**Note to readers**. This is a chapter from my forthcoming book, *Causation with a Human Face: Normative Theory and Descriptive Psychology*. (OUP). As the title suggests, the book is about the interplay between normative ideas about causal reasoning and empirical results due to psychologists and others about how people in fact reason and judge. The chapter accordingly has some discussion of empirical results concerning whether people judge in accord with proportionality type considerations in their causal judgments as well as some suggestions about possible experiments. But much of the chapter is more normative and analytical, exploring various ways of formulating proportionality requirements and their rationales. I’m posting it in advance of the publication of *CHF* because there seems to be considerable current interest in the topic of proportionality. The published version of this chapter may differ slightly from this version so please cite the former if possible.

**Chapter 8: Proportionality**

**8.1 Introduction**[[1]](#footnote-1)

This chapter is about the role of proportionality in causal judgment. This notion was introduced by Yablo in his (1992) and has been the topic of a considerable amount of subsequent discussion. I see proportionality as one criterion among several having to do with the choice of variables for causal analysis[[2]](#footnote-2). The notion is applicable when (and as I will formulate it below, only when) we face a choice among different candidates for the cause variable for a fixed effect or explanandum and where the candidate cause variables are themselves related in a specific way: they stand in what might loosely be described as a non-causal[[3]](#footnote-3) hierarchy of abstractness or more or less fine-grainedness with respect to one another. (The determinate/determinable relation invoked by Yablo --see below-- is one way in which variables or properties in such a hierarchy may be related but it may not be the only way.) Illustrations are provided by two extensively discussed cases to which the notion of proportionality has been applied: a variable representing the presence or absence of a particular shade of red such as scarlet stands lower in such a hierarchical relation to a more generic, less specific variable representing the presence or absence of the color red which in turn is more specific (less abstract ) than an even higher-level variable that just records whether some object is some color or other (rather than transparent). Similarly (one might think) if, for a variable whose values represent mental states, the same mental state can be multiply realized by different neural states, the mental states might be viewed as abstractions from (or coarse-grainings of) these neural states.

The background to this is the observation that both in ordinary life and in science, causal claims can be formulated at different “levels” or “grains”. It is widely (and in my view) correctly believed that some choices of “level” (or of variables associated with certain levels) are better than others for purposes of causal analysis. Moreover, although there is some tendency in philosophical discussion to suppose otherwise, there are strong reasons for thinking that the best choice of level is not always the most specific or fine-grained one[[4]](#footnote-4). Equally it is not always the most abstract or general level that is better or more appropriate. Moreover, as an empirical matter, people do not behave as they believe that either of these choices is always best. Instead, people seem to believe that which choice or choices is best or better depends on the empirical details of the behavior of the system they are trying to understand—an assessment that is supported by normative analysis. This is also reflected in the fact that in science such choices are regarded as non-trivial and as requiring careful thought. Proportionality can be understood as one possible consideration that provides guidance for a good or appropriate choice of level. As such, it addresses a problem which is ubiquitous in causal thinking, both lay and scientific. Thus, as with the other considerations influencing variable choice discussed in this book, proportionality has both a normative and empirical side: there is the question of what normative rationale, if any, supports the use of proportionality (and how to formulate a proportionality condition that has a suitable normative rationale) and there is also the empirical question of the extent to which people conform to such a condition.

Let me add that in what follows I will largely but not entirely focus on the role of proportionality in the assessment of type-level causal claims. Although I will not try to systematically explore this issue, my guess is that, as an empirical matter, proportionality considerations play a somewhat different role in the assessment of actual cause or token level claims than in the assessment of type-causal claims[[5]](#footnote-5). In particular, failures to fully satisfy proportionality are (as an empirical matter) sometimes (and perhaps often) regarded as more acceptable for actual cause claims than for type level claims. An appendix to this chapter describes a possible normative rationale for this difference.

**8.2 Yablo on Proportionality**

In one of Yablo’s illustrations[[6]](#footnote-6) a pigeon is trained to peck at targets of any shade of red and only such targets. The pigeon is presented with scarlet targets and pecks. Consider the following two claims (understood, for the present, as type-level claims):

(8.2.1) The red color of the target causes the pigeon to peck.

(8.2.2) The scarlet color of the target causes the pigeon to peck.

Yablo claims (8.2.1) is superior or “preferable” to (8.2.2) on the grounds that (8.2.2) is “overly specific” and fails to be “proportional” to its effect. We will consider Yablo’s more precise characterization of proportionality as well as some alternative characterizations shortly but his intuitive idea is that causes are proportional to their effects when they do not “contain too little” (where “too little” means being inappropriately narrow and omitting crucial elements) *and* do not contain “too much” (where this means being overly broad and containing irrelevant or superfluous elements). The cause cited in (8.2.2) is overly narrow or too fine-grained (since the pigeon will peck in response to non-scarlet shades of red). (8.2.1) does not have this defect, given the facts stipulated in the example and thus satisfies proportionality. By contrast if we were to describe the cause as the target’s having some color or other (as opposed to being transparent) it would be overly broad.

Yablo’s more precise characterization of proportionality is as follows: First, proportionality is understood as applying to properties that are related as determinates (e.g., scarlet) and determinables (e.g., red), where for our purposes we can think of determinate properties as more specific ways of realizing the corresponding determinable properties or as standing to those properties in a subordinate to superordinate relation. In general the presence of a determinate property necessitates the presence of the corresponding determinable property but not conversely. Yablo’s view is that cause *C* is proportional to effect *E* if and only if *C* is “required” for *E* and *C* is “enough” for *E*. *C* is required for *E* iff none of its determinables screens it off and *C* is enough for *E* iff it screens off all of its determinates. In this context *C1* screens *C2* off from *E* iff, had *C1* occurred without *C2*, *E* would still have occurred (Yablo, 1997)[[7]](#footnote-7). Applied to the pigeon example, the idea is that the presentation of the red target is proportional to the effect of pecking because all of the determinables of red (such as the target’s being some color or other) fail to screen off red from pecking (the counterfactual “if the target had not been red but had been colored, the pigeon would have pecked” is false). Moreover, for any determinate of red such as the target’s being scarlet if that determinate had not obtained but the target was still red, pecking would have obtained. The presentation of a scarlet target is not proportional to the effect (pecking) because there is a determinable of scarlet (namely red) such that if that determinable had obtained, the pigeon would have pecked[[8]](#footnote-8).

Although Yablo does *not* take satisfaction of a proportionality condition to be a strictly necessary condition for the truth of causal claims (1992, p. 277), he does suggest that (8.2.1) and (8.2.2) “compete” with respect to truth and that since (as he supposes) (8.2.1) is true, (8.2.2) must be false. (In other words, the preferability of (8.2.1) to (8.2.2) is understood in terms of their differing truth values.) A number of other writers either interpret Yablo as claiming that satisfaction of some proportionality condition is necessary for a causal claim to be true (e.g., Shapiro and Sober, 2012—see below) or advocate this position themselves (e.g., List and Menzies, 2009).

These claims raise a number of interesting descriptive and normative issues. First, although Yablo’s proposal is normative (he thinks we *ought* to prefer proportional to non-proportional causal claims), it also seems natural to interpret him as suggesting [[9]](#footnote-9) that, as an empirical matter, subjects exhibit a preference of some kind for causal judgments that satisfy a proportionality requirement (on some suitable characterization of that requirement.) Is this descriptive claim correct? If so, precisely what characterization of a proportionality requirement best captures people’s judgments-- Yablo’s or some alternative? Relatedly, if subjects do (as an empirical matter) exhibit a preference for causal claims that satisfy some version of proportionality, what is the nature of this preference? Do subjects think that causal claims that fail to satisfy proportionality are (for that reason) always false? Or do they think that such claims can sometimes be true (assuming other appropriate requirements on causation are satisfied) but are deficient (or comparatively deficient) in some other way—e.g., misleading, unsatisfactory for explanatory purposes, or less informative than they might be. Will a treatment of proportionality that, as a descriptive matter, best captures people’s judgments take the form of a dichotomous, all- or- nothing condition (as Yablo’s condition does) so that a causal claim either satisfies this condition or not, with failure to satisfy the condition being associated with falsity? Or should we think of proportionality as a more graded requirement/condition, that can be satisfied to a greater or lesser degree? Note that this last possibility seems to fit better with the notion that comparative failures of proportionality are associated with failures along some other dimension besides truth/falsity since the latter dimension is “binary”.

To the best of my knowledge there has been relatively little discussion of these empirical issues in the psychological literature. The published research that most directly bears on these issues is Lien and Cheng, 2002, which is discussed below[[10]](#footnote-10). These authors find that their subjects do judge and make causal inferences in a way that suggests that they prefer causal claims that satisfy a proportionality-like condition. However, they do not directly address the issue of whether non-proportional (or less proportional claims) are judged to be false. (They also don’t use the word “proportionality” or cite the philosophical literature on this notion.) The issue of whether subjects judge that claims like (8.2.2) are false of course might be addressed just by asking them a yes/no question, although there are more sophisticated possibilities. It would also be interesting to investigate more systematically whether, as I expect, subjects would give higher causal strength ratings (for some appropriate verbal probe) to claims like (8.2.1) in comparison with claims like (8.2.2), given the empirical facts stipulated in the example. There is also the issue, raised briefly above, of whether proportionality is weighted differently in connection with type-level judgments in comparison with singular or actual cause judgments and if so, why this is the case.

In addition to these empirical question, appeals to proportionality raise a number of important normative issues. Here a comparison with invariance is instructive. Although many accounts of causation pay little attention to invariance-based considerations, few researchers have explicitly argued against the idea that invariance matters normatively in causal assessment. By contrast, a number of philosophers have argued that proportionality is a misguided or normatively indefensible requirement. (For various forms of such arguments, see Shapiro and Sober, 2012, Bontly, 2005, McDonnell, 2017.) Thus the question of whether there is a normative rationale for some version of a proportionality condition and if so, what that rationale is a matter of considerable disagreement.

Obviously whether one thinks that a proportionality condition is normatively defensible is going to depend on just what one takes that condition to involve. Much of the criticism directed against the condition is directed against versions that claim that satisfaction of proportionality is necessary for a causal claim to be true (e.g., Sober and Shapiro, 2012) . Of course, as suggested above, one can agree that proportionality is not necessary for truth while still thinking that, other things being equal, causal claims that do better in terms of proportionality are in some way preferable (because they are more informative, provide better explanations or conduce more to other goals we have) and thus that in this sense proportionality is a normative desideratum. This is my own view, as explained below.

Writers who deny that proportionality is necessary for the truth of causal claims sometimes conclude that proportionality is a “merely pragmatic” virtue or perhaps a condition relevant to the assessment of *causal explanations* (with explanatoriness itself understood as a pragmatic virtue) but not to causation itself. For example, Bontly (2005, 332) writes “ ... proportionality is not in fact a constraint on the causal relation but rather a pragmatic feature of our use of causal language, derived from general principle of language use. If so, it turns out that proportionality has little if anything to do with the nature of causation itself”. In a related vein McLaughlin (2007, p 15) writes: “[f]or the record, I myself think that rather than a constraint on causation, proportionality is a pragmatic constraint on explanation[[11]](#footnote-11)”

Here the reader is urged to recall the remarks on “pragmatic” considerations in Chapter 1. If “pragmatic” is associated with means/ends justification, then of course I agree that to the extent that a proportionality condition has a normative justification, this will take the form of arguments showing that satisfaction of this condition conduces to various ends or goals associated with causal thinking. This is the same form of argument that (according to me) can be used to provide a normative justification for other features like invariance that are relevant to the assessment of causal claims. In this respect (conduciveness to goals associated with causal thinking) I agree that proportionality is a pragmatic virtue. Note, however, that this notion of pragmatic does not support Bontly’s claim that proportionality “derives from general features of language use ” (presumably features having to do with the “pragmatics” of language such as those captured by Gricean maxims for conversational relevance). My own view, defended below, instead sees proportionality as justified in terms of considerations and goals that are distinctively associated with causal thinking rather than with language use in general. For similar reasons, I would resist the invidious contrast (between causation and causal explanation) implied by the claim that proportionality has nothing to do with the “nature” of causation, although perhaps it is connected to causal explanation. In part this is a matter of my misgivings, expressed in previous chapters, about the uses to which philosophers have put the supposed contrast between what does and does not belong to the nature of causation. Putting aside this general concern, my view (recall chapter 1) is that proportionality has to do with what I have called distinctions among causal relationships, in the sense that some true causal claims (or causal relationships) will satisfy proportionality conditions to some considerable extent and other true causal claims will not. Thus if the nature of causation has to do with necessary conditions for a causal claim to be true, then I agree that proportionality is not such a necessary condition. However, on my view, it is also true that a concern with proportionality is closely bound up with goals and functions associated with causal thinking rather than being, as it were, an external add-on to these.

In the remark quoted above McLaughlin explicitly contrasts “causation” with “causal explanation”, with the suggestion that “pragmatic” concerns play a large role in the latter . Of course there are obvious differences that are denied by no one—causation is a relation in the world while “causal explanations” (as well as causal claims, whether or not these are used to explain) involve devices – words, graphs, equations etc.-- that are used to *represent* causal relationships. It is unclear, however, that it follows from this that when causal language is used to provide causal explanations, “cause” in such explanations somehow behaves very differently from (is governed by different rules than) “cause” when it is used in other contexts (“plain vanilla cause”/ “cause with no implication of explanation”?). At the very least such a view requires considerably more spelling out than it has hitherto received. As an empirical matter, it seems to me, as suggested in Chapter 7, that it is at best unclear whether ordinary subjects distinguish between “cause” and “causally explains” in the manner suggested[[12]](#footnote-12). I will add that from a normative perspective, if one thinks of causal explanation along the lines that I favor, with this understood as providing information that can be used to answer what-if-things-had-been different questions, then it is natural to think that one explains causally simply by providing causal information that can be used to answer such questions. Of course one can provide causal information that fails to answer such questions[[13]](#footnote-13) or fails to answer them very well, but on this picture there is no sharp contrast between causal claims and causal explanations, at least of the sort some philosophers claim exists[[14]](#footnote-14). On the basis of these considerations I think we should resist the suggestion that proportionality is a condition that applies only to causal explanation but not to causal claims more generally[[15]](#footnote-15).

Of course this is not to say that proportionality has nothing to do with explanatory goodness—on the contrary. Rather my point is that this connection is not captured very well by adoption of a framework that distinguishes sharply between causal claims and causal explanations in the way Bontley and McLaughlin do. Thus when I say that when a causal claim is defective along the dimension of proportionality the deficiency in question should be understood to pertain to the causal claim itself and should be understood in terms of goals associated with causal thinking rather than in terms of a separate category of causal explanation, governed by different rules and/or embedded within the framework of a pragmatic theory of explanation.

**8.3. A Closer Look at Proportionality**

With this as background, let me now turn to a more detailed exploration of what proportionality involves or how it might be understood. To do this, let’s return to (8.2.1)—(8.2.2) .To represent (8.2.1) -- (8.2.2) within an interventionist framework we need to express them in terms of claims about variables. Neither (8.2.1) nor (8.2.2) are explicit about how we are to think about the variables figuring in them. However, focusing first on (8.2.1), it is very natural to interpret it as employing a binary cause variable *RED*  that can take either of two values, {red, not-red}, and a binary effect variable *PECK* capable of taking the values {peck, not –peck}. (8.2.1) might then be represented as:

(8.3.1) *RED* causes *PECK*

where this is unpacked as implying that

(8.3.1a) An intervention that sets *RED= red* is followed by *PECK=peck*

(8.3.1b) An intervention that sets *RED= not red* is followed by *PECK= not peck*

Thus we are interpreting (8.2.1) as implying both that an intervention that sets the target’s being red leads to pecking and that an intervention that sets the target to a non-red color leads to the pigeon not pecking. Put more simply (and incorporating the idea that causal claims are intrinsically contrastive, which has a natural interventionist motivation[[16]](#footnote-16)): we are interpreting (8.2.1) as the claim that the target’s being red rather than not red causes the pigeon to peck rather than not peck. When interpreted in this way (8.2.1) is true, given the facts specified in the example.

Reasoning in a parallel way, we can interpret (8.2.2) in terms of a cause variable *SCARLET* that takes the values {scarlet, not scarlet} as well as the *PECKS* variable:

(8.3.2a) An intervention that sets *SCARLET= scarlet* is followed by *PECK=peck*

(8.3.2b) An intervention that sets *SCARLET= not scarlet* is followed by *PECK= not peck*

(8.3.2a) is true. How about (8.3.2b)? Here we face a choice point. If we require that for (8.3.2b) to be true, *all* interventions that set *SCARLET= not scarlet* must be followed by *PECK= not peck,*  then (8.3.2b) is false (since interventions that set the color of the target to non-scarlet shades of red are followed by pecking). Thus on this understanding of what it implies (8.2.2) is false.

On the other hand, for (8.3.2b) to be true we might require only that *some* interventions that set *SCARLET= not scarlet*  be followed by non-pecking. Since there are some ways of setting the target to a non-scarlet color (e.g., blue) that are followed by non- pecking, on this interpretation (8.3.2b) and hence (8.2.2 ) are true[[17]](#footnote-17).

To the best of my knowledge, there have been no systematic empirical investigations of whether most subjects interpret (8.3.2b) in terms of the requirement that all interventions that set *SCARLET= not scarlet* must be followed by *PECK= not peck,* or in instead in terms of the requirement that some interventions that set *SCARLET= not scarlet* are followed by *PECK= not peck*  and accordingly whether (8.2.2) is generally judged to be true (although perhaps less than fully perspicuous in some way) or to be false[[18]](#footnote-18).

Of course if (8.2.1) is true and (8.2.2) is false (when these are interpreted along standard interventionist lines), there is no puzzle about why (8.2.1) is preferable to (8. 2.2)— apparently we don’t need to bring proportionality considerations into the picture to explain this preference. Suppose, on the other hand, that (8.2.2) is regarded as true (because, for example, it is interpreted in terms of the requirement that some interventions that set *SCARLET= not scarlet* are followedby *PECK= not peck.)* In this case, we cannot appeal to the falsity of (8.2.2) to explain why (8.2.1) is preferable to (8.2.2) but (I contend) it still seems plausible that (8.2.1) is preferable to (8.2.2). One way of motivating this assessment is to note that within an interventionist framework, there are at least two different ways in which a causal claim might be deficient (or at least limited in various ways or lacking some desirable feature). At this point I will describe these roughly and imprecisely, just to provide a motivating intuition, with a more careful statement (**P**) provided below:

(8.3.3a) A causal claim might falsely claim that some (intervention-supporting) dependency relationship is present when it is not. Call this *falsity*. Violations of this requirement are ruled out by interventionist requirements like (**M**).

(8.3.3b) A causal claim might fail to represent one or more dependency relations that are present in the system of interest and that should be represented. Call this *omission.*

I assume that not all failures to represent dependencies along the lines of (8.3.3b) involve problematic deficiencies or limitations but that some such failures do—failures to represent what *should* be represented. (This is explained in more detail below.) To anticipate, I see failures of proportionality as having to do with certain sorts of failures of this sort.

To apply (8.3.3a- b) to (8.2.1)—(8.2.2), when (8.2.2) is interpreted in accord with the “all interventions” interpretation, it falsely claims that a dependency relationship is present when it is not—hence violating *falsity* (8.3.3a). When ( 8.2.2) is interpreted in terms of the “some interventions” interpretation, it does not violate *falsity* but, as I will interpret *omission* (8.3.3b), it does violate this condition in virtue of failing to represent or convey the information that there are other shades of red besides scarlet that lead to pecking. Again this failure seems to be a defect or limitation of some kind. By contrast (8.2.1) respects both *falsity* and *omission.* This gives us reason to prefer (8.2.1) to (8.2.2) even if (8.2.2) is regarded as true. Here we are supposing that even if true a causal claim can be defective (or less good than some alternative true claim) because of information it fails to provide.

If we confine attention just to a comparison of (8.2.1) and (8.2.2), it may seem simpler to explain our preference for (8.2.1) in terms of (8.2.1) being true and (8.2.2) being false. In other cases, however, this option is not available. Suppose, following Franklin-Hall (2016), that the causal facts involving the pigeon are as described earlier. Consider a new binary variable *C*, which has two values, scarlet and cyan, with the pigeon pecking when the target is scarlet but not when it is cyan. Now consider

(8.3.4) *C* causes *PECKS*

where this is understood as implying that

An intervention that sets *C=scarlet* is followed by *P= pecking*

An intervention that sets *C=cyan* is followed by *P= not pecking*

Putting aside misgivings about whether the variable  *C* is it itself objectionable in some way[[19]](#footnote-19), it seems plausible that (8.3.4) is true—or at least we can’t judge it to be false on the grounds that we judged (8.2.2) (under the “all interventions” interpretation) to be false. (8.3.4) does not falsely claim that dependency relations exist when they do not. On the other hand, there does seem to a deficiency in (8.3.4) along the dimension captured by *omission.* In comparison with (8.2.1), (8.3.4) fails to represent many facts about the dependency relations present in the example: (8.3.4) does not tell us that non-scarlet but red targets lead to pecking and that non-red non-cyan targets are followed by non-pecking. I take this (as well as other considerations described below) to suggest that we need something along the lines of *omission*  to make the kinds of discriminations among causal claims that we want (and need) to make. A properly formulated proportionality condition should accomplish this.

**4. Proportionality Formulated**

Turning now to the task of providing such a formulation, let me re-emphasize the following: we may think of all of the examples discussed above (8.2.1, 8.2.2, 8.3.4) as having to do with *variable choice—*in particular, choices concerning variables representing causes. Some choices of variables seem better (more appropriate, perspicuous, informative etc.) than others in formulating causal relationships, given the empirical facts holding for the system that we want to characterize, and it is this idea that we want to capture. For example, *RED* seems a better choice for the cause variable in the pigeon example than *SCARLET* or *C*. I take proportionality to have to do with a principle governing variable choice which might be characterized as follows[[20]](#footnote-20):

(**P**= Proportionality) Suppose we are considering several different causal claims/explanations formulated in terms of different candidate cause variables *V1..Vn* that are members of a set **V.**  Each of these can be used to represent different claims about patterns of dependency relations involving some target effect or explanandum *E,* which is fixed or pre-specified*.* The variables *Vi* stand in non-causal hierarchal relations (e.g., realization) to one another. Thus we are choosing among pairs one of which is a candidate cause variable and the other of which is an associated dependency relation linking that cause variable to the specified effect, with each such pair at a different level in the hierarchy. For example, we are to choose between the pair {*RED* and (8.2.1)} versus {*SCARLET* and (8.2.2)}. Then a choice of variable *Vi* (and of the dependency claims regarding *E* in which *Vi* figures) satisfies proportionality better than an alternative choice from **V** to the extent that those dependency claims satisfy *Falsity* and *Omission* above—that is, to the extent that (i) non-existent dependency relations involving *E* are not falsely represented as present (as noted earlier this can be understood in terms of satisfaction of **M**) and to the extent that (ii) existing dependency relations (from among the variables in **V**) involving *E* are represented. When we have specified a cause variable and associated dependency relations delineating the conditions under which all possible values of *E* occur, we have fully satisfied **P**.

Obviously (**P**) requires additional explication. Applying it requires that we first fix or pre-specify the effect variable (e. g., *PECKS*) and then choose among different candidate variables for characterizing the dependency relations governing that effect. The motivation for this requirement is that without such a prior specification of the effect variable, the problem of choosing among different cause variables becomes completely indeterminate and unconstrained[[21]](#footnote-21). In other words, what we are interested is choosing among representations of dependency relations (and associated candidate cause variables) governing pecking and not in the representation of dependency relations governing other possible effects or *explananda*—e.g., whether the pigeon coos or blinks. Moreover we are interested just in the conditions under which pecking vs non-pecking occurs and not in, say, fine-grained variations in pecking rate.

Second, in applying (**P**) we are interested just in the problem of choosing among cause variables that bear certain non-causal hierarchical relationships to one another. This includes variables that are related as *SCARLET* is to *RED—*that is*,* via supervenience, coarse –graining, realization, and determinate/ determinable relationships, among others. (Again **P** is *not* intended to apply to cases in which we are choosing among variables that do not stand in such relations. See below for discussion) When variables are related in such ways, there will be pairs of their values (pairs involving one value for each variable) that, for non-causal reasons, cannot hold together, for the same individual or unit. “Non- causal reasons” is meant to capture constraints that exclude possibilities holding for logical, conceptual, supervenience-based or similar relations but not for causal reasons. When this is the case, I will say that this combination of values is not *compossible*. For example, the variable *SCARLET* cannot take the value = *scarlet* for a particular target while the *RED* variable takes the value = *not red* for that same target. This pair of values is not compossible although not because the target’s being scarlet *causally* excludes its being not red[[22]](#footnote-22).

Contrast such failures of compossibility with the following: Suppose (following Franklin-Hall, 2016) that whether the pigeon will peck or not depends (causally) not just on the color of the target but also on whether or not its chin is tickled, represented by a variable *TICKLES*. Each value of *TICKLES* *is* compossible with each value of *RED*—the same pigeon can be tickled and presented with a red target, not tickled and presented with a red target and so on. Similarly, all possible values of *RED* are compossible with the possible values of *PECKS*. (Recall that failures of compossibility have to do with impossibilities that obtain for non-causal reasons.) As this example illustrates, when variables are compossible, they are candidates for variables that can stand in causal relationships; when variables are not compossible, they are not candidates for a cause/effect relationship: an object’s being scarlet cannot cause it to be red.

The reason for this detour into compossibility is that the proportionality requirement (**P**) is to be understood as applying only to choices among variables that are related hierarchically in the way described previously and the values of which are not fully compossible. In other words, in the pigeon example, proportionality has to do with the choice among variables like *RED*, *SCARLET* and *C* in describing the dependency relations bearing on *PECKS*. Proportionality does not have to do with whether or not we should include the *TICKLES* variable along with, say, *RED* in accounting for *PECKS.* *TICKLES* and *RED* do not stand in the kind of hierarchical relationship to which **P** is intended to apply—we excluded such variables in the formulation of **P**. There may or may not be good reasons for including the *TICKLES* variable in an account of pecking, but they will not have to do with proportionality[[23]](#footnote-23). Thus there is a deep difference between the way in which *RED* and *SCARLET* are related and the considerations that are relevant to choosing between them as cause variables and the way in which *RED* and *TICKLES* are related—again **P** just bears on the former[[24]](#footnote-24).

Finally, let me mention an additional constraint that may seem so common-sensical as to be unnecessary[[25]](#footnote-25) but that is sometimes rejected in philosophical discussion: the alternative variables (and the dependency relationships in which they stand) must be such that we are aware of them and can formulate them: we are to choose among *known* alternative variables that we know how to connect via dependency relations we can exhibit to the effect of interest. Philosophers with a taste for “what must be possible in principle” arguments may contend that there “exists” an explanation Q in terms of fundamental physics (e.g., quantum field theory) of the pigeon’s pecking. They may then worry that **P** recommends that we always prefer explanations in terms of Q[[26]](#footnote-26). However, unless we can explicitly formulate the Q-explanation, we are not faced with the problem of choosing between Q and an explanation in terms of a variable like *RED.* **P** is not intended to address the *Q* versus *RED* problem but rather the problem of choosing among variables like *RED* and *SCARLET* and associated dependency relations that are known or that we can exhibit. I will add that to the extent that we are interested in the role of proportionality in the empirical psychology of causal cognition, this constraint will seem trivial—people’s reasoning won’t reflect the influence of variables and dependency relations that they are unaware of or are unable to formulate[[27]](#footnote-27).

Once we understand the proportionality condition to be restricted in the way described, its application is straightforward: ceteris paribus, we should choose the variable from the set **V** that correctly represents more rather than fewer of the dependency relations concerning the effect or *explanandum* that are present, up to the point at which we have specified dependency relations governing all possible values of *E*. The greater the extent to which such dependency relations are represented, the better proportionality is satisfied. Thus *RED* allows for a fuller (indeed as full as possible, given the way in which the effect is specified) representation of the causal relationships present in the pigeon example than either of the other two variables *SCARLET* and *C (*again, assuming that we are confining ourselves to variables that are hierarchically related to red) and hence better satisfies proportionality.

Several other points about (**P**) are worth noting. First, (**P**) is obviously non-binary: it can be satisfied to a greater or lesser degree by a causal claim depending on the extent to which various existing dependency relations are represented. However, if the represented dependency relations bearing on the effect fully specify conditions under which all values of the effect occur, proportionality will be fully satisfied and we can speak of the cause and effect being proportional *simpliciter.* I will suggest below that this is not as difficult to satisfy as some may suppose. By contrast, whether a causal claim is true or not, as assessed by, say, (**M)** is a binary matter. As the above discussion illustrates, a causal claim can be true, in the sense of satisfying (**M**) (or 8.3.3a) and yet not do particularly well (in comparison with alternatives) with respect to proportionality. The extent to which a causal claim satisfies proportionality (and in particular 8.3.3b) is thus a distinct dimension of causal assessment that goes beyond assessment of whether the claim is true. When thus understood proportionality should not be regarded as a candidate for a necessary condition for the truth of causal claims, contrary to the way it is often treated in the literature.

Second, note that (**P**) has to do with “objective” features of causal claims that track how matters stand in the world. That the claim (8.2.1) formulated in terms of the *RED* variable does a better job of satisfying proportionality than the alternative claims (8.2.2 ) and (8.3.4 ) formulated in terms of *SCARLET* and *C* is a reflection of the objective fact that the pigeon pecks at red and only red targets and that (8.2.1) fully captures these facts about the dependency relations present in the example and (8.2.2) and (8.3.4) do not. Finally, note also that on this understanding of proportionality there is no mystery about why proportionality is, ceteris paribus, a virtue or desirable feature of causal claims—its normative justification is straightforward. To the extent that a causal claim satisfies proportionality understood as **P** it will provide more information about dependency relations governing the effect than alternatives that satisfy proportionality to a lesser degree: it will do a better job at providing information that is associated with such distinctive aims of causal thinking as manipulation and control[[28]](#footnote-28). Although proportionality is thus a pragmatic virtue in the sense that it has a means/ends rationale or justification, it is not “pragmatic” in the sense that it depends on the idiosyncrasies of particular people’s interests or other similarly “subjective” factors. Moreover, although, as noted earlier, a causal claim can be true without satisfying (**P**) to any very great extent, the considerations that go into the assessment of proportionality are truth-based in the sense that they have to do with the extent to which a causal claim captures certain truths about dependency relationships. Proportionality is thus not based on considerations that have nothing to do with truth or with what nature is like.

Now consider another case—this one a purported counterexample due to Shapiro and Sober, 2011. Their immediate target is the following characterization of proportionality **P \*** (which is obviously different from **P**):

(**P\***) A statement of the form ‘C caused E’ obeys the constraint of proportionality precisely when C says no more than what is necessary to bring about E. (p. 89)

They assume that **P\*** is intended as a candidate for a necessary condition for a causal claim to be true. In Shapiro’s and Sober’s example, real-valued variables *X* and *Y* are related by some non-monotonic function F which maps two different values of *X*—e.g., 3 and 22—into the same value of *Y* (*y*=6), with other values of *X* being mapped into different values of *Y*. (That is, the function is not injective or 1-1.) There is an obvious sense in which *X*=3 is not “necessary” (or is not “required”) for *Y*=6 since *X*=22 also yields *Y*=6. So proportionality, understood as **P\***, is violated in this case. If we take **P\*** to be a necessary for a causal claim to be true, it follows that F does not truly describe a causal relationship, even if it correctly describes how *Y* responds to interventions on *X* and qualifies as true according to the interventionist criterion **M**. For similar reasons, P\* implies that it is false that *X=3* causes *Y= 6.*

I assume that many readers will regard these assessments (of the above causal claims as false) as “unintuitive”—this certainly seems to be Shapiro’s and Sober’s assessment and one that they expect others to share. More specifically, the intuitive judgments of many of us are that the causal claims above are not false (and arguably are not defective in other ways) merely because they involve non- injective relationships. Of course the methodology advocated in this essay does not license rejecting **P\*** merely because it leads to unintuitive conclusions. On the other hand, given that **P\*** has such consequences one may well wonder whether there is any defensible normative basis for **P\***, either understood as a necessary condition for the truth of causal claims or, more weakly, as capturing a desideratum of some kind—that being that for some reason it is desirable to avoid causal relationships that are not 1-1 or injective.

What does **P** (my preferred explication of proportionality) imply about these examples? That depends on what is included in the cause or explanans side of the examples. Consider the following possibility: in addition to the actual values of *X* and *Y*, we are given a full specification of the function F, relating *X* and *Y* so that we are told for each possible value of *X* what the corresponding value of *Y* will be. This information – call it *A*-- seems to fully satisfy **P**. *Ex hypothesi*, the information *A* does not claim the existence of any dependency relations that do not exist and (in the relevant sense) it fully describes those that do exist. This accords with our judgment that the injective character of F does not by itself show that there is anything defective about F (or *A*) as a characterization of the causal relationship in the example. This also shows that Shapiro and Sober’s **P\*** is not equivalent to **P**, even putting aside the consideration that the former is put forward as a necessary condition for causation and the latter is not. (I say more about this shortly. ) It also follows that certain informal characterizations according to which a causal claim satisfies proportionality to the extent that the cause contains only what is “necessary” or “required” for the effect or according to which causes must be just “enough” for their effects (and no more) are not equivalent to **P** . We can agree that these characterizations in terms of what is required or enough are inadequate while endorsing **P**.

Suppose, on the other hand, that rather than being given the information in *A,* the causal claim with which we are presented is just

(8.4.1) *X*= 3 causes *Y*=6

and nothing more. That is, we are not given the rest of the information contained in F , the function relating *X* and *Y*, beyond (8.4.1).Then there is an obvious sense in which relevant dependency information has been omitted. Although we haven’t been told anything that is false, we haven’t been told what the value of *Y* will be for conditions other than *X*=3. (I assume that it is not plausible to interpret 8.4.1 as claiming that for all values such *X* ≠3, *Y*, *Y* ≠ 6. In any case, even if this interpretation is accepted, (8.4.1) does not tell us which values of *Y* are associated with *X* ≠3.) So in this respect **P** is far from being fully satisfied. However, it is not clear that there is anything wrong with the assessment that, ceteris paribus, it would be preferable to replace (8.4.1) with a claim that provides more information about the dependency relations governing *Y* such as that provided by *A.* That **P** supports this assessment seems a consideration in support of, rather than against, **P**. Note also that this assessment does *not* depend on the idea that there is something unsatisfactory about causal relationships that are characterized by non- injective functions.

In Yablo’s orginal discussion of proportionality as well as a number of subsequent discussions in the philosophical literature, the causal relationships to which this notion was applied were assumed to be binary. Under this assumption, if the relationship between cause and effect is described by a non-trivial function[[29]](#footnote-29), the kind of case considered by Shapiro and Sober which involves a non-injective function cannot arise. However, many variables employed in ordinary causal cognition and in science are non – binary and it is certainly desirable to extend the notion of proportionality to cover such variables. In my view it is an important attraction of **P** that it does this.

**P** differs from Yablo’s understanding of proportionality in several other respects as well. As explained above, **P** is a graded notion and is not proposed as a necessary condition for a causal claim to be true. In addition, to anticipate the more extended discussion below, Yablo’s understanding does seem to be that non 1-1 relationships violate proportionality while as we have noted, **P** does not have this implication.

It is also worth noting that **P** has some straightforward implications for the kinds of functional relationships that will fully satisfy proportionality—one can thus think of **P** as associated with desiderata in the choice of functions to represent causal relationships, just as certain invariance conditions are. First, in the binary case, if the mapping between cause and effect is not a function, then some functional (that is, deterministic) relationship (assuming there is one that truly describes the dependencies present in the example) will better satisfy **P**. For example, the relationship between *SCARLET* and *PECKS* is not a function, since the non-scarlet value of *SCARLET* is mapped into different values of *PECKS* (depending on whether that non-scarlet value is or is not a shade of red). By contrast, the relationship between *RED* and *PECKS* is (described by) a function. When the variables involved are non- binary, **P** will again be best satisfied by some functional relationship (again if one exists that truly describes the dependency relations). As we have seen, such a function need not be one to one. However, to best satisfy **P**, the function should be onto or surjective: every value of the effect variable should be the image of some value of the cause variable. For example, given an effect variable *Y* with 3 possible values, *y1, y2, y3*, a cause variable *X* with two values *x1* and *x2* and an associated dependency relationship F for which F(*x1*) = *y1* and F(*x2*)= *y2*, **P** would be better satisfied by an alternative variable *X*\* (and an associated functional relationship (F\*) which correctly specifies the values of *X\** that are mapped into *y3* assuming there are such. The variable/ functional relationship F\*/X\* tells us more about what the values of *Y* depend on than does F/ *X*. Note also that when the relationship between cause and effect is described by a function, full satisfaction of **P** requires that there be at least as many possible values or states of the cause variable as there are states or values of the effect variable.

Suppose next that the candidate causes and effect variable are both binary but that some or all of the candidate cause/effect relationships are probabilistic (with probabilities strictly between 0 and 1) rather than deterministic. There are delicate issues about how to understand probabilities (and when ascriptions of probabilities are “correct”) in situations of this sort. However, in what follows I propose to put these aside and assume that we are choosing among candidate causes related hierarchically that bear different probabilistic relationships to the effect, where each of these relationships is correct as far as they go, but may differ in informativeness. As an example, suppose that the true probability of pecking given that the target is scarlet is 0.3, the true probability of pecking given that the target is red is 0.6 and that the probability of pecking given that the target is a warm color (which includes red and yellow, but not blue) is 1.0. What understanding of proportionality best fits such contexts? If we think of **P** as motivated by the idea that we should prefer cause variables that allow for the formulation of relationships that tell us more rather than less about the conditions under which various values of the effect occur, then if we have a choice between a cause variable *C* that allows the formulation of a deterministic relationship and an alternative variable *C\** that permits only the formulation of a probabilistic relationship, then a natural thought is that we should prefer the former – this provides more information in the relevant sense. Thus we should prefer “warm color” in the example above. Going further, we might consider generalizing this to a preference for candidate cause variables that assign more extreme rather than less extreme probability values to the effect. That is, we might prefer causes that assign probability values that are as close as attainable to 1 and 0, where “attainable” means that these probability values are empirically correct. If we follow this proposal, then, in cases in which we are choosing among hypotheses each postulating single causes[[30]](#footnote-30) (each satisfying the interventionist condition **M**) at different levels of the hierarchy, proportionality will be better satisfied to the extent that the choice of the cause variable *C* is such as to maximize *∆p = Pr(E/C)- Pr (E/-C).* In other words*,* in cases of this sort,among the candidate cause variables we should choose the one for which *∆p* is maximal, subject to the constraint that the assigned probability values are correct. As we shall see this is also the proportionality-like normative requirement advocated by Lien and Cheng (2002) for such cases. It also is the rule that best describes the behavior of the subjects in their experiments[[31]](#footnote-31).

Proportionality is sometimes presented as the requirement that cause and effect be “commensurate” or that both have a “grain” that allows them to “fit” appropriately with one another. Although this way of putting matters can be misleading[[32]](#footnote-32) we should also be able to see from **P** what is right about this idea. Full satisfaction of **P** requires (among other things) that the “grain” of the cause (understood as the number of possible values of the cause variable) should line up or fit with grain of the effect variable in the sense that the cause variable should have enough values to provide information about the conditions under which each of the possible values of the effect variable will occur, with the ideal case being one in which the value of the cause will determine the value of the effect for all values of the latter. Moreover, even when a candidate cause variable has enough states or values, it can still carve things up in a deficient way, as illustrated by 8.2.2 and 8.3.4, which we can again think of as involving a failure of appropriate graining.

**8. 5**. **P** versus **P+**

So far we have focused mainly on cases in which some candidate cause variable fails to fully satisfy **P** and another choice of variable would constitute an improvement with respect to **P**. There is, however, another possible set of cases (with Shapiro’s and Sober’s example providing one illustration) to be considered which, so to speak, are the other side of the coin. Suppose that we have a cause variable *Z* that fully satisfies **P** with respect to effect *E*. What should we make of alternative causal claims that also fully satisfy **P** but employ a cause variable *X* that is more fine grained than *Z*—that is, that makes distinctions in values or states that are unnecessary from the point of view of accounting for *E*? In other words, *X* goes “beyond” **P** in the sense of making distinctions that are not required by **P**.

Consider the following idealized illustration, which resembles a number of examples that have been discussed in the philosophical literature[[33]](#footnote-33). The effect is a variable *R*  that takes one of three possible values (*monkey reaches right= r1*, *monkey reaches left=r2, monkey does nothing =r3. )* One candidate cause variable for *R* is an “intention” variable *I* which also takes one of three possible values *i1, i2, i3* corresponding to whether the monkey intends to reach right, left or do nothing. The monkey always reaches in the direction he intends to reach, so that the relation *R= F(I)* between *I* and *R* is deterministic. Now suppose that each of these values for the intention variable *I* is multiply realized by the values of a neurological variable *N* with values *nij*: whenever either *n11*, *n12* or *n13* occurs, value *i1* of *I* is realized and correspondingly for the values *i2* (realized by *n21*, *n22*, or *n23*) and *13* (*n31, n32, n33*). These are the only possible values of *N* and *N* is the only neurological variable whose values are realizers for *I*. The relation *R= G(N)* between *N* and *R* is thus also deterministic. As before, assume that the realization relationship is not a causal relationship: the values of *N* don’t cause the values of *I* that they realize. Thus the choice between *N* and *I* is one of those choices among hierarchically related variables to which **P** applies but we will not need to be any more precise than this about what realization involves.

The variable *I* and the associated relationship *R=F (I)* satisfy **P** but so do *N* and *R= G(N),* given the way we have interpreted **P**. The difference between *F* and *G* is simply that the former is injective and the latter is not and our assessment previously was that the non-injective character of *G* was no barrier to its satisfying **P**. So if we are comparing *F* and *G*, **P** does not prefer one to the other: *G* and *N* are not penalized for making unnecessary distinctions. As far as **P** goes, *G* and *N* are just as good as *F* and *I*.

Suppose, however, that we change the comparison: in some actual case, we are given just the information that (8.5.1) *N*= *n11* causes *R=r1* (where this information is correct and is all the information that is conveyed). This claim does leave out a lot of dependency information relevant to conditions under which *R* takes its possible values and in this respect falls short of fully satisfying **P**. (In this respect it is parallel to the claim that *X*= 3 causes *Y*= 6 in the Shapiro/Sober example.) So **P** will judge alternatives to (8.5.1) such as the explanation—call it (8.5.2) -- that presents all of the information in *F* as well as the value taken by *I* as superior[[34]](#footnote-34).

One of the primary motivations for the original introduction of the notion of proportionality by Yablo was to defend the *superiority* of upper-level explanations (in particular explanations that appeal to mental causes like intentions) over lower-level explanations (in particular explanations that appeal to physical or neurological variables). Although it is arguable, as noted above, that **P** supports a preference for the psychological causal claim (8.5.2) over the neurobiological claim (8.5.1), and similarly a preference for (8.5.3) over (8.5.1)-- see footnote 463 -- we also noted that **P** does not tell us to prefer an explanation of *R* that appeals to the psychological variable *I* and the accompanying generalization *F* to the alternative explanation that appeals instead to the neurobiological variable *N* and the generalization *G*. Some (including perhaps Yablo) will think that in this second case as well the psychological explanation *is* preferable and hence that it is a limitation of **P** that it does not yield this result. Let us explore this.

I noted above that one way of thinking about the relationship between the causal claim involving *F* and the claim involving *G* is that the former makes distinctions among values of variables that are unnecessary in accounting for *R* while the latter does not do this. **P** does not treat such unnecessary distinctions as a defect but it might seem plausible that they should be so regarded—this would yield a preference for explanations appealing to *F*. [[35]](#footnote-35)

In adopting this latter view, we are in effect going beyond **P** to adopt a strengthened conception of proportionality that I will call (**P+**): not only do we want variables and generalizations that fully conform to **P** but we also don’t want variables and generalizations that make distinctions or that appeal to information that goes beyond what is required by **P.** Fully satisfying **P+** thus requires use of a cause variable that has exactly the same number of values or states (exactly the same grain) as the effect variable so that in cases in which the cause effect relationship is described by a function, that function must be 1-1. When we find ourselves operating with a cause variable *V* for which two or more values *v1, vn* are mapped into the same value of the effect, then as far as proportionality goes (there may of course be other considerations) it would be an improvement (according to **P+**) to replace *V* with another variable *V\** which collapses all such values of *V* into a single value *v\** of *V\**. (In effect, we can think of the values of *V\** as obtained via a partition of the values of *V* into equivalence classes, where values of *V* belong to the same equivalence class if they lead to the same value of *R*.)

To further illustrate the difference between **P** and **P+** , return to the contrast among the *RED* (8.2.1), *SCARLET* (8.2.2) and *CYAN* (8.3.4) claims about pecking. To explain the superiority of (8.2.1) over these alternatives, we require only **P**. By contrast, to motivate the claim that an upper level psychological explanation that appeals to *F*  is superior to the more fine grained neurobiological explanation that appeals to *G*, we require the stronger **P+**. Put differently, the considerations that support the superiority of the (8.2.1) *RED* causes *PECKS* claim are distinct from the considerations that might be claimed to support the superiority of *F* over *G* in accounting for *R.* One can agree about the superiority of (8.2.1) without agreeing that the psychological explanation appealing to *F* is superior to the neurobiological explanation appealing to *G—*the superiority of (8.2.1) doesn’t support the latter claim .

One way of describing the difference between **P** and **P+** is that the former is more permissive: **P** allows us to use more fine-grained variables and accompanying generalizations when this involves no loss of dependency information but it also does not require this unless there is a gain in dependency information. By contrast **P+** says that the use of the less fine-grained variables is preferable as long as there is no loss in dependency information. Adopting **P** rather than **P+** allows us to avoid taking on the burden of arguing that the upper-level explanation is preferable even when there is no loss of dependency information.

**8.6. Proportionality and Conditional Independence**[[36]](#footnote-36)

There is no doubt more that might be said about the comparative merits of **P** and **P+**. But rather than spending additional space on this[[37]](#footnote-37), I want to turn instead to a related consideration that bears on the motivation for adopting some version of proportionality and which also suggests that, in practical terms, the contrast between **P** and **P+** may, in many cases, be less consequential than one might suppose. As an empirical matter, there are many cases in science in which we are unable (for various reasons having to do with computational and epistemic constraints) to formulate explanations or true causal claims involving upper-level effects that appeal to fine-grained lower-level variables like the neurobiological variable *N*. In such cases, **P** (and for that matter, **P+**) can provide a kind of reassurance: if we can find an upper-level variable (or variables) fully satisfying this condition, we can appeal to these instead (rather than the more fine-grained variables), knowing that as far as proportionality considerations go, we lose nothing by doing so. Suppose (as seems likely) that in many cases we have no idea what fine- grained neural variables and accompanying generalizations are involved in the causation of behavior, but we are in possession of a psychological theory that fully satisfies **P** (or does as well with respect to **P** as any neurobiological theory that we are able to construct would). Then there is no puzzle about why we use (and think that we are justified in using) the psychological theory. I take this to suggest that, insofar as we are interested in understanding why it is reasonable to use the psychological theory, constraints like **P+** are unnecessarily strong; we don’t need them to explain (and justify) aspects of our use of upper-level causal claims and aspects of our practices of abstraction, since **P** plus the non-availability of the lower-level theory provide all the justification that is needed. On the other hand, this consideration also suggests that in many realistic cases, we may not be presented with a sharp conflict between **P** and **P+**: the causal claims that are able to produce or construct may, as in the case of the psychological theory that appeals to *F*, satisfy or come close to satisfying **P+** as well as **P,** simply becausethe many-one claims that would violate **P+** are not ones that we are able to formulate. Of course this line of argument requires that it be true, as an empirical matter, that in some substantial range of cases upper-level generalizations like the psychological explanation appealing to *F* do as well in satisfying **P** as lower level generalizations (of a kind that we are able to produce).

To further motivate this claim, let me put it in a more general context and introduce some additional apparatus. I claim that it is a striking empirical fact that the difference-making features cited in many lower-level theories sometimes can[[38]](#footnote-38) be absorbed into variables that figure in upper-level theories without a significant loss of difference-making information with respect to many of the effects explained by those upper-level theories. (Of course this is not always possible but it is more common than many suppose.) In other words, in many cases with respect to a range of possible *explananda*, the upper-level theory will do just as well as the lower level theory in satisfying **P**. The relationship between the psychological and neurological theories of reaching behavior envisioned above illustrates this possibility. Given the values of the psychological variable *I*, variations of the values of the lower-level variables *N* make no further difference to the values of *R* (as reflected in the generalization *F*), so that from the point of view of providing difference-making information about *R*, *F* and *I*  do just as well as *N* and *G.* As a more realistic illustration, it is almost but not quite true that, given the values of various thermodynamic variables like temperature that are difference-makers for various aspects of the macroscopic behavior of a gas, further variations in the microscopic state of the gas as described by, e.g., the positions and momenta of the individual molecules making it up are irrelevant to its macroscopic behavior of the gas. This is why if we wish just to explain features of the macroscopic behavior of the gas, it is often sufficient to just appeal to the value of the thermodynamic variables and the generalizations in which these figure.

To further spell this out in a more general way, let us say, following (**M**), that a set of variables *Xi*  is *unconditionally causally relevant* (alternatively, irrelevant or independent) to *E* if there are some (no) changes in the values of each *Xi*when produced by interventions that are associated with changes in *Y*[[39]](#footnote-39). A set of variables *Yk*is irrelevant to variable *E* *conditional* on additional variables *Xi*  if the *Xi*are unconditionally relevant to *E* , the *Yk*are unconditionally relevant to *E,* *and* conditional on the values of *Xi ,* changes in the value of *Yk*producedby interventions andconsistent with these values for *Xi* are (unconditionally) irrelevant to *E*[[40]](#footnote-40) . In other words, changes in the *Xi* are causally relevant to *E* in the sense captured by (**M**) and conditional on the values taken by *Xi*, further variations in the *Yk* make no difference to *E*. We can think of this as a kind of generalization of the “screening-off” idea used by Yablo to characterize proportionality, as described in Section 8. 2: the *Xi* screen off the *Yk* from *E* in the sense just described[[41]](#footnote-41). The idea that we are considering is that in such a case, we may appeal just to the *Xi* to explain the values of *E*—the *Yk* (and the generalizations in which they figure) provide no further causal or explanatory information relevant to *E*. One might of course go on to say that the explanation in terms of the *Xi* is preferable to the explanation in terms of the *Yk* (or even that the claims in terms of the *Yk* are false—this would be to adopt **P+** as opposed to **P**.) However, as we have noted, it is not necessary to claim this to vindicate the use of the *Xi* instead of the *Yk* – all that is required (in accordance with **P**) is that the explanation in terms of the *Xi* be as good as the explanation in terms of the *Yk*. In other words, we can think of **P** as reflecting or fitting with the ideas about conditional irrelevance just described. Put in terms of proportionality requirements, the idea is that we can often find upper-level variables, related to lower-level variables, typically by some kind of coarse graining operation, that satisfy the requirements of proportionality just as well as the lower-level variables and thus (as far as proportionality is concerned) are equally satisfactory form the point of view of explanation and causal analysis[[42]](#footnote-42). Although if we accept **P**, there is nothing wrong with the use of variables that make finer grained distinctions, there may also be no motivation for employing such variables—no further gain from the point of view of causal analysis or explanation.

**8. 6. Some Empirical Questions**

As so far described, **P, P+** and the accompanying ideas about conditional irrelevance are normative proposals. However, one can also ask about the extent to which, as an empirical matter, various groups of subjects conform to these principles and this in turn suggests a number of experimental questions.

**8.6. 1. Causal/Explanatory Strength Judgments**

Suppose subjects are presented with information about two sets of variables *X* and *Y* that are related to one another via a non-causal coarsening or hierarchical relation (*X* is a coarsening of *Y,* so that in this sense it is at a higher “level” than *Y).*  *X* and *Y* are both unconditionally relevant to *E* and conditional on *X*, *Y* is irrelevant to *E*. Subjects are then presented with candidate causal/explanatory claims relating *X* to *E* and relating *Y* to *E*. Is it the case, as **P** and the conditional irrelevance idea would suggest, that they would regard causal and explanatory claims about *E* that are framed in terms of the coarser grained variable *X* as just as satisfactory as the those framed in terms of *Y*? That is, would the subjects give strength ratings (for some appropriate verbal probe) to causal and explanatory claims appealing to *X* that are as high as the ratings given to causal and explanatory claims appealing to *Y*? Alternatively, would they regard causal and explanatory claims appealing to *X* as better than those appealing to *Y*, on the grounds that *Y* makes unnecessary distinctions from the point of view of accounting for *E* – that is, do their causal and explanatory judgments follow **P+** rather than **P**? Or would they judge that causal claims appealing to the more fine-grained variable Y on the grounds that more detail is always better (causes need to be maximally specific)?

**8.6. 2.** **Learning variables and relationships at particular levels of abstraction.**

In (8.6.1) we supposed that subjects were presented with alternative causal claims formulated at different “levels” and asked to rate them. That is, they are not asked to discover or formulate these claims themselves but just to rate them once they are provided. Another possibility is not to provide subjects with explicit formulations of these alternative hypotheses but rather to explore whether subjects can discover them, given appropriate dependency information. If subjects can do this, do they then exhibit preferences for hypotheses at certain levels and if so, what principles guide their choices? Suppose, for example, subjects are given information about patterns of dependence that allow for the discovery of relationships between *X* and *E* and between *Y* and *E* where again *X* is related to *Y* via some coarse graining or similar operation. (Subjects are not told, however, that *X*, *Y* or alternative variables are appropriate ones to use.) When asked questions like “What causes (or explains *E*)?” do subjects prefer to cite one of these variables rather than the other? Can they discover these variables (and causal hypotheses formulated in terms of them) from information about dependency or contingency information even if they are not explicitly told about these variables and relationships? To what extent can they discover and to what extent do they prefer to cite variables and accompanying causal claims that satisfy **P** (or **P+** ) better than alternatives? When subjects learn about dependency relations in one set of circumstances, which variables and levels of abstraction do they prefer when asked to generalize to new situations?[[43]](#footnote-43)

Of course it is natural to expect that many of these questions will receive similar results—that is, that subjects will assign higher strength ratings to those causal claims that they are most willing to spontaneously endorse and so on. This is what is found in an empirical study by Lien and Cheng discussed in the following section. However, it is of course an empirical question whether this pattern holds generally.

**8.7. Lien and Cheng on Levels of Abstractness**.

I turn next to some experimental results due to Lien and Cheng (2002) bearing on the extent to subjects are guided by proportionality-like considerations in causal judgment and inference. As noted earlier Lien and Cheng do not use the word “proportionality” in describing their results, but as far as I am aware, theirs is the empirical study that bears most directly on this notion.

In broad outline Lien and Cheng presented subjects with hypothetical soil ingredients that (they were told) were candidates for causes of plant blooming as well as contingency information linking whether plants were fed these ingredients and the rate at which they bloomed. These ingredients were described as varying in ways that fell into hierarchical structures or classes of increasing abstractness, so that these classes were related in the way described by **P**. The experimental questions then had to do with whether the preferentially learned and endorsed claims about the causes of blooming formulated at one of these levels and what principles appeared to be guiding this choice of levels.

In more detail, the experiment (which had a rather complicated design) proceeded as follows: in the *learning* phase, subjects were given information about different hypothetical substances *s* that that were fed to groups of a particular kind of plant as well as the frequency of blooming in each group. They were told that plants were kept in the same constant environmental conditions, so that any variation in their blooming would be attributable to the substances they were fed. Subjects were also given information showing that most plants that were not fed with any substance did not bloom. Each substance was described at varying levels of abstractness. For example, the substances varied as to color and these variations were represented at each of three levels of abstractness – particular shades of color (e.g., pine-green), general type of color (green), and (at a very general level) whether the color was “warm or cool”. Similarly, the substances varied in shape and subjects were given information that represented these shapes both in highly specific ways (e.g., as a triangle of certain dimensions), at an intermediate level of specificity (e.g., a type of shape such as triangular), or in a still more abstract way (regular in the sense of rotationally symmetrical versus irregular).

In this phase of the experiment, although each subject was shown information about the covariation between each of the substances and blooming , the frequency of blooming was manipulated across two groups of subjects. In each such group, the patterns of covariation were chosen in such a way that the contrast *∆p = Pr(E/C)- Pr (E/-C)* was maximized when the candidate cause *C* was at the most abstract category in the hierarchy in which it fell. However the causally relevant variable varied across the two groups: for one group shape was the relevant variable with the irregular shapes in contrast to regular shapes producing the maximum contrast. In the other group color was the relevant variable, with *∆p* being maximized when the color of *s* was warm rather than cool. (Recall that choosing the cause variable for which *∆p* is maximal is the natural interpretation of **P** for binary probabilistic causes, assuming that there is just one cause of the effect.) In contrast, for each group, blooming versus non-blooming covaried less than maximally with more specific descriptions of the soil ingredients. That is, *∆p* was smaller when the substances were described as blue rather than red or as a triangle with an irregular shape rather than a regular one. Then in a “dynamic learning phase” subjects were asked to predict the frequency of blooming for various groups of plants fed various substances on the basis of the information that they had previously received and were given feedback about the correctness of their answers.

Next, the *test* phase: First, subjects in both groups were given a categorization task in which they were asked to sort substances in accord with what caused blooming. The goal here was to determine the level of abstractness of the causal relationship which the subjects recognized and employed in this task. Most subjects gave the highest causal ratings to the most abstract causal relationships. That is, they categorized in accord with the irregular/ regular shape contrast or the warm/cool color contrast, depending on the data they had seen.

Second, subjects in both of the groups previously established were told that the plants were placed in a novel environment and were fed a novel substance *s\** which was irregularly shaped and warm colored and that most of the plants bloomed. Subjects were then asked whether *s\** caused the plants to bloom. Notice that in this case because both the environment and *s\** are new (and perfectly correlated), the covariational evidence from this new experiment alone is ambiguous about whether *s\** or the new environment caused blooming. However, Lien and Cheng hypothesized that the subjects who had learned (during the learning phase) that irregular shapes caused blooming would be more willing to judge that in this new situation *s\** caused the blooming (since *s\** is irregular), in comparison with the subjects who had learned in the test phase that cool colors cause blooming. This is because *s\** is warm colored and hence the claim that s\* caused blooming was “inconsistent” with what the latter group had previously learned. This prediction was born out. Subjects in the former group give a higher causal rating to the claim that *s\** caused the blooming.

The general pattern in the experiment, then, is that when given a choice among causal relationships that might be formulated at different levels (where these levels stand in a hierarchical relationship) subjects appeared to learn or induce causal relationships from the data they were given at the level that maximized *∆ p* and that they were guided by relationships at this level in their willingness to infer new causal relationships. In the particular case explored in Lien and Cheng’s experiment, the level or choice of variable that maximized *∆ p* was the most abstract level but of course the covariational data might have been chosen in such a way that the strategy of choosing the variable or level that maximizes *∆ p* leads instead to the choice of an intermediate or maximally specific level. Lien and Cheng did not do such experiments but in such cases one would expect subjects to prefer these intermediate levels instead. As noted earlier, in the case of binary variables related in the way described in Lien’s and Cheng’s experiment, the choice of variables that maximizes *∆ p* will also be the choice that best satisfies **P**, so that we can also think of their experiment as showing that subjects choose in accord with **P** in their experiment. (Their experiment does not discriminate between **P** and **P+**).

These experiments provide an additional illustration of the interplay between descriptive and normative considerations in causal cognition that is one of the main themes of this book. First, in agreement with the claims of Yablo and others, the experiments show that subjects do sometimes prefer (learn and make use of) causal relationships or descriptions of causal relationships that are not characterized in a maximally specific or detailed way. Instead, they sometimes prefer more abstract characterizations. More specifically, as an empirical matter subjects choose in a way that maximizes the contrast *∆p = Pr(E/C)- Pr (E/-C)* . To the extent that the generalization that subjects choose in a way that maximizes *∆p* holds more generally, this implies that which level of abstractness subjects prefer for the choice of causal variable will depend on the empirical details of the case. Sometimes they will prefer more abstract descriptions and sometimes more specific descriptions. More specifically, at least in these experiments, subjects infer and judge in a way that conforms to a proportionality requirement.

Of course the empirical fact that subjects do this does not by itself show that it is justifiable or normatively correct for them to dos so. However, as Lien and Cheng argue, and is also suggested by our earlier discussion of **P**, there is an obvious normative rationale for this behavior. This suggests that subject preference for variables at a more abstract level in these experiments is not a mistake or due to confusion. Subject preference for characterizations of cause that maximize the contrast *∆p = Pr(E/C)- Pr (E/-C)*  makes sense because this is the level that is most informative about the conditions under which the effect will and will not occur. This raises the following question for those who (typically on metaphysical grounds) claim that more specific cause variables are always better: in what sense , if any, are those who in the appropriate circumstances prefer more abstract or coarser cause variables that better satisfy proportionality to more specific cause variables making a mistake? Why should we be bound by some supposed metaphysical requirement to provide maximally specific descriptions of causes if proportional descriptions are often more informative about matters that we care about?[[44]](#footnote-44)

**8. 7. Conclusion**

A convention of book writing is that books should not end abruptly, but instead should have a substantial conclusion as a kind of gentle leave-taking. But this is already a relatively long book and I see little point in rehashing what I have said earlier. My overall message has been that there is a fruitful collaborative project concerning understanding causation and causal cognition—one that draws together results from philosophy, psychology and other disciplines like statistics and machine learning. I hope to have convinced readers that progress on these topics can be made by cooperation across different disciplines rather than by boundary policing and strategies of dismissal.

**Appendix 1: Proportionality as Applied to Actual Cause Claims**

I noted in 8.1 that a proportionality requirement often seems more plausible or natural (where this is understood as an empirical observation about how people judge) when applied to type as opposed to token (or actual cause claims). For example, (1) “the impact of the rock thrown by Suzy caused the window to break” may seem completely appropriate even though (1) apparently does less well along the dimension of proportionality than (2) “the impact of a rock caused the window to break”, assuming that given the impact it does not matter whether the thrower is Suzy. Here is a conjecture that attempts to make normative sense of this observation: The basic idea is that, as suggested in Chapter 2, Appendix, actual cause claims—particularly those involving human actions, are commonly used to ascribe responsibility. (This does *not* require that “cause” in such cases just means “responsible” but merely that the causal claim is used in the way described.) Ascriptions of responsibility may require non-proportional characterizations of causes because this is the only available way of distinguishing among candidate causes for purposes of ascribing responsibility. For example, if the impact of a rock causes a window to break, then, to the extent that we are interested in attributing responsibility, we may be very interested in whether it was the rock thrown by Suzy or the rock thrown by Billy that caused the breaking. If it was Suzy’s rock, we accept (1) above and regard (2) as less satisfactory because it provides no information about the responsible party. More generally, recall from Chapter 2, Appendix, that in making actual cause claims, we are often interested in discriminating among various candidates for the actual cause of some outcome . In a number of cases, the only characterizations of these candidate causes that we have available and that can be used to distinguish among them may involve some features that are causally irrelevant to the effect. If the candidate causes for the explosion are a gas leak and Jones’ mixing “those yellow chemicals” together in the laboratory (this is the only description we have available) we may settle on the latter as the actual cause despite the fact that it involves a characterization of the cause that does not do very well with respect to proportionality.

We can get some additional insight into what is going on in these examples by considering the variables used to represent them. In the Billy/Suzy rock throwing example, we should employ two variables, one *B* representing whether Billy’s rock impacts the bottle or not, and the other, *S,* representing whether Suzy’s rock impacts the bottle. Notice that *B* and *S* are fully distinct variables with compossible values. (We assume it is possible for both rocks or neither to hit the bottle.) This makes it clear that in the Billy/Suzy causal attribution example, we face a very different problem from the problem that proportionality is designed to address. In the Billy/Suzy case, we want to know which of the values of two fully distinct variables was the actual cause of the outcome. By contrast, as explained above, proportionality has to do with a choice among variables that stand in non-causal hierarchal relations, where the values of these different candidate variables are not fully compossible. It is not surprising that proportionality seems a more appropriate consideration in this second case than in the first.

To reinforce this analysis, consider the following variant of the pigeon example. As before the pigeon has been trained to peck at red and only red targets and this has involved exposing it to training targets of different shades of red. However, now Billy and Suzy are playing a game. Suzy possesses a scarlet target which she can present to the pigeon and Billy a maroon target. The first person to cause the pigeon to peck wins a prize. In this situation we are interested in whether the (3) presentation of the scarlet target or (4) the presentation of the maroon target caused the pigeon to peck since this is what will determine the winner of the game. If (3) is true of what caused the pecking, we are not going to reject it, as a basis for determining the winner, on the grounds that is less proportional than the claim that attributes the causation of the pecking to redness.

**Appendix 2: Proportionality for Effects?**

In the discussion of **P** in this chapter I restricted the application of proportionality to a situation in which the effect variable was fixed or pre-specified, so that proportionality was understood as criterion for choosing among alternative cause variables that are hierarchically related. As I noted, this was a departure from a previous discussion of mine (Woodward, 2010) in which I treated proportionality as a condition on effect variables as well. For example, in discussing a case of Kendler’s which had to do with the best way of characterizing the effects of a particular allele a, I suggested that describing the allele as causing a tendency to take risks was superior to characterizing the allele as causing a tendency to engage in a number of distinct effects including sky diving, cliff jumping, and driving fast. Whatever one thinks about this suggestion, it is now clear to me that it addresses a different question than the question to which **P** is addressed. In Kendler’s case, we have a fixed cause, the presence or absence of the allele a, and different candidates for the effect (or effects) variable(s). Moreover, rather than (as **P** requires) comparing single variables at different levels of abstractness, as in the case of *SCARLET* vs *RED,* we are comparing a single variable *RISKTAKING* with an alternative choice in which there are a number of distinct effect variables *SKYDIVING, FASTDRIVING* and so on, values of which (for different variables) are fully compossible. (Note, though, that the values of the more abstract variable *RISKTAKING* and the more specific variables like *SKYDIVING* that realize itare not fully compossible and seem to be at different “levels”)

I noted above that the problem of choosing variables when both the cause and effect variable are allowed to vary in grain or abstractness seemed hopelessly indeterminate. However, one might wonder about the following: suppose that we take the cause variable to be fixed, as in the allele a case above, and think of ourselves as choosing among different effect variables for that cause, at different levels of abstraction. Is there some condition that might be used to guide such choices? I have no suggestions about this but it is a worthwhile problem.

1. Note for readers familiar with my previous discussions of proportionality in Woodward, 2008 and Woodward, 2010: I now think that these previous treatments are inadequate in various respects, both in failing to restrict the application of the notion of proportionality in certain necessary ways (described below) and in understanding proportionality as imposing the requirement that the functional relation between cause and effect must be 1-1. The formulation I currently prefer—**P** in Section 8.4 -- abandons the 1-1 requirement and is intended to replace what I have said earlier about proportionality. For further discussion, see Woodward, 2018. [↑](#footnote-ref-1)
2. As argued above, invariance is another criterion that legitimately influences variable choice. [↑](#footnote-ref-2)
3. It is crucial to what follows that when variables or properties at one level are more fine-grained realizations of variables or properties at a “higher” level, the relations between these two sets of properties should not be thought of as causal. The relationship is rather of some other kind characterized by notions like supervenience or realization. [↑](#footnote-ref-3)
4. Some writers deny that abstract properties can ever be causally efficacious, claiming that strictly speaking only maximally specific properties can be causes – for discussion see, e.g. Crane, 2007. [↑](#footnote-ref-4)
5. Much of the philosophical literature, including Yablo’s original discussion focuses on the role of proportionality in token causal claims. [↑](#footnote-ref-5)
6. I use this example because it has been widely used in discussions of proportionality. For some illustrations of the use of proportionality that are more scientifically serious, see Kendler, 2005 and Woodward, 2010. [↑](#footnote-ref-6)
7. “Screening off” in this context is thus understood in terms of counterfactuals rather than in terms of conditional independence. For a more precise definition see footnote 480. [↑](#footnote-ref-7)
8. There are many other similar examples in the literature. Suppose (Woodward, 2008) a kind of platform will collapse if more than 2000kg is placed on it. Weights of 3173 kg are placed on a series of such platforms and each collapses. The claim that placing weights of 3173 kg on the platforms causes collapse does not satisfy proportionality or satisfies it less well than some alternatives, while the claim that placing a weight of more than 2000 kg causes collapse does satisfy proportionality. Similarly, to take an example from the psychological literature ( Lien and Cheng, 2002, although they do not use the word “proportionality) “smoking Virginia slims causes lung cancer” and “inhaling fumes causes lung cancer” are respectively too narrow and too broad in the characterizations of their causes, while “smoking causes lung cancer” does much better with respect to proportionality. [↑](#footnote-ref-8)
9. Obviously Yablo expects his readers to respond to examples like (8.2.1) and (8.2.2) in a way that is consistent with the judgments that he thinks are normatively appropriate. That is, he expects that readers will, as an empirical matter, share his assessments of (8.2.1) and (8.2.2). This is one of many cases in which cases or examples are used by philosophers against the background of the assumption that judgments about those cases will be shared, so that, as claimed in Chapter 3, the philosopher uses his own responses to predict how others will judge. [↑](#footnote-ref-9)
10. Issues having to do with granularity of variables and proportionality are discussed in Soo, 2019. After drafting this chapter an anonymous referee for OUP drew my attention to Johnson and Keil (2014). Like Lien and Cheng, this paper also explores how causal judgments are influenced by level- based considerations. However, Johnson and Keil use examples of levels involving what they call partonomic relations . These are cases in which candidate causes and effects at different levels are related by part/whole or component relations—e.g., superordinate chemical reactions having subordinate reactions as parts or subprocesses. This contrasts with the Lien and Cheng’s examples as well as the cases involving proportionality discussed in the philosophical literature, which involve what Johnson and Keil call “taxonomic” relations—relations like that between “red” and “scarlet”. Nonetheless Johnson and Keil find that subjects assign stronger causal ratings to variables at the same level or with matching grain. One can think of this as also reflecting a proportionality-like consideration. Their paper is very interesting on the subject of the role of superordinate/subordinate relations and strategies of matching levels in restricting hypothesis spaces and guiding causal inference. Finally, as I was revising the final version of this manuscript, my attention was drawn to Blanchard et al., 2020. This is an experimental study of ordinary judgments about causal exclusion – judgments about the extent to which causal claims formulated in terms of hierarchically related lower and higher level variables “exclude” one another. They find, in agreement with the views defended later in this chapter, that ordinary judgment does not conform to the exclusionist intuitions advocated by many philosophers. [↑](#footnote-ref-10)
11. These passages are quoted in Harbecke, forthcoming. [↑](#footnote-ref-11)
12. Even if, as suggested in footnote 438, explanation probes are more sensitive to mechanistic information than to covariational information, it certainly doesn’t follow that proportionality has to do with causal explanation rather than causation. Indeed, proportionality considerations often seem to abstract away from mechanistic information and, as observed below, to track covariation information. [↑](#footnote-ref-12)
13. To take a trivial example, “the cause of *e* is the cause of *e*” is (assuming that *e* has a cause) a true causal claim, but (one would think) not explanatory. But I don’t think that we need a sharp distinction between “cause” and “causally explains” to make sense of this. [↑](#footnote-ref-13)
14. In their (2018), Robb and Heil claim that causation is a “metaphysical” notion and causal explanation an “epistemological” notion and that appeals to what they call “explanatory practice” in discussions of mental causation conflate these two notions. But even accepting these characterizations, it does not follow, as they seem to suppose, that causal explanation and causation have nothing to do with one another or that one cannot learn something about how people think (and ought to think) about causation by considering how they think (and ought to think) about causal explanation and conversely. Robb’s and Heil’s discussion thus strikes me as a good example of the misuse of the epistemology/metaphysics distinction that I complain about in the Foreward.

    There are other prominent philosophical account that distinguish between causal claims and causal explanations but in my view these too are problematic. For example, according to Lewis (1986) while causal claims cite causes, a causal explanation of some event *e* can work by citing any causal information about the causal history of *e*, including the information that *e* had no causes, so that the latter counts as a causal explanation even though it does not work by citing causes. I doubt, however, that many people would regard “*e* had no causes” as a causal explanation of why *e* occurred and it is hard to see what the normative justification would be for judging otherwise. [↑](#footnote-ref-14)
15. On the other hand, suppose that for some reason one is convinced that proportionality has to do (just) with causal explanation and not with causation. Then one can understand the discussion that follows as having to do with the role of proportionality in causal explanation. However, I would still want to resist any implication that because proportionality is associated with explanation, it is for this reason uninteresting or unworthy or attention or “pragmatic” in a way that implies it is “subjective”. In this connection it is worth noting that in the quotations from Bontly and Mclaughlin above, the association of proportionality with pragmatic considerations or explanation functions as rhetorical device for not considering it further. It is yet another example of what I called a device of dismissal in the Foreward. [↑](#footnote-ref-15)
16. See Woodward, 2003. [↑](#footnote-ref-16)
17. Interventions that set *SCARLET= not scarlet* are ambiguous interventions in the sense of Scheines and Spirtes (2004). However, the case is different from their total cholesterol case (discussed in chapter 5) in several ways: First, in the total cholesterol case, interventions are ambiguous for all values of the total cholesterol variable. In the pigeon case, one of the two possible interventions-- the *SCARLET= scarlet* intervention is-- not ambiguous. Moreover, in the cholesterol case, unlike the pigeon case, the candidate cause variable cannot take the value “absent” . Perhaps in the case of binary variables we have a higher tolerance for ambiguous interventions involving setting the variable to the value “absent” as long as interventions are not ambiguous for the value “present”. [↑](#footnote-ref-17)
18. Some (perhaps many) readers may find it natural to regard (8.2.2) as false and hence to think that the “some interventions” interpretation according to which it is regarded as true is farfetched. It is worth noting, however, that influential theories of causation apparently judge it to be true. For example, Lewis (1986) tells us at one point that in evaluating a counterfactual of form , “If c had not occurred,…) we should imagine a possible world like the actual world in which c is “wholly excised” (rather than replaced with some event which is similar to c). This can be interpreted as implying that in assessing the counterfactual “if the target had not been scarlet,…” we are to imagine a situation in which no target is presented or something similar, in which case the pigeon will not peck. Under this interpretation (8.2.2) is true. In addition, considerations of charity seem to lead us to regard causal claims that are overly specific in the way that (8.2.2) is as true. To use an example of Glymour’s (1986), suppose that it is claimed that (S) “Shlomo’s smoking 4 packs a day caused his lung cancer”. Here we tend to take the contrast state to be one in which Shlomo does not smoke at all (or very little—his smoking is “wholly excised” ) and, on the assumption that in this case, Shlomo would not have developed lung cancer, we judge (S) to be true. It seems uncharitable to judge (S) to be false on the grounds that if Shlomo had not smoked 4 packs but 3.5 packs instead, he still would have got lung cancer. The interpretation that allows S to be true corresponds to the “some interventions” reading of (8.2.2). [↑](#footnote-ref-18)
19. *C* is what Franklin- Hall calls a “non-exhaustive” variable in the sense that it fails to fully span or exhaust what we intuitively think of as the relevant space of possible values for a “good” cause variable—in the present case, this is the full range of possible colors. One might think *C* is an objectionable variable for that reason. It is not clear to me, however, how to formulate a requirement that cause variables be exhaustive in a way that is clear, normatively defensible, and distinct from the proportionality requirement **P** formulated below. In one sense every variable seems “exhaustive” of its possible values since variables are defined in part in terms of their possible values. If every variable is exhaustive, what is really meant by the complaint of non-exhaustivity is that we should be employing a different variable. [↑](#footnote-ref-19)
20. Two caveats and a clarification: First, as already intimated, proportionality is a ceteris paribus desideratum—a cause variable that satisfies proportionality can be defective in other ways and thus a bad choice for such a variable. Second, although **P** is intended to apply to certain situations in which we choose among candidate cause variables that vary in their level of abstractness, I don’t claim **P** applies to all cases in which candidate variables differ in abstractness – see above for the kinds of cases to which I take **P** to apply. Finally, as remarked earlier, **P** is intended to replace the characterization of proportionality—also called **P** -- in Woodward, 2010. As I said at the beginning of this chapter, I now regard the formulation in Woodward, 2010 as mistaken for several reasons, not the least of which is that it wrongly imposes the requirement that fully proportional relations must be injective. [↑](#footnote-ref-20)
21. If we count any failure to represent causally relevant information about any possible effect (e.g., whether the pigeon blinks or coos) rather than a prespecified one as a failure of proportionality, we will render that requirement effectively empty. Every causal claim (other than a theory of everything) will exhibit arbitrarily large failures of proportionality. [↑](#footnote-ref-21)
22. This reflects the generally accepted requirement that variables can only stand in causal relations when they are “fully distinct” as well as a normative criterion for when variables are “fully distinct” that I and others have defended elsewhere--- see Woodward, 2015b, 2016a, Forthcoming c, Hitchcock (2012). Briefly this requires that we distinguish between variables and their values. While fully distinct variables must have values all of which are compossible, different values of the same variable always exclude one another in the sense that the same individual cannot take two distinct values for the same variable. Thus if the values for a color variable include red and blue, the same target cannot take both values. This gives us a criterion for when we should employ causal representations with distinct variables rather than collapsing those variables into a single variable with distinct values. For an application of this idea to Franklin-Hall’s (2016) discussion of proportionality, see Woodward, 2018a. [↑](#footnote-ref-22)
23. For additional discussion and motivation for restricting **P** in this way see Woodward, 2018a and also Blanchard, 2018. (Although Blanchard does not endorse **P,** he endorses a similar restriction on the application of a proportionality requirement.) [↑](#footnote-ref-23)
24. To use language that is sometimes employed in discussions of proportionality we need to distinguish the “horizontal” problem of choosing among distinct variables from the “vertical” problem of choosing among variables that are hierarchically related. **P** only applies to the latter. [↑](#footnote-ref-24)
25. In formulating a proportionality-like condition, Lien and Cheng, 2002 describe the requirement that the subjects choose among known alternative variables as “obvious” and treat it as not requiring additional discussion for this reason. [↑](#footnote-ref-25)
26. Even supposing that we are able to formulate an explanation in terms of Q, it is not clear that **P** judges it superior to more upper level explanations for reasons discussed in Section 8. 6. [↑](#footnote-ref-26)
27. Both Franklin-Hall (2016) and Weslake (2010) use examples involving the in principle possibility (where this notion is not further explained or characterized) of formulating causal claims or explanations in terms of “fundamental physics” (as with Q above) to object to previous formulations of proportionality by me and others. They seem to think that it should not matter that we cannot actually state or write down such explanations, presumably because from the point of view of the metaphysician, all that matters is in principle possibility. But if we are interested in methodology or empirical psychology, the distinction between causal and explanatory claims that we have some realistic possibility of formulating and those that we do not clearly does matter. Here too my remarks in the Foreward are relevant. [↑](#footnote-ref-27)
28. At the risk of repeating the discussion in previous chapters it may be useful to contrast this rationale or justification for adoption of a proportionality requirement with a more intuition and conceptual analysis-based justification which is nicely described although not endorsed in Dowe (2010). In Dowe’s words

    [This ] is the view that philosophy is conceptual analysis, the aim being to give a theory which best captures the concept of causation that we all share, and so evidence for and against a theory should be the intuitions that any competent user of English will have concerning the truth of sentences of the form ‘this causes that’. Folk intuitions are that sentences like ‘being presented with a red object caused Sophie to peck’ and ‘my deciding to lift my arm is the cause of my doing so’ are true; i.e. ultimately such intuitions are what justify the proportionality principle. (447)

    The rationale I describe above does not claim that a proportionality requirement is built into the concept of causation, does not claim that intuitions concerning proportionality reflect our mastery of the concept of cause and does not claim that such intuitions are what justify proportionality. Rather explained above, the justification for proportionality has to do with the desirable information that causal claims better satisfying proportionality provide. [↑](#footnote-ref-28)
29. Non-trivial means that the function is not a constant function; that it maps some different values of *X* into different values of *Y*. That *Y* will be different for at least two different values of *X* is required by the interventionist criterion **M**. However, this does not imply that just any non-trivial functional relationship in the binary case will satisfy **P** well. The relationship (8.3.4) in Franklin-Hall’s *CYAN* example is such a non-trivial function but does not satisfy **P** very well. [↑](#footnote-ref-29)
30. The restriction to cases in which at each level there is a single candidate cause for *E* is important: unlike the situation for which Cheng’s causal power measure is designed, we are assuming that at each level there are no “other causes” of *E* that operate independently of *C*. Again, we are comparing candidate causes *RED* and *SCARLET* which are at different levels of abstraction In such cases, causal power reduces to *∆p.* [↑](#footnote-ref-30)
31. We can illustrate this idea by returning to the pigeon example. The contrast *∆p* between the probability of pecking given that the target is red and the probability of pecking, given that the target is not red is maximal—it is equal to 1- 0 =1. By contrast, although *Pr (Pecks/ target is scarlet)= 1*, *Pr (Pecks/ target is not scarlet)* is, assuming it is well defined, greater than zero but less than one. Thus *∆p* when the cause is characterized as “red” is greater than *∆p* when the cause is characterized as “scarlet” and on this basis “red” is preferable. [↑](#footnote-ref-31)
32. For example, the variable *C* in 8.3.4 has (in one relevant respect) the same grain as the *explanandum* *PECKS* (since both have two possible states or values) but (8.3.4) is defective from the point of view of **P**. [↑](#footnote-ref-32)
33. See, e.g., Woodward, 2008. [↑](#footnote-ref-33)
34. Consider yet another comparison: the claim that (8.5.3) *I* = *i1* causes *R= r1* (where again this claim is true and it is all the information we are given). This also leaves out dependency information about some of the conditions under which *R* takes its possible values (it does not tell us the conditions under which *R= r2* or *R= r3* ) and in this respect falls short of fully satisfying **P**. However, it might be argued that (8.5.1) does worse than (8.5.3) with respect to **P**: (8.5.1) tells us only what will happen with respect to *R* for one of the nine possible values of *N*, while (8.5.3)  is comparatively more informative in telling us what will happen under one of the three possible values of *I*. Furthermore, in the one case in which *N* is informative about *R*, the corresponding value of *I* conveys the same information about *R*, so that in this sense (8.5.1) does not tell us anything about *R* that we are not told by (8.5.3).Thus in this case it is arguable that **P** tells us to prefer 8.5.3 to 8.5.1. [↑](#footnote-ref-34)
35. It is also sometimes argued that the causal claim formulated in terms of *G* and *N* is defective in failing to show what the various states of *N* that lead to the same state of *R* have in common while the causal claim formulated in terms of *F* and *I* does not have this defect. As formulated this claim is mistaken. Although it is true that if the claim involving *G* just consists of the information described above, it does not tell us what the various values of *N* leading to the same state of *I* have in common, the claim involving *F* and connecting *I* to *R* suffers from the same deficiency. *F* *asserts* that *i1* leads to *r1* and so on, but it does not explain *why* this happens. Nor does *F* tell us what the various neural realizers of *i1* etc. have in common. A non-trivial way of showing what all of the values of *N* that lead to the same value of *R* have in common would be to identify some more abstract common neurobiological characterization of those neural states. [↑](#footnote-ref-35)
36. For more on conditional independence and related themes, see Woodward, forthcoming, a, b and c. [↑](#footnote-ref-36)
37. Nonetheless I can’t resist some additional brief remarks. It might be claimed that **P+** is superior to **P** on the grounds that **P** allows (but **P+** does not) causal claims that involve redundancies or that **P** allows claims that are less simple and compact than the claims permitted by **P**. For example, when *G* and *N* versus *F* and *I* are compared as causal accounts of *R*, *F* and *I* may seem simpler on the grounds that they require only that we keep track of three different states and how they are related to *R* as opposed to nine different states required if we employ *G* and *N*. The account in terms of *I* thus gives us a more compact summary of what is relevant to *R*. Note however that this looks like superiority along the dimension of what is sometimes called descriptive simplicity as opposed to the kind of simplicity that is supposedly relevant to choosing among different empirical alternatives. That is, as far as *R* goes, *G* and *N* and *F* and *I* seem to convey effectively the same empirical information. If this is right the explanation that appeals to *F* and *I* is “simpler” than the explanation that appeals to *G* and *N* in something like the way in which the use of polar rather than Cartesian coordinates might be simpler for a certain problem, even though either choice of coordinates conveys the same information. If so, although causal claims conforming to **P+** may be easier to use, we don’t need to think of them as superior to those conforming to **P** along the dimension of conveying dependency information relevant to the *explanandum*. Of course it is possible that there is some additional dimension that is relevant to causal assessment that supports the use of **P+** and does not reduce to ease of use but so far no candidates have been produced.

    A second relevant observation is due to Hoffman-Kloss (2014). She criticizes previous formulations of proportionality including mine by noting their awkward implications for common cause structures. To use her example, suppose one has a heat source which is used to melt chocolate in a water bath which is also attached to a thermometer , so that the temperature of the heat source is a common cause of the melting and the readings. The chocolate will melt when the temperature is greater than 35 degrees C, so if the dependent variable measuring melting is two valued *E* = {melts, not melts}, the cause variable that best satisfies **P+** and does not make unnecessary distinctions will similarly have just two values corresponding to temperatures above and below 35 degrees. However, the thermometer readings *R* register the water temperature in increments of a single degree, so satisfying **P+** with respect to *R* requires a much more fine grained cause variable. Such a fine grained variable will of course violate **P+** with respect to *M.* Similar problems will arise in many other cases in which variables have multiple effects. **P** avoids this difficulty. [↑](#footnote-ref-37)
38. See Woodward, forthcoming b in support of this claim. [↑](#footnote-ref-38)
39. Ideas that are similar to the account developed here are developed in a machine learning context in considerably more formal detail in Chalupka et al. 2015 and Chalupka et al. 2017. This work has influenced my interpretation of proportionality and conditional independence. Similar ideas can also be found in Ellis (2016) in the context of a discussion of downward causation. [↑](#footnote-ref-39)
40. Some additional observations: First, as originally employed by writers like Reichenbach (1956) and Salmon, 1970, “screening off” is a statistical notion that is characterized in terms of conditional probabilistic (in)dependence. In the present context, I follow Yablo in understanding this notion (and the accompanying notions of relevance/irrelevance) in terms of counterfactuals describing what happens under interventions rather than in terms of statistical dependence —e.g., if  *Y* is conditionally irrelevant to  *E*, given  *X*, then, if (i) one intervenes to fix *X*at some value, (ii) further variations in  *Y*  due to interventions consistent with (i) will not change *E*. (In other words, we are considering nested counterfactuals the antecedents of which involve reference to two interventions: *X*  is fixed at some value via an intervention and then  *Y* is set via a separate intervention at some value that is consistent with the value to which  *X*  has been set. Note that while there is no inconsistency in setting  *X*  to some value and then setting  *Y*  to a value that is consistent with that value of *X,*  we can’t do it the other way around: once we have set  *Y* to a particular value, this determines a value for  *X*; we cannot coherently set *X* to some different value. ) Second, note that conditional irrelevance is much stronger than what is described in the philosophical literature as “multiple realizability”. Taken literally, the latter requires only that *some* different values of the same or different micro-variable(s) realize the same value of a macro-variable. Conditional irrelevance requires that *all* variations at the micro-level consistent with the value of the macro-variables that such variations realize make no difference to *E*. [↑](#footnote-ref-40)
41. [↑](#footnote-ref-41)
42. Note that, as with **(P\*),** relevance and irrelevance are always defined relative to an effect or explanandum *E*. *Y* may be irrelevant to *E* conditional on *X* but this may not be true for some alternative explanandum *E\**. As noted in Woodward, forthcoming b, forthcoming f, it is also possible to relax the requirement of complete conditional causal independence in various ways. For example, *Y* may be conditionally causal independent of *E*, given *X* for all but a very small set of values of *Y* (a set of measure zero.) Or *Y* may be conditionally independent of *E* given *X* for all values of *Y* within a certain large range, although not all such values. In this second case, when the actual values of *X* fall within this range, this may legitimate the use of *Y* rather than *X* to explain *E*. [↑](#footnote-ref-42)
43. An illustration: Suppose subjects are given data about the pecking behavior of a number of pigeons which shows that the pigeons peck in response to targets with various particular shades of red and not in response to colors that are non-red. However, they are not told the pigeons peck in response to red and only red targets. Do they exhibit a preference for causal claims formulated in terms of the cause variable *RED?* (One but not the only form such a preference might take is higher strength assignments to *RED* causes *PECKING* than to *SCARLET* causes *PECKING.*) Do subjects spontaneously discover that the *RED* variable is the most appropriate one? (Presumably the variable *RED* is *“*available” to them in the sense that they have this word in their vocabulary. We might also ask about cases in which the variable which will best satisfy **P** is a new, unfamiliar variable which they need to discover.) Relatedly we might also ask about which variables and level of abstraction guides their willingness to generalize about new situations. For example, suppose that in the set up just described, subjects are not explicitly encouraged to formulate a generalization containing the variable *RED* but are instead told that the pigeon has been presented with a new target that is a shade of red that has not been previously seen—e.g., maroon. Will subjects judge that the presentation of this target will cause the pigeon to peck and will they describe the redness of the target as the cause of the pecking? (This would be another form that a preference for *RED* might take) [↑](#footnote-ref-43)
44. At the risk of belaboring a point made earlier, again note the difference between asking, on the one hand, whether proportionality is part of the “nature of causation” or whether it is part of “our concept” of causation and, on the other hand, asking instead whether there is a normative rationale for preferring causal claims that better satisfy proportionality. Again issues of the latter sort tend to be more tractable. [↑](#footnote-ref-44)