Philosophy of Science (2025), **92**, 1117–1127
doi:10.1017/psa.2025.10114

PSA PHILOSOPHY OF SCIENCE ASSOCIATION

PHILOSOPHY OF SCIENCE

CONTRIBUTED PAPER

Thick Concepts and Impartiality

Mark Risjord[1], Kareem Khalifa[2] and Jared Millson[3]

[1]Philosophy Department, Emory University, Atlanta, GA, USA, [2]Philosophy Department, UCLA, Los Angeles, CA, USA and [3]Philosophy Department, Rhodes College, Memphis, TN, USA
Corresponding author: Mark Risjord; Email: mrisjor@emory.edu

(Received 22 January 2025; revised 15 June 2025; accepted 24 July 2025; first published online 28 August 2025)

Abstract

Thick concepts have both descriptive and evaluative dimensions to their meaning. Some have argued that because the descriptive and evaluative dimensions cannot be separated (they are "blended"), the implicit values influence the confirmation of any "mixed claims" containing the thick concept. Using the development of the concept of hypersegregation as a case study, we argue for a distinction between the semantic function of definitions and the epistemic function of indicators. While thick concepts are semantically blended, the evaluative meaning of a thick concept does not influence the function of epistemic indicators. Therefore, mixed claims can and should be tested impartially.

1. Introduction

Consider two men in Chicago. Juan, who is Hispanic, lives in Pilsen on the Lower West Side. Zion, who is African American, lives in Chatham on the South Side. Despite living in neighborhoods in which they are members of their respective ethnic and racial majorities and despite having similar education, incomes, and occupations, Juan and Zion are likely to experience Chicago in significantly different ways. Zion must traverse far more neighborhoods and far greater distances before he encounters a white neighborhood. The neighborhoods he travels are likely to be similar to his own: predominantly Black, with more food deserts, higher infant mortality, higher incarceration rates, fewer hospitals, and fewer educational and employment opportunities than city-wide averages across the Chicago metro area. In short, the very structure of Chicago's neighborhoods imposes greater burdens on Zion than on Juan for many basic amenities.

The differences between Juan and Zion reflect the sociological phenomenon known as *Black hypersegregation* (Denton, 2007; Massey and Denton, 1989, 1993; Massey and Tannen, 2015; Wilkes and Iceland, 2004). To call Chicago hypersegregated both describes and evaluates the environments within which Juan and Zion live; it is a paradigmatic *thick* concept (Abend, 2019; Alexandrova, 2017; Alexandrova, 2018; Alexandrova and Fabian, 2022; Anderson, 2002; Djordjevic and Herfeld, 2021; Dupré,

© The Author(s), 2025. Published by Cambridge University Press on behalf of Philosophy of Science Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

2007; Risjord, 2007; Sundstrom, 2004; Van der Weele, 2021). Philosophers of science broadly agree on two points about thick concepts. First, thick concepts cannot be cleanly divided into descriptive and evaluative components. They are semantically "blended." Second, attempting to purge science of thick concepts would be a mistake. Without thick concepts, scientific inquiry would be unable to identify the phenomena that are important to us. These two points entail that nonepistemic values[1] are ineliminable from science. In an earlier phase of the discussion, most held that the presence of thick concepts did not influence the confirmation of scientific hypotheses. That is, a scientific inquiry could be nonneutral and still be impartial (Lacey, 1999; Anderson, 2004; Risjord, 2007).

Starting with Alexandrova (2018), however, views of the relationship between neutrality and impartiality have begun to change. Elliott expresses the contemporary view succinctly:

> "[S]ome fields of science employ concepts that incorporate both empirical and normative components, and so good scientific reasoning requires that scientists appeal at least in part to non-epistemic values when assessing hypotheses in those fields of science" Elliott (2022, 33).

This presentation of the argument is obviously elliptical, but the blended character of thick concepts fills the gap. The intuition is that because the descriptive and evaluative components of thick concepts cannot be separated, any empirical testing of a hypothesis containing a thick concept—a "mixed claim"—would depend on the implicit values. When this *blending intuition* is made explicit, the argument is straightforward:

(1) Thick concepts' descriptive and evaluative components are blended.
(2) **Blending intuition:** If a concept *C* is blended, then appeal to the values implicit in *C* must be part of the justification of any mixed claim in which *C* occurs.
∴ The values implicit in a thick concept must be part of the justification of any mixed claim in which that concept occurs.

While the blending intuition is quite plausible, it generates a puzzle. On the one hand, thick concepts are needed to ensure that science addresses questions of social concern. On the other hand, if the blending intuition is right, empirical confirmation of mixed propositions requires commitment to the values implicit in the thick concepts. It follows that different evaluative stances would yield different bodies of scientific results; unfavorable empirical results would always be just more "fake news." The puzzle arising from the blending intuition is that thick concepts enable science to address socially relevant questions while at the same time undermining its ability to provide satisfying and useful answers.

In this article, we will argue that the blending intuition is false. Using "hypersegregation" as a case study of scientific concept development, we will argue

[1] Because the only interesting thick concepts in this context imply moral or political values, the distinction between epistemic and nonepistemic values is not in play. Hence, we will drop the "nonepistemic" qualifier in what follows.

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

that the blending intuition conflates the semantics of conceptual content with the epistemology of identifying instances of a concept and testing mixed claims. Once these have been distinguished, we argue that nonneutrality is consistent with impartiality and thereby recover the sense in which scientific inquiry can provide impartial evidence for policy.[2]

2. Semantic blending and the development of "hypersegregation"

Many have argued that the descriptive and evaluative dimensions of a thick concept cannot be fully disentangled (Anderson, 2002; Gallie, 1955; McDowell, 1981; Root, 2007; Taylor, 1967). In an often-cited definition of thick concepts, Elizabeth Anderson makes blending explicit:

> "A concept is thickly evaluative if (a) its application is guided by empirical facts; (b) it licenses normative inferences; and (c) interests and values guide the extension of the concept (that is, what unifies items falling under the concept is the relation they bear to some common or analogous interests or values)" (Anderson, 2002, 504–5).

Anderson's final clause expresses the way in which thick concepts are blended. It also appears to support the blending intuition. After all, if values guide the extension of a concept, C, and assessing the truth of a proposition containing C requires attending to C's extension, then the values would seem to guide assessment. The blending intuition thus turns on what it means for values to "guide the extension of the concept," but this meaning is far from clear.

To make the character of blending clear, let us turn to the development of the concept of hypersegregation. Douglas Massey and Nancy Denton's intentional creation of a new thick concept clearly illustrates the sense in which the determination of the extension of the new concept was "guided." It also shows that there is an important difference between the *semantic* project of defining a new concept and the *epistemic* project of developing *indicators* that measure or identify objects in the concept's extension. In section 3, we will argue that the deployment of such indicators does not depend on the values implicit in the thick concept, and therefore that confirming (or disconfirming) mixed propositions does *not* require attending to such values.

The development of "hypersegregation" followed a pattern that is common in the social sciences. The phenomenon of scientific concern, segregation, is already recognized and described in ordinary language. Scientists initially frame their questions in these common terms, but to answer them, they need to develop specialized methods. This prompts a methodological question: Do such-and-such methods really identify *segregation*? In the social sciences, this is recognized as a question of "measurement validity" or "operationalization." Robert Adcock and David Collier (2001) synthesized a commonly used analytic framework—originally articulated by Paul Lazarsfeld (1958)—for discussing issues of measurement validity.

[2] While he does not explicitly invoke the blending intuition, our conclusions also run contrary to Ronald Sundstrom's analysis of the values implicit in segregation studies in (Sundstrom, 2004).

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

They distinguish between a *background concept*, a *systematized concept*, and an *indicator*. Because background concepts are drawn from nonscientific discourse, they typically bear a number of different, sometimes conflicting meanings, and they are often thick. Systematized concepts carve out determinate aspects of the background concept's meaning, but they do not attempt to capture its full range of implications. Indeed, a single background concept may spawn several distinct systematized concepts. Because the point of developing systematized concepts is to make scientific questions answerable in a more consistent and reliable way, systematized concepts are typically developed along with measurement techniques. Adcock and Collier call these *indicators*, which include "any systematic scoring procedure, ranging from simple measures to complex aggregated indices not only quantitative indicators but also the classification procedures employed in qualitative research" (Adcock and Collier, 2001, 530). The historical development of the systematized concept of "hypersegregation" and its associated measurement indices out of the background concept of "segregation" nicely fits Adcock and Collier's framework. In so doing, it clearly shows how empirical and evaluative considerations worked together to create the new, systematized concept of hypersegregation.

In the first decades of the 20th century, studies of segregation were content to rely on the concept of segregation as it was used in ordinary discourse. However, the ambiguities of the concept were always apparent. In one of the earliest multicity studies of residential segregation, Thomas J. Woofter represented patterns of segregation with maps (Woofter, 1928). Because of the ambiguities, Woofter felt compelled to produce *two* maps of each city. One map blocked out areas where the Black population was in the majority, whereas the other placed dots, each representing 100 Black residents. Later scholars would distinguish these maps as showing the *evenness* and *concentration* in the distribution of Black residents. But Woofter had no systematized concepts capturing these different meanings of "segregation."

By the late 1930s, mathematical techniques for measuring variation in data sets began to emerge (Wright, 1937). Segregation researchers recognized the value of such segregation indices, but they disagreed among themselves on the best mathematical representation. As a result, researchers were unclear about whether different indices were two ways of representing the same thing or whether they were capturing different senses of segregation. Otis Duncan and Beverly Duncan showed that a number of the indices could "be regarded as functions of a single geometrical construct, the 'segregation curve'" (Duncan and Duncan, 1955, 210). As a consequence, they argued, while the indices were not equivalent, the dissimilarity index (eq. 1 later in the article) contained most of the information carried by the other indices. Duncan and Duncan's work led to the wide adoption of the dissimilarity index in segregation studies. However, their achievement did not succeed in resolving the ambiguity of the background concept.

Although the dissimilarity index dominated segregation studies for 20 years, in the 1970s, new indices were proposed. Some of these aimed to capture aspects of segregation not represented by the dissimilarity index. In the late 1980s, Massey and Denton analyzed 20 different indices in order to disambiguate "segregation" and create more clearly defined, specified concepts. Using a factor analysis, they found that the 20 indices loaded onto five factors. Using conceptual arguments to

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

complement these mathematical results, they identified the five factors with five recognizable patterns of residential segregation:

> Minority members may be distributed so that they are overrepresented in some areas and underrepresented in others, varying on the characteristic of *evenness*. They may be distributed so that their *exposure* to majority members is limited by virtue of rarely sharing a neighborhood with them. They may be spatially *concentrated* within a very small area, occupying less physical space than majority members. They may be spatially *centralized*, congregating around the urban core, and occupying a more central location than the majority. Finally, areas of minority settlement may be tightly *clustered* to form one large contiguous enclave, or be scattered widely around the urban area. (Massey and Denton, 1988, 283)

In Adcock and Collier's (2001) terms, each of the five dimensions is a systematized concept of segregation. Massey and Denton associated each of the five dimensions of segregation with a preferred segregation index. Each segregation index provided a way of measuring a particular kind of spatial distribution. For instance, they argued that evenness was best measured using the dissimilarity index:

$$D = 0.5 \sum_{i-1}^{n} \left| \frac{x_i}{X} - \frac{y_i}{Y} \right| \tag{1}$$

Here, x_i is the proportion of a minority group in census tract i,[3] X is the proportion of that same minority group in the entire metropolitan area in which i is located, y_i is the proportion of non-Hispanic whites in census tract i, and Y is the proportion of non-Hispanic whites in the entire metropolitan area in which i is located. Intuitively, the dissimilarity index measures the proportion of a minority group that would need to move from neighborhoods in which that group is overrepresented to neighborhoods where that group is underrepresented in order for each neighborhood to have the same demographic composition as the broader metropolitan area. Similar mathematical representations were used to characterize preferred indices for the other four dimensions.

In the language of Adcock and Collier (2001), the segregation indices are *epistemic indicators* for the specified concepts of segregation. Their function is to measure the distribution of residents in a metropolitan area and assign it a number between 0 and 1. That number is then evidence for inferring that the distribution in a metropolitan area is, for example, more or less even. In general, Massey and Denton preferred indices that loaded more heavily and discretely on one specific dimension of segregation, that were easy to compute, and that were consistent with past studies. It is important to note that while Massey and Denton chose preferred indices for each dimension of segregation, they did *not* argue that each dimension had a unique indicator—a point to which we will return.

[3] Census tracts are relatively permanent statistical subdivisions of a county or statistically equivalent entity. They have 1,200 to 8,000 residents and are typically taken to be operational definitions of neighborhoods.

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

The systematization of the concept of segregation into five specified concepts was based on both epistemic and evaluative considerations. On the epistemic side, the factor analysis suggests that there are five distinct patterns of segregation. The fact that each of the 20 indices loaded onto one of the five factors was evidence that segregation, too, had five forms. However, this would not have been sufficient grounds for systematizing segregation into a concept with five dimensions. After all, they could have decided to exclude one or more of the factors from the extension of "segregation." Their decision to include five dimensions also rested on conceptual arguments that made explicit appeal to values associated with segregation. For example, when they distinguished concentration from centralization, they invoked explicitly normative considerations:

> A fourth dimension of segregation is related to concentration, but is conceptually distinct. Centralization is the degree to which a group is spatially located near the center of an urban area. In most industrialized countries, racial and ethnic minorities concentrate in center city areas, inhabiting the oldest and most substandard housing. (Massey and Denton, 1988, 291)

Centralization is presumably a relevant form of racial segregation because the central city was an undesirable place to live. Massey and Denton's creation of a systematized concept of segregation thus used both evaluative and empirical considerations to establish its extension. In so doing, they created a concept that was unambiguous in its definition and therefore more scientifically useful than the ambiguous background concept. However, they also preserved the thickness of the original.

In a separate article, Massey and Denton (1989) proposed a novel systematized concept to be used in segregation research: hypersegregation. Hypersegregation occurs when a minority group is highly segregated on four of the five dimensions of segregation. A group is considered to be highly segregated along a given dimension if the dimension's associated index has a value of at least 0.6. The justification for defining hypersegregation as a high level of segregation on most dimensions is that the dimensions intersect to create harms over and above the harms associated with each of the separate forms of segregation (Massey and Denton, 1989, 373). Massey and Denton express the normative consequences of the concept of hypersegregation when they argue that because Black Americans are hypersegregated, they "occupy a unique and distinctly disadvantaged position in U.S. urban society" (Massey and Denton, 1989, 374).

The development of systematized concepts out of the ambiguous background concept of segregation shows what it means for evaluative considerations to "guide the extension of" a thick concept (Anderson's third condition). As the background concept was systematized, evaluative commitments about the evils of segregation were relevant to deciding whether a particular residential distribution was or was not to be counted as a specific type of segregation. The definitions of the different dimensions of segregation and the resulting concept of hypersegregation were thus motivated by evaluative commitments. While values are not explicitly mentioned in the definitions, Massey and Denton's (1989) text shows that the dimensions of both "segregation" and "hypersegregation" license normative inferences, satisfying Anderson's second condition. The systematization of "segregation" was therefore

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

not an attempt to develop nonevaluative technical terms for scientific use (contra Nagel, 1961). Systematized concepts blend the descriptive and evaluative aspects of their meaning. However, this does not yet confirm the blending intuition: It remains to be shown that blending entails that the implicit values must play a role in the empirical justification of mixed propositions.

3. Epistemic indicators and mixed propositions

This section will argue that even though systematized concepts are blended, the blending intuition is mistaken because it conflates the semantic definition of a concept with the epistemic indicators providing evidence for mixed propositions. Once the distinction between definitions and indicators is appreciated, it becomes clear that mixed propositions can be empirically supported without appeal to the values implicit in thick concepts.[4]

Let us begin with definitions. The function of a scientific definition is to clearly delineate a concept's extension. Paradigmatically, it does so by providing necessary and sufficient conditions for either inclusion of instances within the extension (e.g., *x* is a triangle if and only if *x* is a closed, three-sided figure) or for the determinants of its value range (e.g., the temperature of *x* is the mean kinetic energy of *x*). By providing clear conditions for inclusion within and exclusion from the concept's extension, scientific definitions provide truth conditions for propositions containing the concepts. In this sense, then, scientific definitions have a *semantic* function.

When a thick concept like hypersegregation is defined, the *extension* does not contain values. The extension of hypersegregation is a set of metropolitan areas—precisely, those scoring 0.6 or higher on four of the five dimensions of segregation. Therefore, the values said to be "implicit" in thick concepts must be a component of their semantic content over and above their extension as specified by the technical definition. It follows that the technical definitions of thick, specified concepts capture only *part* of their semantic content. Insofar as thick concepts license evaluative inferences, there is a further dimension of their meaning over and above their extension.

While a scientific definition of a concept delimits its extension, it cannot, by itself, tell us whether an actual object falls within that extension. *Knowing* whether a given metropolitan area is hypersegregated requires evidence. Therefore, something beyond the definition must fulfill the *epistemic function* of providing such evidence and thereby help to test empirical propositions that include the concept. When the definiens are all easily observable, the indicator is simply *the observation* of the definiens. In most scientifically interesting cases, however, evidence for whether a definition is satisfied by an object or state of affairs cannot be read off the definiens. As the history of "hypersegregation" illustrates, contemporary science devotes substantial resources to the development of specialized measurement instruments and procedures. These scientific artifacts are specifically designed to provide evidence that a particular scientific definition (or class of definitions) is satisfied.

[4] This section will thus flesh out Anderson's first condition: that the *application* of a thick concept be "guided by empirical facts."

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

Indicators are those scientific artifacts that have this epistemic function of providing evidence.

Systematized concepts and indicators thus have distinct functions. The explicit definitions of systematized concepts disambiguate the background concept. Scientific definitions are mutually exclusive; a single concept does not have more than one definition at the same time in the same theoretical context. Eliminating such vagueness and ambiguity is the point of replacing a background concept with one (or more) clearly defined systematized concept(s). The function of indicators, by contrast, is epistemic: They provide evidence that an object or state of affairs satisfies the concept, given its definition. The definition functions semantically to provide truth conditions for propositions in which the concept occurs. Together, the indicator and definition let scientists infer that propositions deploying the concept are true of some actual situations and false of others.

Segregation indices are epistemic indicators, and as such, they are distinct from definitions. One might object, at this point, that segregation indices, as illustrated in equation 1, are propositional. Hence, they look like definitions. However, unlike scientific definitions, segregation indices are not mutually exclusive. As noted earlier, Massey and Denton did *not* associate each specified concept of segregation with a unique indicator. Their pluralism recognizes that there are multiple ways of epistemically identifying whether a metropolitan area is, for example, even in its distribution. But the pluralism does not introduce a semantic ambiguity. Segregation indices are therefore not definitions.

The blending intuition holds that the values implicit in a thick concept must figure in the confirmation of mixed propositions in which it occurs. We have seen so far that while systematized concepts can remain thick, their definitions delimit unambiguous extensions, and the indicators provide evidence for mixed propositions. If the blending intuition is correct, the use of epistemic indicators must somehow depend on the values implicit in a mixed proposition. There are three points where the implicit values might play a role in the use of segregation indexes to confirm mixed claims about hypersegregation: (1) the choice of index, (2) the use of a measurement procedure to supply numeric values (simple measurement outcomes) to the index's variables, and (3) the calculation of the index's value from the simple measurement outcomes. Once the index values have been calculated, all of the evidence relevant to the concept of hypersegregation is in, and the definition of hypersegregation will let us determine whether the target city is hypersegregated.

To determine whether values need to intervene at any of these points, let us turn to an example of testing a mixed claim. Rima Wilkes and John Iceland (2004) used data from the 2000 census to ask whether any of four minority groups—Blacks, Hispanics, Asians, or Native Americans—were hypersegregated in major American cities.

Let us begin with the choice of index. Wilkes and Iceland calculated the values of 19 of Massey and Denton's 20 indices using the 2000 census data. Calculating multiple indices for each of the five dimensions let them triangulate and identify any results that might be artifacts of the index. In the article's final analysis, for "ease of interpretation" (Wilkes and Iceland, 2004, 26), they settled on one index for each dimension. They remark that the choice was based on an "assessment of the indices, Massey and Denton's recommendations, and earlier research" (Wilkes and Iceland, 2004, 26). Hence, pragmatic values are playing a role. However—and crucially for the

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

blending intuition—these pragmatic considerations are not drawn from the values implicit in "hypersegregation."

While Wilkes and Iceland did not, in fact, rely on the values implicit in hypersegregation in their choice of indices, it is also the case that they *ought not* to have done so. Massey and Denton used empirical and pragmatic considerations to determine the best indicators, although they noted that in some cases, particular features of the target system, such as very small minority populations, might necessitate alternative choices. The indices are thus selected for their epistemic characteristics. If these choices were also good in light of the values associated with hypersegregation, then the values would be both harmless and irrelevant. Conversely, if values associated with hypersegregation contradicted these choices and motivated indicators that were less reliable, then the values would clearly be hindering the inquiry. Therefore, the values implicit in a thick concept ought not influence the choice of indicators used to test mixed propositions.

Second, consider the collection of simple measurement outcomes. Wilkes and Iceland used the 2000 census data. Now, while it is possible—even likely—that the collection of this data was value laden in various ways, those values were not consequences of the thickness of "hypersegregation" because the concept played no role in the 2000 census methodology. Wilkes and Iceland generated simple measurement outcomes of the following form:

> In year y, n% of geographic region R's residents were members of racial/ethnic group G.

There is no reason to believe that the values associated with hypersegregation specifically influenced these simple measurement outcomes. And as before, if the empirical results conformed to the values, the values would be irrelevant, whereas if the values motivated a modification of the outcomes, they would hinder the inquiry by undercounting or overcounting different racial groups. Therefore, values ought not play a role in deriving simple measurement outcomes.

Finally, complex measurement outcomes are calculated from the indices using the simple measurement outcomes as inputs. The simple measurement outcomes give values to the variables on the right side of an index like the dissimilarity index (eq. 1), and the math determines the value of D. The calculation of these complex measurement outcomes is thus straightforwardly mathematical. Clearly, no values are necessary for addition, subtraction, and so on. Should the values implicit in hypersegregation be invoked, they would again either be irrelevant or hinder the attainment of reliable results.

At this point in the process, Wilkes and Iceland had all the evidence necessary to answer the question of whether Black, Hispanic, Asian, and Native American minority populations were hypersegregated in American cities in 2000. They concluded that "twenty-nine metropolitan areas could be classified as having black-white hypersegregation in 2000" (Wilkes and Iceland, 2004, 29), and six of these scored above 0.6 on all *five* dimensions. By contrast, Native Americans and Asians were not hypersegregated in any cities, and Hispanics were hypersegregated in two. These are clearly mixed claims involving the thick concept of hypersegregation. Wilkes and Iceland thus provided clear, direct evidence for mixed claims, and they did so without the values implicit in

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

"hypersegregation" playing any role. We conclude, then, that the blending intuition is false. Moreover, if the values had influenced the justification of the mixed claim so as to change the empirically derived results, such influence would have hindered the inquiry. The confirmation of mixed claims therefore *ought* to be value-free.

4. Conclusion

Our argument against the blending intuition points the way out of the puzzle with which we began. By defining systematized concepts, scientists fix parts of the background concept's extension and develop epistemic indicators. With these tools, they can frame inquiries relevant to our political debates, thus confirming the idea that thick concepts make science relevant to social concerns. Of course, because the systematized concept may remain thick, it is always possible for someone with contrary politics to reject the relevance of the inquiry to their concerns. This is a motivation to initiate *other* lines of inquiry with different systematized concepts. Complex political issues thus appropriately spawn a variety of inquiries. However, because the results of these inquiries can (and ought to) be impartially supported, adherents of one political position are not entitled to reject the answers generated by another inquiry. Because mixed propositions can and should be impartially tested, the sciences can do justice to the richness of social problems while providing solutions that do not treat evidence and ideology as interchangeable.

Acknowledgments. Dr. Risjord's work on this article was supported by the Czech Science Foundation (GAR), GX20-05180X, "Inferentialism Naturalized: Norms, Meanings and Reasons in the Natural World."

References

Abend, Gabriel. 2019. "Thick Concepts and Sociological Research." *Sociological Theory* 37 (3):209–33. https://doi.org/10.1177/0735275119869979

Adcock, Robert, and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95 (3):529–46. https://doi.org/10.1017/S0003055401003100

Alexandrova, Anna. 2017. *A Philosophy for the Science of Well-Being*. Oxford: Oxford University Press. https://doi.org/10.1111/phc3.12203

Alexandrova, Anna. 2018. "Can the Science of Well-Being Be Objective?" *British Journal for the Philosophy of Science* 69 (2): 421–45. https://doi.org/10.1093/bjps/axw027

Alexandrova, Anna, and Mark Fabian 2022. "Democratising Measurement: Or Why Thick Concepts Call for Coproduction." *European Journal for Philosophy of Science* 12 (1):1–23. https://doi.org/10.1007/s13194-021-00437-7

Anderson, Elizabeth. 2002. "Situated Knowledge and the Interplay of Value Judgments and Evidence in Scientific Inquiry." In *In the Scope of Logic, Methodology and Philosophy of Science*, vol. 2 of 11th International Conference of Logic, Methodology, and Philosophy of Science, Cracow, edited by Peter Gärdenfors, Jan Wolenski, and K. Kijania-Placek, 497–517. Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-94-017-0475-5_8

Anderson, Elizabeth. 2004. "Uses of Value Judgments in Science: A General Argument, with Lessons from a Case Study of Feminist Research on Divorce." *Hypatia* 19 (1):1–24. https://doi.org/10.1111/j.1527-2001.2004.tb01266.x

Denton, Nancy A. 2007. "Hypersegregation." In *The Blackwell Encyclopedia of Sociology*, edited by George E. A. Ritzer. New York: John Wiley and Sons. https://doi.org/10.1002/9781405165518.wbeosh057

Djordjevic, Charles, and Catherine Herfeld. 2021. "Thick Concepts in Economics: The Case of Becker and Murphy's Theory of Rational Addiction." *Philosophy of the Social Sciences* 51 (4):371–99. https://doi.org/10.1177/00483931211008541

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press

Duncan, Otis Dudley, and Beverly Duncan. 1955. "A Methodological Analysis of Segregation Indexes." *American Sociological Review* 20 (2):210–17. https://doi.org/10.2307/2088328

Dupré, John. 2007. "Fact and Value." In *Value-Free Science?: Ideals and Illusions*, edited by Harold Kincaid, John Dupré, and Alison Wylie, 42–57. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195308969.003.0003

Elliott, Kevin. 2022. *Values in Science*. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009052597

Gallie, Walter Bryce. 1955. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56 (1):167–98. https://doi.org/10.1093/aristotelian/56.1.167

Lacey, Hugh. 1999. *Is Science Value Free?: Values and Scientific Understanding*. New York: Routledge. https://doi.org/10.4324/9780203983195

Lazarsfeld, Paul. 1958. "Evidence and Inference in Social Research." *Daedalus* 87 (4):99–130.

Massey, Douglas S., and Nancy A. Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces* 67 (2):281–315.

Massey, Douglas S., and Nancy A. Denton. 1989. "Hypersegregation in US Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions." *Demography* 26 (3):373–91. https://doi.org/10.2307/2061599

Massey, Douglas S., and Nancy A. Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

Massey, Douglas S., and Jonathan Tannen. 2015. "A Research Note on Trends in Black Hypersegregation." *Demography* 52 (3):1025–34. https://doi.org/10.1007/s13524-015-0381-6

McDowell, John. 1981. "Non-Cognitivism and Rule-Following." In *Wittgenstein: To Follow a Rule*, edited by Steven H. Holtzman and Christopher M. Leich, 141–62. London: Routledge & Kegan Paul.

Nagel, Ernest. 1961. *The Structure of Science*. London: Routledge & Kegan Paul.

Risjord, Mark. 2007. "Scientific Change as Political Action: Franz Boas and the Anthropology of Race." *Philosophy of the Social Sciences* 37 (1):24–45. https://doi.org/10.1177/0048393106296541

Root, Michael. 2007. "Social Problems." In *Value-Free Science?: Ideals and Illusions*, edited by Harold Kincaid, John Dupré, and Alison Wylie, 27–41. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195308969.003.0004

Sundstrom, Ronald R. 2004. "Racial Politics in Residential Segregation Studies." *Philosophy & Geography* 7 (1):61–78. https://doi.org/10.1080/1090377042000196029

Taylor, Charles. 1967. "Neutrality in Political Science." In *Philosophy, Politics and Society*, vol. 3, edited by Peter Laslett, 25–75. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139173490.003

Van der Weele, Simon. 2021. "Thick Concepts in Social Research: What, Why, and How?" *International Journal of Qualitative Methods* 20:1–11. https://doi.org/10.1177/16094069211066165

Wilkes, Rima, and John Iceland. 2004. "Hypersegregation in the Twenty-First Century." *Demography* 41 (1):23–36. https://doi.org/10.1353/dem.2004.0009

Woofter, Thomas J. 1928. *Negro Problems in Cities*. Beloit, KS: McGrath.

Wright, John K. 1937. "Some Measures of Distributions." *Annals of the Association of American Geographers* 27 (4):177–211. https://doi.org/10.1080/00045603709357166

Cite this article: Risjord, Mark, Kareem Khalifa, and Jared Millson. 2025. "Thick Concepts and Impartiality." *Philosophy of Science* 92 (5):1117–1127. https://doi.org/10.1017/psa.2025.10114

https://doi.org/10.1017/psa.2025.10114 Published online by Cambridge University Press