

Hormone Research as an Exemplar of Underdetermination*

P.D. Magnus

e-mail: [pmagnus\(at\)fecundity.com](mailto:pmagnus@fecundity.com)
web: <http://www.fecundity.com/job>

Published in *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(3): 559-567. September 2005.

<http://www.elsevier.com>

This is my final draft and may not reflect copyediting done by the staff of the journal.

Abstract

Debates about the underdetermination of theory by data often turn on specific examples. Many cases are invoked often enough that they become familiar, even well-worn. Here I consider one such commonplace: the connection between prenatal hormone levels and gender-linked childhood behavior. Since Helen Longino's original discussion of this case a decade-and-a-half ago, it has become one of the stock examples of underdetermination. However, the case is not genuinely underdetermined. We can easily imagine a possible experiment to decide the question. The fact that we would not perform this experiment is a moral, rather than epistemic, point. Further, I argue that Longino need not have appealed to 'underdetermination' to establish her central claim about the case.

1 Underdetermination

Arguments about the underdetermination of theory by data often depend on examples. Philosophers list theory choices which are allegedly underdetermined to show that underdetermination is serious and widespread. Consider, as illustration, a recent exchange between Ron Giere and Janet Kourany.

*This paper is based in part on work supported under a National Science Foundation Graduate Fellowship, work which I presented at the University of Alabama, Birmingham in Feb 2001 and in [Mag03, §4.3.3]. My thinking on these issues developed in response to feedback from Philip Kitcher, Helen Longino, and some number of anonymous referees.

Giere notes, “it is often argued that the underdetermination of theory by data provides space for the social determination of theory” [Gie03, p. 19]. He invokes some of the “many arguments that the extent of such underdetermination is not nearly so great as Kourany assumes” [Gie03, p. 19]. Giere cites Laudan and Leplin [LL91], and in her reply Kourany cites Kukla [Kuk98]—yet these are *en passant* references to general arguments by those other philosophers. Kourany suggests instead that the matter can be decided by examples. “Abstract arguments aside,” she writes, “the scientific and philosophical literatures contain many examples of underdetermination in ‘real scientific practice’” [Kou03, p. 23]. She goes on to mention a number of cases, including research on the connection between IQ and race [Gly98] and between gender and hormones [Lon90]. Giere and Kourany share this concern for examples. She notes, “Giere himself... has pointed to one of the applications in biology— Longino’s critique of the linear-hormonal research program... Longino illustrates the kind of choice to which feminist principles lead” [Kou03, p. 24].

Of course, there is a *prima facie* gap between examples and general claims about science. What justifies the inductive generalization from *these* cases?¹ Nevertheless, as Earman writes, “the production of a few concrete examples is enough to generate the worry that only a lack of imagination on our part prevents us from seeing comparable examples of underdetermination all over the map” [Ear93, p. 31]. So it is important to consider examples directly, lest they be uncritically traded around as commonplaces. In the remainder of the paper, I want to consider the specific example from Longino [Lon90] [Lon02, pp. 126–7, 183, 199–200]: the rôle of prenatal hormone levels in determining childhood behavior. This case is not genuinely underdetermined, as we can see by considering an experiment that would decide the matter. In the final section, I argue that underdetermination was not essential for Longino’s purposes anyway.

2 Hormones and gender-linked behavior

A news item heralding British research on the source of boyish behavior in young girls announced, “Forget nurture, the tendency for a girl to behave like a tomboy is all up to nature— specifically, the amount of testosterone a baby is exposed to during pregnancy.” Quoted in the story, a researcher says, “Because hormones influence basic processes of brain development, they also exert permanent influences on behaviour” [ABC02].² This research and the uncritical reporting of it presuppose that chemicals effect behavior by effecting brains. Differentiation of behavior between the sexes is taken to be the result of differing brain chemistry, and this in turn is presumed to result from differing levels of hormones. Whatever rôle environment plays in the differentiation of behavior is (according to the model) independent of hormone levels and brain development.

¹Some authors have argued that inductive fallacies are responsible for prominent but degenerate debates in philosophy of science; cf. [MC04], [How00, p. 52–4], and [Lew01].

²The study touted is Hines, et al. [HGR⁺02]. Claims about brains, which the researcher makes to the media, are not made explicitly in the published study.

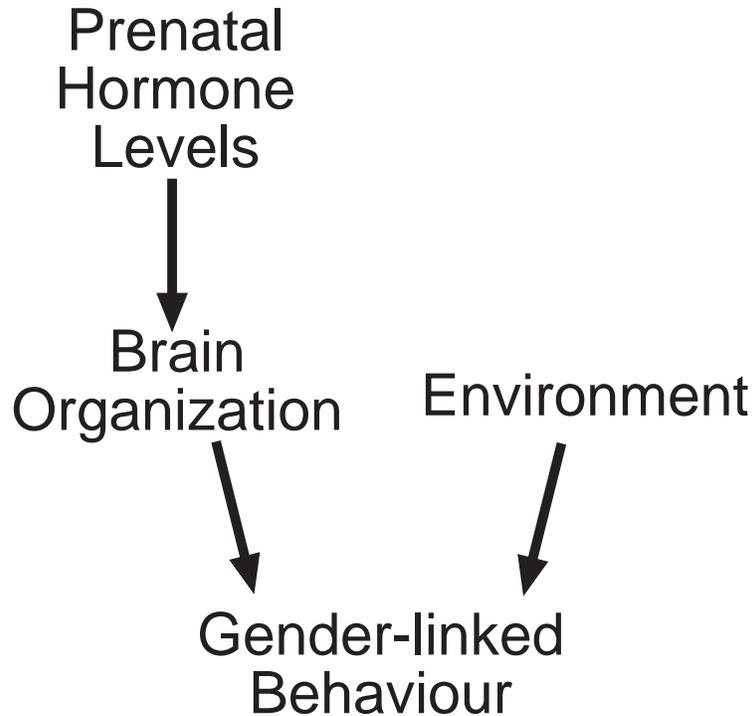


Figure 1: The LH model.

Longino calls this the Linear-Hormonal (LH) model. It posits a causal structure like the one in figure 1.³

Two sorts of evidence are used in support of the LH model: animal and human studies. In the animal studies, hormone levels are manipulated directly and behavior is then observed. For example, testosterone correlates with fighting behavior in rats. In human studies, groups with hormone disorders might be observed and compared to control groups. For instance, girls with abnormally high levels of prenatal androgens are observed to exhibit more boy-like behavior than girls with normal levels.

Animals, Longino argues, are sufficiently unlike humans that the animal studies should not be assumed to describe how the hormones being studied affect

³This is Longino's figure 2 [Lon90, p. 138]. The arrows should not be taken to imply immediate causation without intermediate causes, but only to imply that the elements represented in the graph do not, according to the model, causally influence one another except where there are arrows. The graphs I employ here can be understood as directed, acyclic graphs in the manner of Spirtes, Glymour, and Scheines [SGS93]. Glymour uses the same methodology in work lauded by Kourany [Gly98, cited above].

human behavior. Even monkeys, who lack many of the complications present in humans, have been shown to react differently to hormones than rats. [Lon90, p. 157] I find Longino's argument convincing, but that undercuts claims of underdetermination and feminism. She argues that the animal evidence should not be used to underwrite conclusions about humans, but she does so without appealing in any obvious way to a rival conception of science. She does not argue from the premise that she is a feminist to the conclusion that humans are different than rats, but rather she argues against the analogy by pointing to disanalogies. Would the disanalogies be unconvincing to anyone working within the LH model? Wouldn't LH partisans be unreasonable if they were unconvinced?

The human studies, Longino believes, only stand as evidence for a particular hormonal connection if one begins with the LH model. Longino suggests as an alternative a selectionist model developed to explain human memory, learning, and self-awareness. On this model, "experience... and self-image... play a primary role in the *biological* explanation of the behavior-action of species with a highly developed cortex" [Lon90, p. 148].

Although this model "is not in use to explain any particular category of behavior" [Lon90, p. 143], Longino suggests that it may provide an alternate explanation for the results of the human studies. First, the girls with abnormally high levels of prenatal androgens were aware that they had such a condition. This medical history might make them feel self-conscious, feel unlike other girls, and feel uncertain of their femininity. These factors could result in behavior seen as more boy-like. Second, testosterone is known to affect muscle development. More rambunctious behavior by the girls might simply have been an expression of the consequent need for exercise. [Lon90, p. 150] Research employing the LH model cannot be criticized as bad science, Longino thinks, because intra-scientific considerations alone give no reason to prefer the selectionist model. She writes:

Both rest on explanatory models that involve metaphysical assumptions about causality and human action. Neither theoretical perspective can muster constitutively based arguments sufficient to exclude the other— thus both can continue to generate studies that are used to support microhypotheses about the etiology of particular forms of behavior that are consistent with one or the other broader model. [Lon90, p. 161]

Rather than deciding based on strictly scientific concerns, Longino thinks we should prefer the selectionist model because it accords with political ideals of autonomy and responsibility. Whereas the LH model connects behavior to a simple biological variable, the selectionist model emphasizes "the enabling rather than the limiting aspects of biology" [Lon90, p. 175-6].

Longino laments that "work on cognition shows no sign of reflection or analysis on the part of researchers. Correlation after correlation is produced with no attempt to understand just what it is that is being measured or its relation to associated phenomena" [Lon90, p. 167]. Indeed, she is right to insist that an

observed correlation alone does not establish any particular causal connection. Nevertheless, a correlation is something to be explained— Longino admits as much by providing a rival explanation.⁴ She accepts, just as adherents of the LH model do, that the general causal structure is such that prenatal hormone levels are at one end of a causal chain that has gender-linked behavior at the other end. Longino offers only schematic and complicated depictions of the selectionist model, but she offers specific causal hypotheses:

As for the . . . children who provide the strongest evidential base for the hormonal model, one can say that their behavior is a specific response to their situation as they perceive it. They are . . . persons who (1) have a medical history productive of greater self-consciousness and self-knowledge than is usual for young people, (2) are quite aware of their uniqueness, and (3) are very likely uncertain of their “femininity” and possibly of their femaleness as well. This combination of circumstances may account for their choosing as children behavior less restrictive in its implications for adult life than traditional girllike behavior is. Finally, such direct effects of testosterone as do exist could be on muscle development, with a consequent need for exercise. [Lon90, p. 150]

These are the specific hypotheses summarized above and in figure 2.

Longino makes much of the causal loops in the selectionist model. For instance, a child’s psychological state both effects and is effected by her actions.⁵ Yet there is no possibility of reciprocal causation between prenatal hormone levels and gender-linked behavior, since the former plays its rôle years before the latter occurs. If there is some causal connection, it must flow one way and not the other. The correlation might result from some other causal structure— e.g., if both were effects of a common cause— but, provided one is a causal ancestor of the other, it must be hormone levels causing childhood behavior and not *vice-versa*. The only question is, What is the mechanism? Longino suggests that hormone levels effect the girls’ personal history in ways that, in turn, effect their self image. The girls then make different choices than other girls. Choices are causes of actions, of course, so this means that self image is taken to be one causal factor controlling behavior. There is an additional causal pathway from hormones to behavior, one that goes through muscle development. It is this causal scenario— and not the selectionist model *tout court*— that is represented in figure 2.

Is Longino right to think that the choice between the causal account represented in figure 1 and the account represented in figure 2 is underdetermined on the basis of evidence and that this determination could only be made on the basis of contextual values? It is true that the mere correlation between hormone levels and gender-linked behavior does not favor one over the other, but there are

⁴Some correlations are spurious and need no special explanation, of course, but none of the parties suggest that here.

⁵Cf. her figures 6 and 7. [Lon90, pp. 147, 149]

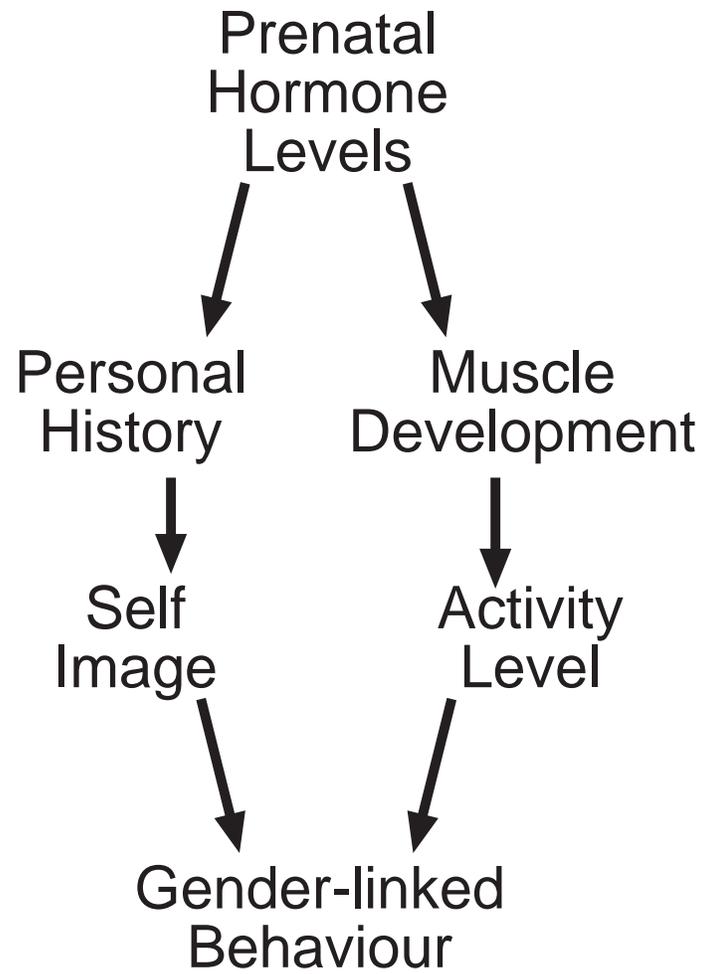


Figure 2: Longino's selectionist model.

testable differences nonetheless. Notice that on the selectionist account (figure 2), hormones only influence behavior through intermediate causes of personal history and muscle development. Two children with identical histories and muscle development should thus exhibit the same degree of gender-linked behavior *regardless* of their prenatal hormone levels. Contrast this with the LH model (figure 1), in which there is a causal path from hormones to behavior which goes only through brain organization, such that gender-linked behavior should be strongly correlated with hormone levels even among children in identical environments with identical muscle development.

Imagine an experiment, then, in which one each out of pairs of identical twins is treated with hormones in utero. The twins are never told which was which. They are not even told that they are part of a medical study. As they grow up, careful records are kept of their muscle development, using some objective measure like muscle mass. When they reach school age, case workers judge whether the childrens' behavior is boyish or girlish.

The question here would not be simply whether or not prenatal hormone levels correlate with behavior, but whether they do so conditional upon history and muscle development. If a correlation remains conditional on these other factors, then it could not be explained by a causal structure like figure 2; some other causal path would have to exist between hormone levels and behavior. Conversely, if no significant correlation remains after conditioning on history and muscle development, then a causal structure like figure 1 could be ruled out. Neither result would prove what the causal structure actually is *in an absolute sense*, but either would eliminate one of the two rivals. Controlling directly for personal history and conditioning on muscle development would eliminate the correlation between those variables and prenatal hormone levels. If a correlation remains between prenatal hormone levels, then there must be some other causal path between the two. Some variable X may intervene—perhaps brain organization.

There may be concerns about this experimental design. There are familiar worries about twin studies, for instance that there are never enough natural twins and that there are sample selection biases. So imagine a more extreme version of the experiment in which twins are forcibly harvested and raised in a clinical facility to fully control for history. Some variation on this double-blind structure should yield robust results. The alleged underdetermination yields no significant conclusion, since further evidence can close the inferential gap. This is just the ordinary situation of enquiry. Cost might make such a study impractical, but more importantly moral considerations would make it reprehensible. What this shows, though, is that there are things we are not willing to do in the name of science. Empirical considerations *could* decide between the two theories, if we were willing to pay the price for the knowledge.⁶

Of course, these empirical considerations do not settle the matter deductively. There are still assumptions that, as Longino says, “establish the evi-

⁶It's not clear that such an extreme study is necessary. The study mentioned above [HGR⁺02] involved healthy children rather than children with hormone disorders. One of the factors that Longino appeals to (peculiar medical history) is absent.

dential relevance of data to hypotheses” [Lon02, p. 126]. Yet the assumptions are not morally charged assumptions about humanity and agency. It is unclear how the assumptions involved in a suitably designed experiment would have a controversial sociopragmatic dimension.

3 Postscript

Longino argues that we should pursue research programmes within the selectionist model and abandon present work within the LH model. I am sympathetic with this position, but I don’t see how it requires the choice between the models to be underdetermined. Let me suggest a way we might reach her conclusion, but without relying on claims of underdetermination.

The experiment proposed in the previous section presumes that only one of the two models is true. Even if we suppose that hormones have *some* influence on brain development, it is certain that such influence is not the only thing going on— otherwise the correlations would be perfect. So it seems that environment is also playing a rôle. Suppose that the LH model and the selectionist model both told part of the causal story. Prenatal hormones would influence behavior in a number of ways: by the mechanisms posited in the LH model, by those posited in the selectionist model, and perhaps by other ways as well. If that were so, the choice would not be underdetermined, because there would not need to be any choice at all. We would be justified in accepting elements of both models.⁷

Yet this would not settle Longino’s question: How should we structure research programmes that study the effects of hormones? As Longino notes, scientific research is selective. It explores some features of the world and not others. As such, we should pick programmes not only on the basis of whether their posits conform to the world but also on the basis of whether or not they explore features of the world that we care about.

Longino gives us good reason to think that the features picked out by the LH model aren’t ones we should care about, *even if* those features do represent part of the whole causal story. As such, our concerns should lead us to prefer the selectionist model.

References

- [ABC02] Tomboys made, not born— by testosterone. ABC Science Online [⟨http://www.abc.net.au/science/news/stories/s725131.htm⟩](http://www.abc.net.au/science/news/stories/s725131.htm), November 2002. Australian Broadcasting Corporation. Accessed 14xi2002.

⁷Even if the LH model accurately describes one causal path, the researcher cited at the beginning of §2 is still careless to infer from the strength of the correlation to the strength of one specific causal path.

- [Ear93] John Earman. Underdetermination, realism, and reason. In *Midwest Studies in Philosophy*, volume XVIII, pages 19–38. University of Notre Dame Press, 1993.
- [Gie03] Ronald N. Giere. A new program for philosophy of science. *Philosophy of Science*, 90(1):15–21, January 2003.
- [Gly98] Clark Glymour. What went wrong? Reflections on science by observation and *The Bell Curve*. *Philosophy of Science*, 65(1):1–32, March 1998.
- [HGR⁺02] Melissa Hines, Susan Golombok, John Rust, Katie J. Johnston, Jean Golding, and the Avon Longitudinal Study of Parents and Children Study Team. Testosterone during pregnancy and gender role behavior of preschool children: A longitudinal, population study. *Child Development*, 73(6):1678–1687, November/December 2002.
- [How00] Colin Howson. *Hume’s Problem: Induction and the Justification of Belief*. Clarendon Press, Oxford, 2000.
- [Kou03] Janet A. Kourany. Reply to Giere. *Philosophy of Science*, 90(1):22–26, January 2003.
- [Kuk98] André Kukla. *Studies in Scientific Realism*. Oxford University Press, 1998.
- [Lew01] Peter Lewis. Why the pessimistic induction is a fallacy. *Synthese*, 129:371–380, 2001.
- [LL91] Larry Laudan and Jarrett Leplin. Empirical equivalence and underdetermination. *The Journal of Philosophy*, 88(9):449–72, 1991.
- [Lon90] Helen Longino. *Science as Social Knowledge*. Princeton University Press, 1990.
- [Lon02] Helen Longino. *The Fate of Knowledge*. Princeton University Press, 2002.
- [Mag03] P.D. Magnus. *Underdetermination and the Claims of Science*. PhD thesis, University of California, San Diego, 2003.
- [MC04] P.D. Magnus and Craig Callender. Realist ennui and the base rate fallacy. Forthcoming in *Philosophy of Science*, 2004.
- [SGS93] Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search*. Springer-Verlag, New York, 1993.