Thought coordination as concept identity

Keywords: concepts, coordination, coreference, relationism, Campbell's regress

William James (1890/2016, p. 459) observed that "the same matters can be thought of in successive portions of the mental stream, and some of these portions can know that they mean the same matters which the other portions meant." He claimed that this "sense of sameness is the very keel and backbone of our thinking" and argued convincingly that the "subjective sense of identity" between the objects of our thoughts is not explained by the objective identity between the "outer things", if any, those thoughts are about. James' "sense of sameness" anticipates the contemporary notion of coordination between thoughts (Fine 2008; Gray 2017; Goodman & Gray 2020, 2024; Murez 2023). "Coordination" (roughly) designates the relation between representations which the thinker is immediately disposed to treat as coreferential, without the mediation of any further representation of identity or coreference.

What, if not objective coreference, explains thoughts being coordinated or representing-as-the-same? A simple answer is that thought coordination is grounded in concept identity. According to this coordination-as-concept-identity view (henceforth, "CCI"), we represent things as the same by redeploying the same concept to think about them. Thus, we may conflate different things by representing them using the same concept (Landrum 2022), or conversely, represent one thing using different concepts, not realizing we are thinking of the same thing again (Frege 1892).

The main competitor of CCI is *relationism* (Fine 2008; Heck 2012, 2014; Goodman & Gray 2020, 2024; for review, see Gray 2017). Relationists deny that coordination is grounded in identity, equivalence, or any *internal* relation between representations, i.e., any relation determined by those representations' intrinsic properties. In particular, relationists deny that coordination between a pair of representations is explained by each independently having the same (or a similar) content or vehicle as the other. Coordination is claimed to be "irreducibly relational" and grounded in informational processes which connect token representations of different semantic and vehicular types (Goodman & Gray 2020, 2024).

CCI can be fleshed out in different ways, by appealing to different notions of "concept". My goal will be to defend against relationism versions of CCI that construe concepts as robust representational vehicles. In doing so, I will draw on empirical considerations and focus on issues relevant not only to philosophers, but also to cognitive scientists.

First, I will clarify the explanandum. A theory-neutral way of understanding "coordination" is as shorthand for a complex of personal-level inferential/behavioral dispositions. I distinguish normative epistemological questions concerning the rational standing of those

dispositions from descriptive psychological questions about the nature and (broadly) causal explanation of coordination. Among the latter, I further distinguish the question of how coordination dispositions are acquired diachronically from the question of which functional organization or mechanisms underlie them synchronically. I propose that CCI is best construed as sketching an answer to this last sort of question.

I then underscore a relationist insight: it cannot generally be assumed that the capacity to represent some relation R between the referents $o_1...o_n$ of some representational vehicles $V_1...V_n$ is grounded in the fact that R obtains between $V_1...V_n$. This assumption is illegitimate even when R is identity: representing-as-the-same need not be grounded in sameness between representations. Yet the proper lesson to draw, I argue, is not that CCI is false, but only that its truth cannot be established a priori. The debate between CCI and relationism thus becomes an open empirical debate about how coordination actually works in human cognition.

We should seek accounts that secure "explanatory distance" (Taylor 2023) between their explanandum and the explanantia they offer. This speaks in favor of accounts which appeal to *robust* mental entities, i.e., entities (including processes/relations) whose characteristics and identity can, at least in principle, be ascertained through multiple means, rather than only based on whether thinkers superficially exhibit coordination dispositions. I argue that extant versions of relationism, and many versions of CCI, fail to satisfy this desideratum.

By contrast, versions of CCI that construe concepts as (re)identifiable by signature properties of their psychological and/or neurological vehicles make it possible to independently ascertain, through empirical means, when the same or different concepts are being deployed. I present two arguments – one theoretical, the other more empirical – in favor of such *robust* versions of CCI.

The theoretical argument amends influential regress arguments from Campbell (1988) and Millikan (1993). Supposing a plausible computational-representational theory of thought, there must be *some* relation between token thought-representations that does not itself need to be represented explicitly as the content of a dedicated representational vehicle. Instead, that relation must be encoded implicitly, in the sense that cognitive/computational processes are causally sensitive to its being instantiated by the vehicles themselves. Because *identity* between vehicles is guaranteed to be exploitable in the relevant manner by cognitive/computational processes, this makes it likely that such a relation should be exploited by well-designed cognitive systems to encode the ecologically crucial relation of identity between outer things.

The empirical argument develops the idea that relationism and CCI, if interpreted as making claims about robust mental entities, can be associated with competing empirical predictions. Unlike CCI, relationism predicts no syntactic (algorithmic level) or implementational (neural level) identity or resemblance between the vehicles of coordinated thoughts. For instance, relationism does not predict that personal-level coordination dispositions should reliably correlate with similarity in the patterns of neuronal activation that underlie separate deployments of concepts. I argue that such correlations in fact exist. As an illustration, I appeal to recent neuroscientific research on the processing of pronominal anaphora, which is a paradigmatic way in which coordination between (singular) thoughts is linguistically expressed. Dijksterhuis et al. (2024) show that pronouns reactivate the same neuronal representations (known as "concept cells") as the antecedent nominals with which they are interpreted as coreferential. This is what CCI predicts, if concept cell reactivation is taken as a neurological signature of (singular) concept redeployment (Quiroga 2012). Beyond this particular empirical hypothesis, a more general lesson emerges: robust versions of CCI have powerful heuristic value as guiding hypotheses for multi-level empirical investigation of the "sense of sameness".

References

Campbell, J. (1988). Is sense transparent? *Proceedings of the Aristotelian Society* 88:273-292.

Dijksterhuis D.E. *et al*. (2024) Pronouns reactivate conceptual representations in human hippocampal neurons. *Science* 385,1478-1484.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100(1), 25-50.

Goodman, R. & Gray, A. (2022). Mental filing. *Noûs* 56 (1):204-226.

Goodman, R. & Gray, A. (2024). Mental filing, continued. *Synthese* 204 (1):1-26.

Gray, A. (2017). Relational approaches to Frege's puzzle. *Philosophy Compass* 12 (10):e12429.

Heck, R. (2012). Solving Frege's puzzle. *Journal of Philosophy* 109 (1-2):132-174.

Heck, R. (2014). In Defense of Formal Relationism. *Thought: A Journal of Philosophy* 3 (3):243-250.

James, W. (1890). *The Principles of Psychology*. London, England: Dover Publications.

Landrum, K. (2023). Coordination, Content, and Conflation. *Australasian Journal of Philosophy* 101 (3):638-652.

Millikan, R. G. (1993). On mentalese orthography. In Bo Dahlbom, Dennett and His Critics. Cambridge, Mass., USA: Wiley-Blackwell.

Murez, M. (2023). The transparency of mental vehicles. *Noûs* 58 (4):877-904.

Quiroga, R. Q. (2012). Concept cells: the building blocks of declarative memory functions. *Nature Reviews Neuroscience*, *13*(8), 587-597.

Taylor, E. (2023). Explanatory Distance. *British Journal for the Philosophy of Science* 74 (1):221-239.