without the need of positing any semantic or metaphysical unifier (a core meaning or essential property). And for “hardness” and analogous concepts we have reasons to suspect that there indeed is no such unifier.

Regarding pointless debate, Taylor and Vickers vacillate between scientific and philosophical debates about a concept. If they mean scientific debates, they provide no evidence that material scientists engage in pointless debates about “hardness” which justify eliminating the term. My approach suggests that researchers avoid pointless debates by following the reliability, homogeneity and significance constraint when using “hardness”. If Taylor and Vickers mean philosophical debates, then the patchwork approach shows that philosophical confusions can arise from a faulty semantic picture of the reference of the polysemous concept. For example: Descartes’ and Reid’s debate about the subjective or objective nature of hardness dissolves once we recognize the patchwork of tests used to probe the resistance of a material, both in everyday life and science (Wilson [2006], p. 335). Pointless debates can be avoided by updating our semantic picture of a patchwork concept without eliminating that concept. The success of material property charts in reflecting the worldly correlates of “hardness” suggests that adequate semantic pictures play an important role in science. Philosophers of science can contribute to this role by articulating better semantic pictures of patchwork concepts (Haueis [2018], [2020], Burnston and Haueis [2021], Novick and Doolittle [2021]).

Taylor and Vickers may admit that relations between patches and semantic pictures are important but insist on selective elimination when polysemy does lead to miscommunication and pointless debate. The concept “cortical column” in neuroscience seems to exemplify this situation: it has multiple different meanings which can lead to miscommunication and which affect the debate about the functional significance of columns. A patchwork approach, however, shows that that elimination of “cortical column” is not necessary (Haueis [2021]). First, as Rakic ([2008], p. 12099–100) points out, “the term “column” is used in so many ways that it

can be very confusing to the non-specialist". Like in the "hardness" case, it is unclear why miscommunication should really affect neuroscientific experts trained to use the term in a scale-dependent, technique-involving, domain-specific and property-targeting manner.

Second, neuroscientists resolved the debate about functional significance by updating their semantic picture of "cortical column". According to the 'building block picture', "cortical column" refers to a module that processes information in the same way in every cortical area. Subsequent discoveries, however, showed that not all areas have a columnar architecture, and that scale-dependent columnar properties occur independently or are insignificant. In response to these discoveries, researchers now hold that "cortical column" refers to different *kinds* of scale- and area-specific columnar structures. When neuroscientists criticize that the column is a "structure without a function" (Horton and Adams [2005], p. 837), they target the old building block picture of "cortical column". This picture entails that columns are the basic functional entity of the cortex. The new semantic picture, however, acknowledges that some columnar structures are functional whereas others are not. So "cortical column" does not refer to a basic functional entity, but partially refers to different columnar structures of variable functional significance. Researchers should keep "cortical column" to understand how knowledge about these structures can or cannot be combined to achieve epistemic goals (Haueis [2021]). Rather than being pointless, this debate shows that it is crucial to determine which semantic picture of a concepts' reference accords with the empirical facts and which one does not.

My response to the eliminativist challenge does not reflect deep disagreements with Taylor and Vickers but rather differences in emphasis. While focusing on the danger of verbal disputes and miscommunication, eliminativists can (and should) acknowledge that relations between patches are often crucial to achieve epistemic goals (see Hampton [2010], Machery [2010], 238, Pérez Carballo [2020], p. 24). And while focusing on the merits of linking patches

together, patchwork approaches do (and should) acknowledge the importance of dissolving philosophical debates that rest on confusions about the meaning of a concept.

5 Conclusion

In this paper, I defended the legitimacy of polysemous concepts in science, in part, by introducing a generalized patchwork approach to scientific concepts. Patchwork concepts are characterized by multiple patches, i.e. scale-dependent, technique-involving, domain-specific and property-targeting uses of a word. Yet, there is no core meaning to which all these uses are related. Individual patches are legitimate if they involve reliable techniques, are applied to homogenous domains, and refer to properties which are significant to describe, classify or explain the behavior of entities in the domain. To reach these epistemic goals, researchers relate multiple patches of a concept by reusing techniques across patches, letting domains legitimately overlap and combining scale-dependent properties in multiscale models or scale-bridging explanations. Together with a general reasoning strategy, these normatively constrained local relations integrate patches to give patchwork concepts a pragmatic form of unity. The general structure of patchwork concepts shows how scientists maintain seemingly equivocal uses of the same scientific term without eliminating the legitimate patchwork concept of which they are instances. Researchers avoid miscommunication by following the reliability, homogeneity and significance constraints, and they can dissolve pointless debates by articulating better semantic pictures of what a patchwork concept refers to.

The normativity of the generalized patchwork approach makes it a potential alternative to theories of scientific concepts which posit a unifier that underlies conceptual variation. I already mentioned natural kind essentialism, which posits an essential property that unifies different uses of a natural kind term like “gold” (Nimtz [2018]). Similarly, semantic holism posits a unifier because it holds that the meaning of a concept like “force” is determined by its relation to axioms of a unified theory (Andreas [2010]). The preceding discussion suggests that

scientists could violate normative constraints if they follow essentialist or holist principles. For example, concerning “gold”, Nimtz ([2018], p.15) claims that the atomic property of ‘being Au’ is the pivotal explanatory property which explains why “paradigmatic samples of gold exhibit compellingly many of those characteristics chemistry itself deems most relevant”. But singling out ‘being Au’ as the only pivotal explanatory property may violate the significance constraint: whereas atomic properties are most relevant to explain the inertness of macroscopic gold lumps, molecular geometry is most relevant to explain the reactivity of gold nanoparticles.

Semantic holists like Andreas ([2010]) claim that the method of determining the extension of “force” involves $F=ma$ as a theoretical axiom. One may worry that singling out $F=ma$ as the only fundamental equation violates the reliability constraint. In fact, physicists combine many different modeling techniques to reliably calculate the forces acting on an object. For billiard ball collisions, standard Newtonian techniques become unreliable once high-impact velocities plastically deform the balls. Physicists switch to continuum mechanics to calculate the impact of internal stress waves, or to laminate modeling to calculate interfacial forces when the balls crack internally after impact. Thus “force” seems a patchwork concept which partially refers to different scale-dependent behaviors described in a multiscale modeling schema (Wilson [2006], pp. 180–3; [2017], p. 38, pp. 339–40).

Further work is needed to compare the generalized patchwork approach to these and other theories of scientific concepts. But if the normative constraints developed here can be fruitfully applied to paradigmatic cases of such theories, then the patchwork approach provides a genuine alternative of how concepts can be structured in our best scientific practices.

6 References

- Altenhoff, A.M., Glover, N.M., Dessimo, C. ([2019]): ‘Inferring Orthology and Paralogy’. In M. Anisimova (Ed.), *Evolutionary Genomics*, New York, NY: Springer New York, pp. 149–75
- Ashby, M. ([2010]): *Granta CES EduPack Resource Booklet 2: Material and Process Selection Charts*.
- Andreas, H. ([2010]): ‘Semantic Holism in Scientific Language’. *Philosophy of Science*, **77**, pp. 524–43.
- Arnell, D. ([2010]): ‘Mechanisms and Laws of Friction and Wear’. In Rahnejat, H. (ed.). *Tribology and Dynamics of Engine and Powertrain*. Sawston: Woodhead Publishing, pp. 41–72.
- Batterman, R. ([2013]): ‘The Tyranny of Scales’, in R. Batterman (Ed.), *The Oxford Handbook of Philosophy of Physics*, Oxford University Press, pp. 255–86
- Brigandt, I. ([2010]): ‘The Epistemic Goal of a Concept: Accounting for the Rationality of Semantic Change and Variation’, *Synthese*, **177**, pp. 19–40.
<https://doi.org/10.1007/s11229-009-9623-8>.
- Burnston, D., & Haueis, P. ([2021]): ‘Evolving Concepts of "Hierarchy" in Systems Neuroscience’, in F. Calzavarini & M. Viola (eds.), *Neural Mechanisms: New challenges in the philosophy of neuroscience*. Cham: Springer. <https://doi.org/10.1007/978-3-030-54092-0>
- Bursten, J. R. ([2018]): ‘Smaller than a Breadbox: Scale and Natural Kinds’, *The British Journal for the Philosophy of Science*, **69**. <https://doi.org/10.1093/bjps/axw022>.
- Creighton, H. B., & McClintock, B. ([1931]): ‘A Correlation of Cytological and Genetical Crossing-Over in Zea Mays’, *Proceedings of the National Academy of Sciences of the United States of America* **17**, pp. 492–497.
- Colaço, D. ([2018]): ‘Rethinking the Role of Theory in Exploratory Experimentation’, *Biology & Philosophy*, **33**, 38. <https://doi.org/10.1007/s10539-018-9648-9>.
- Craver, C. F. ([2007]): *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*, Oxford: Oxford University Press.
- da Costa, N. M., & Martin, K. A. C. ([2010]): ‘Whose Cortical Column Would that Be?’, *Frontiers in Neuroanatomy*, **4**, 16. [https://doi.org/10.3389/fnana.\[2010\].00016](https://doi.org/10.3389/fnana.[2010].00016).
- Ereshefsky, M. ([1992]): ‘Eliminative Pluralism’, *Philosophy of Science*, **59**, pp. 671–90.

- Egel, R. ([2000]): 'How 'Homology' Entered Genetics', *Trends in Genetics*, **16**, pp. 437–9. [https://doi.org/10.1016/s0168-9525\(00\)02097-7](https://doi.org/10.1016/s0168-9525(00)02097-7).
- Ebbs, G. ([2003]): 'Denotation and Discovery', in Schmitt F. F. (ed.), *Socializing Metaphysics: The Nature of Social Reality*. Lanham: Rowman and Littlefield, pp. 247–68.
- Falkum, I. L. ([2015]): 'The How and Why of Polysemy: A Pragmatic Account', *Lingua*, **157**, pp. 83–99. [https://doi.org/10.1016/j.lingua.\[2014\].11.004](https://doi.org/10.1016/j.lingua.[2014].11.004).
- Field, H. (1973): 'Theory Change and the Indeterminacy of Reference', *The Journal of Philosophy*, **70**, pp. 462–81. <https://doi.org/10.2307/2025110>.
- Feest, U. ([2010]): 'Concepts as Tools in the Experimental Generation of Knowledge', *Spontaneous Generations* **4**, pp. 173–90.
- Feest, U. ([2011]): 'Remembering (Short-Term) Memory: Oscillations of an Epistemic Thing', *Erkenntnis*, **75**, pp. 391–411. <https://doi.org/10.1007/s10670-011-9341-8>
- Hampton, J. A. ([2010]): 'Concept Talk Cannot be Avoided', *Behavioral and Brain Sciences* **30**, pp. 212–13 <https://doi.org/10.1017/S0140525X1000036>.
- Haueis, P. ([2018]): 'Beyond Cognitive Myopia: A Patchwork Approach to the Concept of Neural Function', *Synthese*, **195**, pp. 5373–402. <https://doi.org/10.1007/s11229-018-01991-z>.
- Haueis, P. ([2021]): 'The Death of the Cortical Column? Patchwork Structure and Conceptual Retirement in Neuroscientific Practice', *Studies in History and Philosophy of Science Part A*, **85**, pp. 101–13, <https://doi.org/10.1016/j.shpsa.2020.09.010>.
- Havstadt, J. C. ([2018]): 'Messy Chemical Kinds', *British Journal for the Philosophy of Science*, **69**, pp. 719–43. <https://doi.org/10.1093/bjps/axw040>.
- Hendry, R. F. ([2006]): 'Elements, Compounds, and Other Chemical Kinds', *Philosophy of Science*, **73**, pp. 864–75. <https://doi.org/10.1086/518745>.
- Horton, J. C., & Adams, D. L. ([2005]): 'The Cortical Column: A Structure without a Function', *Philosophical Transactions of the Royal Society of London B*, **360**, pp. 837–62. [https://doi.org/10.1098/rstb.\[2005\].1623](https://doi.org/10.1098/rstb.[2005].1623).
- Hochstein, E. ([2016]): 'Categorizing the Mental', *The Philosophical Quarterly*, **66**, pp. 745–59. <https://doi.org/10.1093/pq/pqw001>.
- Hu, Y., Linz, D. M., & Moczek, A. P. ([2019]): 'Beetle Horns Evolved from Wing Serial Homologs', *Science*, **366**, pp. 1004–7. <https://doi.org/10.1126/science.aaw2980>.

- Hubel, D. H., & Wiesel, T. N. ([1977]): ‘Ferrier Lecture 1972. Functional Architecture of Macaque Monkey Visual Cortex’, *Philosophical Transactions of the Royal Society of London B*, **198**, pp. 1–59. <https://doi.org/10.1098/rspb.1977.0085>.
- Hvolbæk, B., Janssens, T. V.W., Clausen, B. S., Falsig, H., Christensen, C. H., & Nørskov, J. K. ([2007]): ‘Catalytic Activity of Au Nanoparticles’, *Nano Today*, **2**, pp. 14–8. [https://doi.org/10.1016/S1748-0132\(07\)70113-5](https://doi.org/10.1016/S1748-0132(07)70113-5).
- Machery, E. ([2010]): ‘Précis of Doing without Concepts’, *Behavioral and Brain Sciences*, **33**, pp. 195–206; discussion pp. 206–44. <https://doi.org/10.1017/S0140525X09991531>.
- Milton, O., Macomber, R. S., Pinhas, A. R., & Wilson, R. M. ([2011]): ‘Atomic Orbital Theory’ in M. Orchin (ed.), *The vocabulary and concepts of organic chemistry*, 2nd ed., Hoboken, N.J: Wiley-Interscience, pp. 1–24.
- Muskens, O. L., Billaud, P., Broyer, M., Del Fatti, N., & Vallée, F. ([2008]): ‘Optical Extinction Spectrum of a Single Metal Nanoparticle: Quantitative Characterization of a Particle and of its Local Environment.’ *Physical Review B*, **78**, <https://doi.org/10.1103/PhysRevB.78.205410>.
- Neurath, H., Walsh, K. A. & Winter, W.P. ([1967]): ‘Evolution of Structure and Function of Proteases: Amino Acid Sequences of Proteolytic Enzymes Reflect Phylogenetic Relationships’, *Science*, **158**, pp. 1638–44.
- Newey, C. and Weaver, G. ([2013]): *Materials Principles and Practice: Electronic Materials, Manufacturing With Materials, Structural Materials*. Amsterdam: Elsevier.
- Nimtz, C. ([2018]): ‘How Science and Semantics Settle the Issue of Natural Kind Essentialism’, *Erkenntnis*, **86**, pp. 149–70. <https://doi.org/10.1007/s10670-018-0098-1>.
- Novick, A. ([2018]): ‘The Fine Structure of ‘Homology’’. *Biology & Philosophy*, **33**. <https://doi.org/10.1007/s10539-018-9617-3>.
- Novick, A. and Doolittle, F. W. ([2021]): ‘‘Species’ without Species’, *Studies in History and Philosophy of Science Part A*, **87**, pp. 72–80. <https://doi.org/10.1016/j.shpsa.2021.03.006>.
- Oppenheimer, D. I., MacNicol, A. M., & Patel, N. H. ([1999]): ‘Functional conservation of the wingless–engrailed interaction as shown by a widely applicable baculovirus misexpression system. *Current Biology*, **9**, pp. 1288–96. [https://doi.org/10.1016/S0960-9822\(00\)80050-0](https://doi.org/10.1016/S0960-9822(00)80050-0).
- Pietroski, P. M. ([2018]): *Conjoining Meanings. Semantics Without Truth Values*, Oxford: Oxford University Press.

- Pérez Carballo, A. ([2020]): ‘Conceptual Evaluation: Epistemic’, In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual Engineering and Conceptual Ethics*. Oxford: Oxford University Press, pp. 304–332.
- Rakic, P. ([2008]): ‘Confusing Cortical Columns’, *Proceedings of the National Academy of Sciences of the United States of America*, **105**, pp. 12099–100.
<https://doi.org/10.1073/pnas.0807271105>.
- Ravenhall, M., Škunca, N., Lassalle, F., & Dessimoz, C. ([2015]): ‘Inferring Horizontal Gene Transfer’, *PLoS Computational Biology*, **11**, e1004095. <https://doi.org/10.1371/journal.pcbi.1004095>.
- Reeck, G. R., de Haën, C. Teller, D.C., Doolittle, R.F. Fitch, W. M. Dickerson, R. E., Chambon, P. *et al.* (1987): ‘‘Homology’’ in Proteins and Nucleic Acids: A Terminology Muddle and a Way out of It’, *Cell* **50**, p. 667. [https://doi.org/10.1016/0092-8674\(87\)90322-9](https://doi.org/10.1016/0092-8674(87)90322-9).
- Rouse, J. ([2014]): ‘Temporal Externalism and the Normativity of Linguistic Practice’, *Journal for the Philosophy of History* **8**, pp. 20–38.
- Sapp, J. ([2009]): *The New Foundations of Evolution: On the Tree of Life*, Oxford: Oxford University Press.
- Stanford, P. K., & Kitcher, P. ([2000]): ‘Refining the Causal Theory of Reference’, *Philosophical Studies*, **97**, pp. 97–127. <https://doi.org/10.1023/A:1018329620591>.
- Tadmor, E. B., & Miller, R. E. ([2012]): *Modeling Materials: Continuum, Atomistic and Multiscale Techniques*, Cambridge: Cambridge University Press.
- Taylor, H., & Vickers, P. ([2017]): ‘Conceptual Fragmentation and the Rise of Eliminativism’, *European Journal for Philosophy of Science*, **7**, pp. 17–40.
<https://doi.org/10.1007/s13194-016-0136-2>.
- Taylor, J. R. ([2003]): ‘Polysemy’s Paradoxes’, *Language Sciences*, **25**, pp. 637–55.
[https://doi.org/10.1016/S0388-0001\(03\)00031-7](https://doi.org/10.1016/S0388-0001(03)00031-7).
- Vicente, A. ([2017]): ‘Polysemy and Word Meaning: An Account of Lexical Meaning for Different Kinds of Content Words’, *Philosophical Studies*, **175**, pp. 947–68.
<https://doi.org/10.1007/s11098-017-0900-y>.
- Wagner, G. P. ([2014]): *Homology, Genes, and Evolutionary Innovation*, Princeton: Princeton University Press.
- Williams, S. R. ([1942]): *Hardness and Hardness Measurements*, Cleveland: American Society for Metals.

Wilson, M. ([1982]): ‘Predicate Meets Property’, *The Philosophical Review*, **91**, pp. 549–581.

<https://www.jstor.org/stable/2184801>

Wilson, M. ([2006]): *Wandering Significance: An Essay on Conceptual Behavior*, Oxford: Oxford University Press.

Wilson, M. ([2017]). *Physics Avoidance: Essays in Conceptual Strategy*, Oxford: Oxford University Press.

Acknowledgments

I thank Daniel Burnston, Carl Craver, Marie Kaiser, Edouard Machery, Christian Nimtz, Aaron Novick, Christian Feldbacher-Escamilla, Rose Trappes, Mark Wilson, and audiences at the GWP Conference at HHU Düsseldorf 2019, the I²SOS Colloquium and the Research Seminar Theoretical Philosophy at Bielefeld University 2019/2020, and the Lunchtime Talk Series at the Center for Philosophy of Science 2020 for helpful feedback on previous versions of this material. Two anonymous reviewers provided insightful feedback which greatly improved the structure and argument of the paper. This project was supported by a stipend of the German Academic Research Council (grant no. 408804291) for a Visiting Scholarship at the Center for Philosophy of Science, University of Pittsburgh.

Philipp Haueis

Department of Philosophy

Bielefeld University

Germany