Constraints on Constraints: 
Surveying the Epigenetic Landscape

FRANK C. KEIL
Cornell University

It is now relatively commonplace to advocate the need for some sorts of constraints on learning and knowledge acquisition. The critical issues to cognitive science concern the sorts of constraints that are able to best model various phenomena of learning and development. Four types of constraints on learning are proposed to be used as an interpretative framework within which to: 1. Better understand the nature of current research; 2. Allow the exploration of alternative models of learning related phenomena, and 3. See more clearly needs for further research. Superficially similar learning phenomena can be modeled by very different configurations of underlying constraints with strong implications for the sorts of representational states that are involved. Each of the five papers in this issue (Brown, Gelman, Markman, Newport, and Spelke) is considered in terms of the configuration of constraints after which each author intends to model their phenomena and in terms of alternate configurations. The papers are construed as illustrating a diverse set of models of how constraints might guide learning, and while the evidence generally favors the configurations suggested by the authors, in each case alternative models are possible and motivate quite specific future research questions.

More broadly, it is suggested that asking detailed questions about the sorts of constraints types that could potentially model complex cases of natural knowledge acquisition helps motivate fundamental questions about learning and the nature of knowledge and that the five papers in this issue are superb examples of how adopting this kind of perspective has been fruitful research orientation.

Preparation of this paper and some of the research described herein was supported by NSF grant BNS83-18076 and NIH grant 1-RO1-HD23822-01 to the author. Much thanks to Jim Greeno, Doug Medin and Rochel Gelman for helpful comments on early drafts of this paper and to all of the authors in this issue for reactions to my comments. Several of the papers in this issue have gone through a great deal of evolution over the course of when I wrote my initial commentary over two years ago. I have tried to keep pace with these changes as much as possible, but a little insight into that process may help readers understand why I sometimes go on about issues that now are only lightly mentioned in some of the current versions of these papers.

Correspondence and requests for reprints should be sent to Frank Keil, Department of Psychology, Cornell University, Ithaca, NY 14853.
I. Classifying Constraints

If pushed hard enough, almost all researchers from the staunchest empiricists to the most ardent nativists will ultimately concede the need for some sorts of constraints on learning. The arguments made by philosophers (e.g., Pierce, 1931–1935; Quine, 1951, 1963), information theorists (Watanabe, 1969) and linguists (Chomsky, 1980) demonstrate that, from the viewpoint of logically specified properties, no known learning procedure can succeed at learning a fixed set of categories or rules without a priori constraints. Explicit mention of the need for such constraints is less common, however, and many researchers seem to act as if they are irrelevant. They may lurk as necessary presuppositions in almost all research programs; but they are often so strongly presupposed that they are neglected in analyses of learning. The papers in this issue of Cognitive Science illustrate the central importance of being more explicit about constraints and how such explicitness leads to highly specific questions, and answers, about the acquisition and representation of knowledge.

Arguments for the inevitability of constraints can take many forms, several of which are seen in the papers in this issue and in Gelman’s opening remarks. Elsewhere (Keil, 1981), I have made the argument in terms of the problem of induction (see also Markman, this issue; and Markman, 1989). Roughly a century ago, Peirce (collectively published, 1931–35) eloquently illustrated how learning through induction was doomed in an organism that did not possess a priori biases on hypothesis generation; biases that had acquired an affinity with environmental regularities through the process of natural selection. For virtually any naturalistic learning situation, it is easy to formally demonstrate the indefinitely large number of different hypotheses that can be made over a set of data. Somewhat less intuitive, but equally important, is the demonstration that there is no absolute metric of simplicity that can be used to order those hypotheses. There are, of course, simplicity metrics relative to the specific endowments of each organism, but such metrics are embodiments of the constraints on learning specific to that organism. Perhaps the simplest demonstration of the need for constraints on induction can be seen in attempts to discover the function describing a set of points in a $x,y$ coordinate system. For any set of points, there is an indefinitely large number of possible functions describing lines that can be drawn through those points. New data points may exclude large sets of possible functions, but there always remains an indefinitely large set of possible candidates. Again, one is tempted to argue that some functions are intrinsically simpler to others, but such claims are notoriously hard to substantiate.

Given these sorts of in principle arguments, the broadest possible definition of constraints would simply be any factors intrinsic to a learner that result in a nonrandom selection of the logically possible characterizations of an informational pattern. It is perhaps more intuitive to state the definition in terms of limits on sampling of a hypothesis space, but in the most general form, one wants to account for constraints on the acquisition of mental states not normally considered hypotheses.

Although more limited characterizations of all constraints are desirable, it is difficult to achieve consensus on a single formulation. There is, for example, the intuition that biases away from nonrandomness arising from general physiological properties of an organism (e.g., the need for sleep) should not have the same status as those giving rise, for example, to categorical perception; but making this distinction straightforward and reliable is much more of a challenge. Perhaps the most practical way of reducing the scope of the term is to look at the range of normal uses in descriptions of empirical phenomena.

There are other ways of talking about constraints, such as those outside of the learner. Thus, in many areas of artificial intelligence research, constraints are discussed as properties of a domain, problem, or task. The learner, or problem solver, is seen as trying to discover those constraints so as to be able to either solve the problem more efficiently, or sometimes solve it all. For example, in cryptarithmetic problems, if a problem solver notices parity as a constraint, many possible “hypotheses” can be ruled out (Greeno, personal communication, 1989). Such formal properties of a problem space, if understood, will greatly aid a learner; but they are not considered cognitive constraints until they are internalized, either explicitly, or implicitly, by the learner. This is not to diminish their crucial importance in a full account of problem solving and learning, but it is critical to keep these senses of constraints distinct.

The contrast between external and internal, cognitive, constraints is clearly seen in architecture. There are clearly many important external constraints that limit an architect’s “hypothesis space” of possible buildings. These would include physical constraints imposed by strengths of materials, the load bearing capacities of various geometric configurations as well as constraints imposed by budgets, the size of humans, the nature of air flow, and so on; and clearly an excellent architect should have knowledge of, or at least immediate access to those with knowledge of those constraints. Without such knowledge, much effort would be wasted in designing structures that could not stand or be afforded (witness the initial attempts at design and construction of the Sydney opera house). These “task” constraints, however, can be importantly different from constraints that guide the generation of general forms and patterns of layout. Consider, for example, some partners in large firms who sketch broad patterns and then turn them over to more junior staff and engineers to fit those patterns within allowable physical and financial constraints (sometimes to have them returned as impossible). This contrast is not to say that internal and external constraints are unrelated. Often, expertise in a domain might consist of acquiring a rich knowledge such that one’s acquired domain-specific constraints capture
critical relations in the task. Similarly some innate constraints, for example on the perception of objects, might reflect the evolutionary internalization of certain geometric and physical constraints on objects. (See Shepard, 1984 for a discussion of such possibilities).

Even if there is at least a tacit agreement by all researchers on the need for learner constraints, there are profound and often passionate disagreements about their nature. Much of the controversy can be reduced to debates about the innateness and domain specificity of constraints. There are other important contrasts as well such as whether constraints are stated symbolically or subsymbolically, or whether they are stated over process versus structure, or whether the constraints are mild probabilistic biases, or powerful strutures against the formation of certain classes of representations; and some striking relations between these contrasts and those of domain specificity and innateness are mentioned below; but these two dimensions are surprisingly effective in helping us understand many fundamental controversies. Thus, many of the disagreements about the nature of learner constraints can be reduced to questions about which of the four cells are emphasized in the $2 \times 2$ matrix shown in Figure 1.

The matrix certainly oversimplifies the nature of both contrasts. There have been many heated discussions about when, if ever, the term “innate” can be used (e.g., Johnston, 1988; Lehrman, 1953; Lerner, 1983; Oyama, 1985). Development is obviously an interactional process between the organism and the environment; and this fact has led some to claim that a particular structure or concept cannot be innate since it is always a product of such an interaction. This concern is less applicable to constraints, however, because constraints can be construed as intrinsically interactional in nature: they are constraints on what sorts of knowledge representations an organism will construct given a range of environments. They are not statements about some knowledge structure as “built in.” The recognition of the interactional nature of constraints is nicely illustrated in all of the papers in this issue.

In Chomsky’s terms (1975), we can think of constraints as partially embodying a function from environments to mental representations. Constraints describe how the array of possible environments becomes mapped onto the array of possible mental representations. If, for example, they are powerful constraints, they result in a very strongly convergent mapping such that a very wide range of dissimilar environments yields an outcome set of mental structures that are all highly similar to each other. When such convergences are very powerful and the endstate representations are extremely similar across all but the most bizarre environments, it is common to say that the representations themselves are innate. To deny that some constraints themselves are innate is to deny that there are any fixed principles governing how classes of organisms interact with their environments. At the same time, not all constraints on learning are predetermined; acquired knowledge can come to exert its own influences on future learning with wide differences across individuals as a function of expertise and culture. These represent the opposite side of the innateness contrast. They specify an organism/environment interaction as well, but they emerge with development in a way that is not maturational, but which is highly sensitive to particular kinds of environments.

The domain generality contrast is oversimplified since it is clearly a continuum that is artificially dichotomized in the table. Thus, although some constraints might apply to tightly bounded bodies of knowledge and others might apply with complete generality to all kinds of knowledge, still others might well apply in an intermediate fashion to some, but not all domains. In addition, the constraints might be influential to differing degrees in different domains.

The notion of domain varies considerably across researchers. In some cases, most notably with those demarcated by innate constraints, the domains cover very broad areas of cognitive competency such as the representation of space or physical objects. In other cases, the domains may be locally circumscribed bodies of expertise. The critical common factor in all cases is that domain specific constraints are predicated on specific sorts of knowledge types and do not blindly constrain any possible input to learning. Thus, the working definition of domain is in terms of patterns of learning. If, for example, restrictions on possible hypotheses are unique to a specific body of the knowledge, that knowledge is considered a domain. Attempts to define domains independent of the pattern of learning are vastly more difficult since there would seem to be infinitely many possible domains that could be formally specified as sets of features and relations.

Domain general constraints are not purely indifferent to inputs; but they do constrain all inputs in the same way with the possible consequence of thus filtering out some classes of input because they are hard to process or encode under those general constraints. If they happen to result in different types of outputs for different classes of inputs, it is only because those classes
have properties that interact differently with the constraints in question. This can result in illusory domain specificity.

Domain general constraints are usually constraints on process with any resulting constraints on knowledge as derivative products of those process constraints, while domain specific constraints are usually stated over specific knowledge structures. It might therefore seem that any structural constraint on development is domain specific; but except for somewhat unusual senses of "structural," this may not be correct. This is perhaps best illustrated by recent developments in connectionism where constraints are posited that operate on what are presumably completely domain general bodies of knowledge. Thus, restrictions on the optimal number and on the nature of hidden units (e.g., Hinton, in press) would seem to refer to specific structural configurations; but they are domain general because learning in any domain whatsoever must honor those constraints. They don't operate only for inputs in one bounded area of knowledge and not for isomorphic configurations of inputs in a different domain.

Consequently, the relation between domain specificity and structurally stated constraints is not logically necessary. Moreover, structure/process distinctions are notoriously difficult to make (Newell, 1972). Nonetheless, the correlation is a strong one and it heavily influences researchers who adopt one bias versus another. Medin, Wattenmaker and Michalski (1987), for example further elaborate on this correlation between what they call process and product constraints and domain generality versus specificity. It is a fascinating correlation that is yet to be fully explained; a partial reason may be that, although domain-general constraints can easily be stated over either process or structure, it is more difficult to have constraints applying uniquely to a domain without at least some partial structural specification of that domain.

There are two other correlations with domain generality: degree of explicitness of constraints and the power of the constraints. There seems to be a general trend for domain-general constraints to be less explicit, often less stateable symbolically. The clearest case of this would be in some connectionist models that explicitly wish to eliminate symbolic levels of processing while retaining domain-general constraints at the "subsymbolic" level (Smolensky, 1983). Domain-general constraints are also usually less powerful in that they seem to narrow down the class of options less dramatically.

One possible reason for this tendency is that, although an extensive narrowing is clearly needed for learning to succeed, it is accomplished in a much too heavy handed a manner at the domain-general level, running roughshod over important differences in a world of heterogeneous information.

This $2 \times 2$ matrix should therefore be construed as a kind of interpretative framework within which to compare and contrast current theories of cognitive development rather than as a theory in itself. In this essay, I consider the consequences of different emphases on these four cells for theories of cognitive development. These emphases are then illustrated more concretely by considering how Brown, Gelman, Markman, Newport, & Spelke would probably fill out the matrix and why. In addition, I will discuss the extent to which alternative patterns of emphases might also be compatible with each of these researcher's findings.

II. SOME DIFFERENT THEORETICAL PERSPECTIVES ON DEVELOPMENT

The full range of views of how constraints guide cognitive development can be best appreciated by first considering some extreme perspectives and then moving to more moderate accounts. None of the articles in this issue are clearly allied with either of the fully extreme views, but such extremes serve as important anchor points within which to view other more moderate positions and, indeed, why they are moderate.

Extreme Nativism

The most extreme nativist view, shown in Figure 2(a) envisions as much development as possible as occurring in the Domain Specific Innate cell. In addition, each cognitive domain has its own unique structural constraints as well as having its own knowledge acquisition device (shown in Figure 2(a) as KAD's). Note that the knowledge acquisition device and the constraints might each exert different sorts of constraints on the knowledge output. Thus, the device itself might have certain unique properties governing the timing of input strings, how many are held in a memory buffer before an induction is made, and what sorts of confidence levels to use; while the structural constraints might result in a weighting preference for classes of output knowledge structures from such a device. A different approach might try to have all the constraints emerge from unique domain specific processing properties of the KAD. Finally, the uniqueness may lie solely in the structural constraints, with the processing component of KAD being more generic across domains. (A version of this last case is discussed below as a more moderate account.)

Under this extreme nativist view, learning is minimized. When it does exist, it consists in the filling in of slots in a richly structured innate conceptual system for each of several domains (or perhaps as the setting of parameters, (Chomsky, 1988). When forced, such a view concedes a little activity in each of the other cells but tries to keep it to a minimum. The environment may help trigger the unfolding of complex cognitive structure but does not play much of a role itself. In Fodor's (e.g., 1981) "triggering" view of concept development, concepts are not really learned at all; rather they are simply brought from a latent to an active form. Moreover, Fodor argues that the
<table>
<thead>
<tr>
<th>Innate</th>
<th>Acquired</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain Specific</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>local KAD's structural constraints of the richest sort</td>
</tr>
<tr>
<td><strong>Domain General</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \emptyset )</td>
</tr>
</tbody>
</table>

Figure 2 (a). An extreme nativist perspective

<table>
<thead>
<tr>
<th>Innate</th>
<th>Acquired</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain Specific</strong></td>
<td></td>
</tr>
<tr>
<td>only those that are caused by the sensory transducers and other &quot;qualia&quot;</td>
<td>parasitic off of DGI constraints and same in all domains</td>
</tr>
<tr>
<td><strong>Domain General</strong></td>
<td></td>
</tr>
<tr>
<td>fewest possible general learning procedures</td>
<td>?</td>
</tr>
</tbody>
</table>

Figure 2 (b). An extreme empiricist perspective

<table>
<thead>
<tr>
<th>Innate</th>
<th>Acquired</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain Specific</strong></td>
<td></td>
</tr>
<tr>
<td>local skeletal frameworks laid down by sets of constraints</td>
<td></td>
</tr>
<tr>
<td><strong>Domain General</strong></td>
<td></td>
</tr>
<tr>
<td>varies: from fewest possible general learning procedures to many heuristics, biases, etc.</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2 (c). A moderate nativist perspective

<table>
<thead>
<tr>
<th>Innate</th>
<th>Acquired</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Domain Specific</strong></td>
<td></td>
</tr>
<tr>
<td>constraints imposed by sensation and perception; and by more cognitive fragments and vestiges, often only accidently correlating with domains</td>
<td></td>
</tr>
<tr>
<td><strong>Domain General</strong></td>
<td></td>
</tr>
<tr>
<td>general learning procedures, heuristics, biases</td>
<td>?</td>
</tr>
</tbody>
</table>

Figure 2 (d). A moderate empiricist perspective
environmental events that occasion the manifestation of a concept may bear little or no relation to the represented concept. It is one of the hallmarks of triggering that such a correspondence is only optional. This point of view has its closest affinity with a number of older ethological models of "learning" (see Johnston, 1988 for an array of such views). A somewhat different extreme nativism view is modeled on the development of immune responses, and is discussed below.

Extreme Empiricism
In the extreme empiricist view Figure 2(b), the major activity in cognitive development occurs in the Domain General Acquired cell, in which, ideally, all learning is governed by the smallest number of possible principles. Perhaps the most archtypical proponent of such a point of view was Hull (1943), who envisioned reducing learning to a very small set of axiom-like principles out of which all patterns of learning could be modeled. Hull explicitly compared his program to the axiomatic structure of classical mechanics. Although Hull's program has waned in popularity, there are many current substitutes that espouse similar points of view. Many learning theorists view themselves as cognitive psychologists who are fundamentally concerned with internal representations and information processing but who are also biased towards models utilizing the most domain general constraints possible. (See Anderson, 1983 for a particularly clear statement of this view.) The sorts of general laws more currently in vogue would include: prototype abstraction, the Rescorla-Wagner law and its connectionist variants, and some general decision biases.

In this extreme empiricist view, the Domain Specific Innate cell is nearly empty containing only those constraints imposed by properties of the sensory transducers, such that visual input might result in centrally represented information that is formatted somewhat differently from that which is auditorily inputted; although, in general, such accounts try to minimize these effects as well, assuming convergence on a common representational format as information flows to more and more central levels of processing.

There are a variety of subviews on the Domain Specific Acquired or, what I call the "expertise cell." In Figure 2(b), all of the expertise domains are structured in the same way and constrain learning in the same way as the Domain General Innate constraints that gave rise to them; thus, once one understands how expertise constrains learning in one domain, one knows how it is done in all of them. Such constraints are, in fact only domain specific in that they occur on a domain by domain basis and usually show little or no transfer. They are domain general in the sense that once expertise is attained it is constrained in the same way in all domains.

In addition to these extremes there are several more moderate points of view, two of which simply emphasize each of the other remaining cells, and others of which simultaneously emphasize more than one cell or talk about interactions between them. It is assumed that there are no serious models of knowledge acquisition in which the only constraints are of the acquired sort, whether they be domain specific or domain general; consequently, the remaining models all involve more than one cell at a time as well as interactions over time.

Moderate Nativism
A more moderate nativist viewpoint also has Domain Specific Innate constraints but they are not as complete as for the more extreme view. Instead they set up more of a framework or skeleton within which cognitive development proceeds (see also Gelman's remarks in this issue on "skeletal principles"); they talk less about triggering, and have more of a tendency to see the environmental inputs as having a sensible structural mapping onto endstate representations. Thus, the knowledge structures in each domain must honor these constraints, but the constraints still allow for considerable developmental change and growth. This view also acknowledges a number of Domain General Innate constraints on learning (often needed to explain knowledge growth) and usually embraces Domain Specific Acquired constraints as well, which will also be called expertise constraints.

The domains for Domain Specific Acquired and for Domain Specific Innate constraints are very different. There are by most accounts only a dozen or two of the Domain Specific Innate domains, corresponding to broad areas of cognition that have presumably been important to humans in the course of evolution and different enough from each other in terms of the natural laws describing them that different sets of constraints are optimal for learning each. These domains might include knowledge of spatial layout, naive mechanics, naive theories of persons or "folk psychology," aspects of language (e.g., syntax and phonology), moral judgments, and so on. Moreover, the constraints corresponding to these domains tend to be more central in nature and less sensory and perceptual. As we have seen, even the staunch empiricist might acknowledge some domain specific constraints at the levels of the sensory transducers, but quickly wants to abandon them at any more central levels.

In contrast with the few innate specific domains, there may be an unlimited set of acquired expertise specific domains ranging from knowledge of chess to knowledge of antique chairs, really about as many as there are ways in which one person can be more specialized than another. In addition, it might be claimed that each of these areas of expertise exerts unique constraints on how new knowledge is acquired in that domain. Such a claim assumes that, as one acquires expertise, there is an internal structure to the expertise that is uniquely different from other sorts and which uniquely constrains further learning in that domain. Consequently, if one wants to know
what is constraining the hypotheses for an expert in a domain, one has to know the domain in detail.

This is not the only possible view of expertise, as described above in the extreme empiricist case, it may also be formally structured in the same manner regardless of content and exert the same sorts of constraints on learning. In such a case, expert knowledge has the same sort of "look" or distinguishing characteristics from novice knowledge in just about any domain. Local knowledge could still help learning in only one domain, but the way in which it helped would be the same in all cases. Thus, shifts to an expert "format" might occur on a domain specific basis, but the nature of that format might be domain general.

The representational status of expertise is currently a matter of considerable debate; one that has potentially dramatic consequences for how to conduct future research. For example, if the structure is unique to each local area of expertise, the principles of learning might have to be uncovered empirically in each and every distinguishable domain; and researchers will have to be cognitive "naturalists" documenting the special patterns of learning associated with each type of present and future expertise. Such a view is probably committed to the notion that such patterns of learning closely reflect the structure of an extraordinarily heterogeneous and diverse set of patterns of learnable information in the world.

Activity in the Domain General Acquired (DGA) often suggests the acquisition of new ways of learning and representing knowledge that sweep across virtually all domains of cognition. The grand stage theories of developmental psychology such as those of Piaget, Bruner, and Vygotsky are compatible with such an account although there often seems to be some influence from other cells, particularly from the Domain General Innate (DGI) one where there might be maturation of a predetermined competence. (e.g., Piaget might be considered an extreme empiricist if all DGS constraints directly and inevitably unfolded from DGI ones). This cell contains a question mark because one of the greatest single controversies of current work in cognitive development is to ask whether any domain general constraints of this sort emerge that result in new cross domain mental organizations and processes. There has been an extensive series of studies over the last 15 years or so empirically challenging the existence of such domain general changes in manner of representation and computation. (see Gelman and Baillargeon, 1983 for a summary). Later in this essay, I will suggest that some apparent DGA constraints are really caused by special classes of domain specific ones.

A Moderate Empiricist View

The moderate empiricist accounts usually allows for some central fragmentary constraints or biases that are domain specific; but see them as more or less random accidents of evolution or vestiges of an earlier process such that they fail to coalesce and define coherent domains. For example, some vestiges from early spatial abilities might haphazardly have been incorporated into restrictions on syntax, and so on (Shepard, 1975). The moderate empiricist view might also acknowledge expertise specific constraints on learning and structuring of knowledge and the range of possible views here might be quite similar to those for the moderate nativist. The domain general acquired cell is again a puzzle with the moderate empiricists perhaps putting more of an emphasis on the acquisition of general metatrategies that might be seen as domain general acquired knowledge (e.g., "rehearse and practice what you are trying to learn"). The domain general innate cell would allow for a much wider repertoire of biases, and heuristics than would be seen for the extreme empiricists. Many of the biases discussed by Kahneman, Slovic, and Tversky (1982), such as neglect of base rate information, might fit into this category. Thus, the moderate nativist and moderate empiricist views may differ little on the roles of constraints in the acquired cells but differ much more strongly with respect to the innate ones.

III. TYPES OF CONSTRAINTS AND TYPES OF REPRESENTATIONAL CHANGE

In addition to showing how different emphases in the 2 x 2 matrix (Figure 1) can have contrasting affinities with general theoretical perspectives in developmental psychology, these matrices can be used to explore more specific accounts of how knowledge structures might change with development. This is illustrated in Figure 3 (a–d). In 3(a), we have a classical learning theory account, which stands for one version of an extreme empiricist view in which there is only differentiation and elaboration of the same structural type, such as an associationist net. In case 3(b), global stage views are modeled. Here the most obvious account would have a great deal of activity with the Domain General Acquired cell, some activity with the Domain General Innate cell, and perhaps a small level of activity in the Domain Specific Innate cell representing some primitive reflexes that seem to be adapted for only restricted types of input (e.g., the sucking reflex). These views would be either moderate or extreme empiricist as a function of postulated mechanisms in the Domain Specific Innate cell. In case 3(c), a version of the moderate nativist account, knowledge structures differentiate according to invariant constraints, but no qualitative restructuring occurs. Finally, in case 1d, which might model all development as novice/expert shifts, there is an increasing movement away from domain general innate constraints to domain specific acquired ones. In this case, I have shown a particular pattern suggested by Quine (1969), with moderate empiricist leanings. These combinations hardly exhaust the possible models of representational change, but they do at least illustrate the range of possibilities and how emphasis in different cells translate into inferences about the representations themselves. It
may be that each of these can be found in relatively pure forms in various situations and content domains, although I doubt one will ever find case 3(b).

In sum, the point of this introduction to types of constraints and how they have consequences for patterns of representational change is to provide a context within which to evaluate each of the five papers in this issue, which is the topic to which I now turn.

IV. RELATIONS TO SPELKE, GELMAN, MARKMAN, AND NEWPORT

How do the five papers in this issue fit into this 2 x 2 matrix? Perhaps the only completely clear point is that they do not represent either extreme nativism or extreme empiricism. Neither of these extremes are "strawpeople"; they are both currently represented by eloquent and forceful advocates. For example, Piatelli-Palmarini (1989) argues that there is rarely any real learning; instead, in a manner analogous to the development of immune responses, there is a selection of preexisting mental structures. Experience may serve to bring some structures into the foreground, but there is no learning. As extreme empiricists, one might again cite Anderson (1983).

Conceding that the five authors in this issue are not representatives of either extreme, we are confronted with trying to discover which of the more moderate views is suggested by their data. In several cases, it may not yet be clear in which cell or cells the predominant constraints reside, but resolution of such ambiguities will carry important implications for further theorizing about development in each case. It is not always easy to pigeonhole the array of views; I will therefore examine each researcher in turn, first in terms of how she would apparently like her 2 x 2 table to be filled out, and second, in terms of other possibilities that would seem to fit with her empirical findings.

Spelke

Spelke is the closest to an advocate of pure DSI constraints in her account of the infant's apprehension of objects and, in the nativist sense, since she wants the constraints not just to be on low level perceptual processes or even on quasi-low level Fodorian modules with their encapsulation, lack of cognitive penetrability, and so on (Fodor, 1983). Instead, she has advocated elsewhere that constraints emerge from theories at the "highest level of conceptual structure," a naive theory of objects (Spelke, 1988). She, therefore, does not agree with Quine's scepticism (1953, 1963) about there being no innate constraints on theories. In this paper her discussion has turned mostly to perceptual principles, but it remains staunchly opposed to the knowledge of objects being "bootstrappable" from constraints arising out of basic Gestalt laws or a simple minimum principle. She suggests instead that appeals to such mechanisms cannot account for the rich knowledge that infants have about objects. By rejecting not only the radical empiricists such as Berkeley but also the Gestalt theorists such as Wertheimer (1958) and Koffka (1935), she is arguing for constraints at more central levels of representation. Equally important is her argument for domain specificity, when she suggests that although young infants do not use such principles as good continuation to perceive object unity and boundaries, they do easily detect such things as misaligned contours in arrays with aligned contours, apparently using such principles in contexts not related to object perception. Only later can they be fully used to perceive objects. The constraints moreover seem to be on knowledge structures, not on process. Thus, the DSI constraints heavily predominate; and they give rise to such beliefs as that: objects are cohesive, bounded, and do not act on each other at a distance. Moreover, these beliefs are interdependent and interconnected and give rise to other beliefs such as that objects move on connected paths and do not pass through each other, forming a kind of coherence that fosters the attribution of an intuitive theory of objects to young infants.

Is this the only possible account of what is going on? The consistency across Spelke's subjects and the systematicity of their knowledge about objects as compared to other things would seem to support a focus on the DSI cell. One might wish to explain the acquisition of physical object knowledge in terms of domain general constraints; but one is then confronted with the problem in explaining the infant's rapid progress of learning about some aspect of objects as compared to other aspects that seem no more formally complex and no less common, such as the effect of gravity (Maccamber, Spelke, & Keil, 1988). Gravity presupposes some aspects of objects but not all, hence it is still of interest that infants "know" about solidity long before they know about gravity.

One might try to acknowledge a few domain specific biases, but say they are more perceptual and sensory; perhaps there are principles governing dynamic properties of occlusion and edge detection that isolate objects from other sorts of things. Then, by more general principles of induction, the infant develops a naive physics. As noted, however, Spelke argues against this in her demonstrations that infants, although sensitive to some Gestalt laws that might pick out objects, fails to use them in the context of dealing with objects.

Although Spelke may put the largest emphasis on the DSI cell; she clearly also allows for learning which is guided by more general principles. Thus, infants develop new ways of perceiving objects in accord with Gestalt relationships. Contrary to the Gestalt theorists, Spelke explicitly suggests that such relations can be learned by building on the principles that she advocates. She tells us little about the learning process itself, other than that it must occur; but her account is fully compatible with those learning mechanisms being domain general.
There is a temptation to interpret Spelke's early principles as evidence for infants' possession of intuitive theories of physical objects (Spelke, 1988). Although I am sympathetic with such an interpretation, I am also concerned about what should constitute evidence for the presence of a theory, an issue important not just to Spelke's paper but to several others in this issue as well. Spelke (1988) suggests that theory tells you what sorts of things there are, that is, what are the meaningful cuts to make in the world. But theories are also said to be connected sets of beliefs. If an entity merely behaves in such a way that only certain inputs are "meaningful" to it, we might not want to attribute theory to it: at least without unpacking the meaning "meaningful" in a way that does not circularly rely on theory. I am reluctant to grant spiders intuitive theories of the mechanisms of physical lattices like webs, even though their behavior displays a precise honoring of such principles. Similarly, cockroaches and other cognitively "simple" creatures also seem to pick out objects and follow their trajectories and the like, yet one cringes at calling them object theorists. They don't, I suspect, anticipate collisions for screened events and make inferences about numbers of objects involved based on seen and unseen trajectories; but in other respects, they do seem to think there are objects "out there."

Thus, there are some unresolved questions concerning what constitutes evidence for theory versus less belief-laden systems of representation. It would seem that coherence and systematicity are necessary parts of a theory, but they may not be sufficient. The problem is in trying to specify what else is needed. Gentner (personal communication, 1988) suggests that theories may also require explicit recognition of and correction of inconsistencies and contradictions. Gentner speculates that perhaps mental models should be construed as not having to honor such requirements. If so, then we probably wouldn't want to grant theories to nonprimates and the question of whether young infants have theories or merely mental models becomes a real one. Spiders and cockroaches clearly have representations in Gallistel's (1989) sense of being functioning isomorphisms that are computationally linked to environmental regularities; but invocation of theory would seem to call for something beyond a functioning isomorphism.

In sum, Spelke's word suggests a strong emphasis on the domain specific innate cell, with relatively little spillover to the other three cells. Constraints on theory building about physical objects may be extensive and specific from the start. Although the possibility is still present that all of this is induced through more general means, the onus seems to have shifted to those who wish to make such claims to provide empirical evidence for DGI constraints and against DSI ones.

Markman

Markman is carefully agnostic about whether her constraints are either domain specific or innate. This caution is fully appropriate given the current status of empirical findings; and is sometimes misunderstood by others who have interpreted Markman as proposing DSI constraints. This agnosticism, however, does not undermine the importance of ultimately knowing what sorts of constraints are at work; for such a determination can have a dramatic impact on additional predictions and future, more detailed, models. In fact, Markman's constraints are particularly interesting because they nicely illustrate the subtlety and complexity of the contrasts, especially the dimension of domain generality.

Linguistic labels suppress thematic thinking and promote taxonomic thinking. A class of things when referred to by a linguistic label is understood differently by the child than when it is not so labeled. Thus, the hypothesis space on what constitutes a word meaning may be constrained in a special way that does not extend to other sorts of nonlabeled concepts. This is an intriguing sense of domain specificity not yet considered, the domain is not restricted to be about any particular conceptual content domain such as number, or spatial concepts, or physics; we obviously have word meanings in all these domains, rather the constraint says that, once you know a concept is a word meaning, be sure to organize it in a way that you don't normally do for nonlinguistic concepts. In a sense it is a domain general principle of conceptual organization that is only invoked when one is learning in a particular domain, such as words. The taxonomic over thematic constraint is only domain specific because it narrows the hypothesis space in a special way for a well defined subset of all natural concepts, but it is also very different from naive physics because other than being marked by a label, the specific content of those concepts is nearly irrelevant.

One might try to argue that the taxonomic organization constraint is not domain specific at all. Markman's work appears to demonstrate how a word label strongly influences guesses about meaning and clearly illustrates how, without such labels, very different sets of judgements result. But perhaps other nonlinguistic devices could also result in a shift towards more taxonomic structure. Perhaps words are a special case of a more general class of "highlighting" devices that tell the child to look beyond the thematic and the superficial. Perhaps, but the problem is trying to come up with feasible accounts of other such devices. We know from Markman's work that simple deictic devices are not adequate. It would seem that words and words alone may promote certain constraints on the structures that are not seen elsewhere.

Markman's constraint on mutual exclusivity also appears to be special to the domain of word meanings. When a new label is used, children assume that its referents contrast with referents of other known labels with little or no sharing of referents across labels. This constraint implies that, when setting up categories of the world, if those categories are lexically labelled, one will show more of a tendency not to have them overlap. The structure of natural categories in the world presumably does not reinforce such a con-
strait; it is solely a consequence of the label (see Simon, 1962, however, on
the pervasiveness of hierarchical structure in nature). Notice again that this
is a special sort of constraint for which the content of the word meaning
may be relatively unimportant.

Are more domain general interpretations also possible here? Do word
labels uniquely promote mutual exclusivity? Perhaps there is a general cog-
nitive principle having to do with contrasting and sharpening, something to
the effect that when you are trying to acquire a concept at a certain level
of analysis, you attempt to maximally contrast to other related concepts in
the same multidimensional space. Basic level categories have been argued to
have such properties independent of language (e.g., Rosch, Mervis, Gray,
Johnson, & Boyes-Braem, 1976). Perhaps labels imply to the young child,
often wrongly, a basic as opposed to super or subordinate level of categori-
ization and the constraint on mutual exclusivity arises out of a more domain
general one concerning the basic level. This alternative could be addressed
by future studies that explicitly explore the power of the mutual exclusivity
constraint in the face of labels versus non-labels and different levels of cate-
gorization. See, however, Mandler & Bauer’s (1988) recent challenge to the
privileged role of the basic level in cognitive development.

Whether or not these constraints are uniquely triggered by words or more
broadly by other devices, once they are triggered, they are very different
from other sorts of DSI constraints in that, from that point forth, they are
indifferent to the conceptual domain involved. Moreover, we need not feel
so compelled to think of these constraints as innate. Perhaps the child has
discovered through observation that, whenever words are used, they seem
to refer to concepts with a more taxonomic-like structure. The constraints
of interest could be residing primarily in any one of three cells and possibly
even in the DGA one. This ambiguity should not be interpreted as saying
that it doesn’t matter what sorts of constraints are involved, however. Quite
the contrary; depending on which interplay among these cells is correct, very
different empirical patterns should fall out. For example, do children who
learn labels much earlier than normal, such as ASL, speakers show a corre-
sponding benefit for focusing on taxonomic relations? If so, then one would
be more inclined to think that there are DSI constraints that are specifically
triggered by linguistic labels and that the bias towards taxonomic relations
is not an outgrowth of more general patterns of cognitive development.

Markman’s paper also dramatically illustrates the ways in which con-
straints can fruitfully interact in development. In fact, one might take such
rich interactions as suggestions that the constraints are moving towards the
sorts of coherence indicative of domains. We see how each constraint on
its own can lead the child astray, but how they form a system of checks and
balances in their interaction. This is especially clear with the introduction of
the whole object bias. Thus, the taxonomic and whole object assumption
biases the child towards assuming labels refer to relatively homogeneous,
nonthematic classes of objects. This is a very useful bias for learning what
sorts of things are called and how they can be embedded in other sorts; but
as Markman points out, many language terms don’t refer to objects but
rather to properties and relations. Fully unrestrained combinations of the
taxonomic and whole object biases will blind a child to such relational and
property terms; and it is here that mutual exclusivity performs such a power-
ful role. In concert with the whole object and taxonomic biases, it ensures
that, once an object does have a label, the child will look for other sorts of
meanings for new terms that are applied to it.

I suspect that it is not merely coincidence that these three constraints
work together so well. The problem of word meaning acquisition may well
have motivated their emergence as an interconnected system. Moreover,
although it may never be possible to show such constraints in infants, I am
inclined to suspect that they are not acquired. They are interestingly depen-
dent on other earlier emerging clusters of constraints, most notably those
such as Spelke’s governing the isolation of whole objects; but this depen-
dence does not rule out the possibility that their emergence is also matura-
tionally predetermined. The point here, however, is not that Markman’s
constraints must be either innate or some special sense of domain specific; it
is that such a determination will bear directly on how they will work and on
our predictions about novel learning situations, such as ASL.

Newport
Newport presents converging forms of strong support for a critical period
for the acquisition of language. She does not, however, conclude as many in
the past have (e.g., Lenneberg, 1967) that this is a result of maturationally
guided DSI constraints. Newport offers an entirely different pattern of ex-
planations, one so clever and original that you then wonder where else it
might work. Remember that she notices an additional intriguing fact that
the child’s ability to learn language declines at just about the same time as
the child suddenly becomes a much more capable learner in many nonlin-
guistic domains. This linkage leads Newport to suggest that the responsible
constraints may not be specifically tailored for language per se but rather
that the general cognitive machinery may be limited in such a manner as to
make certain aspects of language especially salient. Newport proposes that
the general cognitive limitations of the young child may be critically helpful
for learning languages such as English and ASL but detrimental for many
other nonlinguistic types of information: her so called ‘‘less is more hypoth-
esis.’’ By way of analogy, she points to Turkewitz and Kenny (1982) who
suggest that the young infant’s limited acuity may make him or her tend to
focus more on relatively larger objects and not isolated features or detailed
patterns, thus perhaps giving him or her a very useful bias for later percep-
tual and cognitive development.
This is a fascinating notion for it is an embracing of domain general innate constraints governed by a maturational timetable; they can emerge at various points in development, thus in principle being able to explain global stage-like reorganizations in development and thereby mimicking a pattern thought to be caused by DGA constraints. It is perhaps the most developmental account of constraints of all the papers in this issue. One gains a kind of domain specific constraining here but only via domain general mechanism. There are no special constraints favoring a particular class of representations; instead, early general patterns of processing result in certain kinds of information being learned more easily; and then, as the general patterns of processing change, that filtered information can now constitute a coherent domain. Although, Newport raises the provocative possibility that DG constraints are causing critical period effects, this should not be construed as an advocacy for a completely domain general language acquisition device. Newport is carefully neutral on this point, allowing for an interaction between both DG and DS constraints (see pp. xx and xx of Newport).

Newport’s account may not be precisely right, of course, but it elegantly illustrates the complexity of trying to find out what sorts of constraints are involved and when. It should also be noted that a “less is more” (LIM) position, while perhaps suggestive of certain types of constraints, does not always entail a specific kind. For example, Newport mentions the developmental shift in ability to store items in short-term memory as one possible way of instantiating the LIM hypothesis for language. She notes that there is considerable controversy over the reasons for the change in STM, but argues that the LIM hypothesis is indifferent to this controversy. Newport is certainly right that the LIM hypothesis could exist regardless of such details, but such details are nonetheless important in understanding what sorts of constraints might be involved and how the LIM hypothesis would actually work. Thus, if the changing STM limitations are maturationally governed and apply indiscriminately to all information, they reflect DGI constraints of the sort Newport is provocatively suggesting. If, however, those changing limitations are consequences of emerging general metamemorial strategies, they might reflect DGA constraints, which could show some cross-cultural variations and perhaps generally more of a range across individuals. The changes could also be caused by emerging bodies of expertise in local areas and not general strategies, with the constraints thereby being of the DSA sort (Chi, 1988). Finally, it is possible that there is a maturationally guided change in STM capacity that is linked specifically to storage of linguistic information, suggesting DSI constraints. LIM allows a much easier way of understanding how DGI constraints might be causing critical period effects of language; but without other demonstrations of concomitant changes in nonlinguistic domains, we cannot automatically rule out DS alternatives. Different demonstrations would be required to rule out the DGA alterna-

tive. Newport is clearly sensitive to these unresolved issues and they should not detract from her compelling demonstration that the LIM hypothesis certainly does move one away from a strong DSI bias in explaining critical period effects in language acquisition. It is an excellent vehicle for understanding how DGI constraints could cause the effect, but it does not automatically entail them.

Newport’s LIM hypothesis also suggests possible reinterpretations of other accounts. For example, perhaps in Spelke’s case an early ability to learn about objects especially in terms of their dynamic properties stems from an immature perceptual system that is primarily sensitive to movement patterns and relatively large visual objects rather than detailed patterns. These sorts of DG processing constraints might be focusing the child on just those sorts of information most relevant to a naive physics, or at least providing some assistance in that direction. I have suggested above that Spelke’s findings are unlikely to be explained in such a way, but until careful models are built along these lines, the possibility remains. Perhaps most impressive about Newport’s work is that her model of constraints makes several counterintuitive predictions about future studies, the most dramatic of which is that one might be able to erase critical period effects by presenting linguistic information to adults that has been passed through a “kid filter” that yields just those linguistic elements salient to young children. Studies now underway in Newport’s lab suggest that such an outcome is a real possibility.

Brown
Brown discusses recent work on surface and deep similarity and systematicity, an area that has received extensive attention in recent years (cf. Gentner, 1983; Gentner & Toupin, 1988). These general contrasts alone could be results of domain general constraints, but Brown also embraces domain specific innate constraints in her discussion of infants’ causal theories and of the sorts of intuitive theories they might possess. She endorses the early emergence of the two classes of naive theories postulated by Carey (1985) and suggests that these in turn provide constraints on the search for causal explanations about physical and animate kinds. How then might we model this complex array of developmental patterns in terms of the 4 cells in our matrix? Brown’s account is in many ways the most explicitly eclectic in that she posits activity in all but the DGA cell and perhaps even there in terms of general metacognitive acquisitions and strategies. Her discussion of infants being predisposed towards particular patterns of causal explanations strongly implies that they have innate constraints on the formation of systems of explanatory beliefs in such realms as physical objects. Brown is careful, however, not to maintain that constraints on theories themselves are innate. By citing Leslie’s (1988) and Gelman’s (this issue) views, she acknowledges that infants’ theories may arise out of more perceptual sorts of biases.
Do Brown's data, however, require reference to innate domain specific constraints on physical causality? In fact, do they require any reference to the DSI cell at all? An answer will largely depend on how causality is encoded. It will also depend on the ways in which Brown sees explanation as a mechanism for going beyond the superficial and providing the child with deeper relations that allow transfer. Quine (1977), for example, certainly believed in the development of theoretical beliefs that restructured similarity relations and which even co-existed with the perceptual/phenomenal ones. Thus, there is a Quinean account which could apparently model Brown's notion of transfer by reference to deeper relations without any assumptions of innate DS theories. I doubt Quine's account can work, but for complicated reasons having to do with how theories might emerge; Brown's data on their own don't seem incompatible with a Quinean view. Fortunately, further work can help us to untangle these options. For example, one might examine cross-cultural patterns of the emergence of causal beliefs.

Brown also supports the idea of concepts as embedded in theories. Again this has been an increasingly popular notion in recent years (e.g., Carey, 1985; Murphy & Medin, 1985) and one that I have been increasingly stressing in my own work. (e.g., Keil, 1989); in fact, some have gone so far to complain that the claim has become a cliche (Atran, 1989). Brown also supports Medin and Ortony's (1989) view that surface, probabilistic, similarity spaces are useful because they are often diagnostic of deeper essences (see also Medin, Wattenmaker, & Hampson, 1988; Rips, 1989; S. Gelman, 1988; Keil, 1989 and others) and that we use them as fallback options when other information is not available. What does all this recent fuss about concepts and theories and assumptions of hidden essences entail for the sorts of constraints that guide knowledge acquisition? At a general level, the entailments are weak; for it would seem that models with very different emphases on types of constraints could discuss such phenomena. What is required are more specific commitments of what theories look like, how concepts are embedded in them and what it means to use “surface” information to look deeper.

My hunch (spelled out further in Keil, 1989) is that eventually we will see theories and causal explanation as even more important than most current accounts would claim. Natural concepts, at least, may not merely be associated with the theories that help illuminate them, they may not be able to exist without such theories in that there are no purely theory neutral similarity spaces in humans. Only in the most contrived artificial cases does it even makes sense to talk about concepts independent of the networks of relations in which they are embedded. For natural kinds, these relations may be largely causal/explanatory in nature. For other concepts, such as those of space, time, and number (see Gallistel, 1989), causal/explanatory relations are less relevant; nonetheless, the notion of an individual concept deriving much of its meaning from a broader network of relations may well still exist. It may be silly to even talk about the infant having concept of an object per se; that concept may emerge from the network of causal relations that yield a naive physics. Knowing what a physical object is may be knowing the sorts of causal relations it can enter into. (Incidentally, for this view of human natural concepts to go through, we probably have to argue that pigeon concepts for example (e.g., Hernstein, Loveland, & Cable, 1976) are of a fundamentally different sort just because they aren't embedded in theory. I think this is a fair move but not one that all would accede to. Pigeons seem to be purer Roscheans because, beyond limits sensory transducers and “belongingness” constraints (cf. Seligman, 1970), their concepts are organized purely around typicality distributions. Pigeons, however, will never acquire new knowledge that fundamentally alters their early similarity space because of theoretical insight.)

There are other sorts of constraints of the acquired expertise variety at work in Brown's studies as well, such as where the acquisition of any kind of rich causal structure in a domain enables transfer. Thus DSA constraints arising out of local knowledge systems can also predict patterns of transfer. Brown explicitly refers to DSA influences when she suggests that rapid transfer is often dependent on the differentiation of local theories in relevant domains.

One might be tempted to conclude from some of Brown's remarks that there are different sorts of causality, perhaps reflecting different innate causal biases. It may be wrong, however, to focus on causality in a domain specific way or perhaps as a separable aspect of cognition at all. Causal thinking may be a useful building block throughout many diverse sorts of conceptual structures and more like a basic vocabulary item in structural descriptions of many concepts rather than a separate domain. It may make no more sense to talk about a separate domain or domains of causal thinking than it does to talk about a disjunction operator independent of the logical calculus into which it enters. In fact, one can think of the causal relation as a special kind of logical operator. I suspect the tendency to look for causal explanatory relations is the most domain general of cognitive skills and that, although there may be DG constraints on the causal connective, most of the interesting restrictions on causal thinking may arise out of the contents in the local domains. Brown seems to support this view elsewhere in her paper.

Brown's impressive demonstrations that young children can go beyond surface similarity to engage in transfers based on deeper relational/causal knowledge do not on their own narrow down the probable class of constraints. They suggest several exciting possibilities, but further work looking at origins of the causal beliefs that mediate transfer in needed. These studies do argue strongly for a model in which almost all the possible sorts of con-
straints come into play at the same time; although there are also different degrees of influence over the course of learning as is seen in Brown's references elsewhere to the changing influence of Newell’s weak methods (e.g., Brown, 1986). An important implication of these mixed models of constraints is that the representations themselves are probably more heterogeneous in nature with different aspects reflecting different sorts of constraints, hence the increasingly popular view of concepts as blends of associative structure (the fall back information) and causal structure. (This view of concepts is discussed at length in Keil, 1989).

Brown’s paper nicely illustrates both the importance of a constraints perspective on learning at all age and the diversity of constraints types that may be simultaneously at work in many settings. Her arguments for domain specificity point out a special neglect of such constraints, especially in older children, but she also does not want to argue for the exclusive role of such constraints; rather that a more balanced perspective is needed. This should not be confused however with a call for unpredictable eclecticism (“it’s a little bit of everything”); Brown’s studies make crystal clear the importance of discovering exactly what sorts of constraints are at work and how they interact.

Gelman
In accord with the general line of argumentation for constraints, Gelman suggests that young children focus on some properties and relations in a task and not others because they are being guided by “skeletal versions” of different sets of principles, in this case, by principles for counting and for reasoning about causality. She further suggests that, since their knowledge in the domains is about structure, it can organize exploration of the environment for domain-relevant inputs. This is a clear statement of the importance of domain specific and presumably innate constraints. In her prior work on the child’s concept of number, Gelman offers strong arguments for DSI constraints; and I have long viewed that work as a paradigm case of domain specific innate constraints (Keil, 1981). Gelman’s elegant demonstrations of differences in preschooler’s inferences about mechanisms of movement between animate and inanimate, however, may not yet be at a point where commitments can be made about what sorts of constraints are most active.

For causality in general and for the causal factors influencing the movements of animates and inanimates in particular, the precise nature of the necessary constraints is still unclear. Gelman suggests that the children have principled concerns about the causal conditions for movement, principles that apparently demarcate a domain of causal thinking. She argues that her demonstrations of dramatically different responses to identical questions about the insides of superficially similar kinds are incompatible with a model invoking only domain general constraints. Gelman clearly also endorses a role for general constraints as can be seen in her reference to Rescorla & Wagner (1972), but she also wants there to be more domain specific biases as she argues that those general biases cannot be the only ones involved in concept acquisition. This raises the complicated question of how such domain specificity might be manifested; and, as has already been seen, there are many ways to fill in the DSI cell and to specify its interactions with other cells. Do these biases spring from constraints on a theory of biology, are they constraints from a separate domain of causality, or do they represent less coherent piecemeal and fragmentary biases that just happen to yield very different expectations about animates and inanimates, and around which coherent belief systems coalesce?

Gelman suggests causal thinking may be especially invoked when we search for a mechanism. This is presumably one reason for her making it domain specific; but except for a few artificial concepts where there are no implicit causal structures, it is in fact astonishing how pervasive causal thinking is, even to the point of creating illusory correlations (cf. Murphy & Medin, 1985). As I argued for Brown’s work, causal relations may pervade the representations of natural concepts. Moreover, to say that the search for mechanisms fosters causal thinking runs the risk of being circular without further noncausal specifications of what mechanism is. The two principles about movement that Gelman presents, the innards principle for animates and the externals principle for inanimates, are fascinating demonstrations of very young children’s abilities to think causally about even unseen forces; but perhaps the mechanisms in those accounts are arising simultaneously with causal explanation, not being fostered by it. Gelman is clearly sensitive to these issues and is wrestling with the question of how various causal principles differ from explanatory beliefs.

My emphasis on cause as possibly being pervasive for almost all natural concepts at all points in development is not the same as a stronger claim of “no cognition with causation.” There seem to be many aspects of our mental activities that do not require causal representations. These might include representations of spatial layouts, such as cognitive maps, representations of principles concerning number and language, and various strategies and heuristics. For concepts about natural phenomena, however, cause may be much more central, as well as for concepts about other kinds that have natural kind components. If, as mentioned in the discussion of Spelke, the concept of a physical object can only exist in terms of the roles objects play in an intuitive theory of mechanics; then concepts about any entities involving physical objects will involve some set of beliefs about causal relations and about entailments from those relations. This would be the minimal level of involvement for totally unfamiliar physical objects, such as novel artifacts; but in almost all cases much more elaborate sets of causal beliefs come into play as well. Cause makes many of the probabilistically afforded
cuits in world meaningful and is fuel that drives the inductive engine. In short, I am suggesting that it is a primitive in virtually all concepts about natural phenomena and even for many concepts about other kinds and only recedes completely for the purest of nominal kinds such as "triangle" and "prime number."

"Having a concept" should not be construed as knowing a static set of properties, features and frequencies. Instead, for most real world categories it fundamentally and centrally involves knowing a set of dynamic causal relations that help us understand why a cluster of properties have become segregated as an interactive unit and what processes continue to maintain that segregation and thereby continue the presence of the kind. Our concepts of markedly inanimate objects such as a cactus are much more dynamic than they seem because they involve an understanding of the ecological forces at work that make cactus's cluster of features a superb adaptation for its environment. The idea of static objects being represented in more dynamic terms has been argued for in the realm of perception as well. Freed, for example, (1987), has conducted a series of experiments demonstrating that the perception of static forms such as handwriting and photographs often involves an encoding of the dynamic movements involved either in creating the perceived form or in predicting its likely future movements. While the relations of perception of objects to concepts of kinds are quite indirect, both lines of work do serve to illustrate how mental representations of non-events might nonetheless have an event-like flavor to them.

One also cannot yet tell if the principles are the initial basis for demarcating animates and inanimates or are the product of such a distinction that arises on another basis. Does the major difference between animates and inanimates critically hinge on the two principles? Perhaps most three-year-olds also see their moving toys as mobile because of innards but at the same time do not think those toys are truly animate. If so, then the principles might be domain specific acquired heuristics arising out of perhaps more perceptual means of initially distinguishing the two kinds. The studies described by Gelman do illustrate that young children easily distinguish animates from inanimates, animals from machines, and that they see internal parts as more important for animals; but we don't yet know if the innards principle is the original primitive; or, if it is primitive, whether it might not even be initially collapsed with the external agent principle into a single more domain general one such as "Assume that all motions must have causes such that, if there are no observable external ones, there must be internal ones."

More broadly, Gelman's work illustrates the need for more systematic investigation of how expectations about different kinds emerge. In general, as with Brown's work, demonstrations of an early ability to go beyond surface similarity relations to see different structures "beneath" are suggestive of domain specific constraints of an innate sort; but literature documenting similar effects in adult novice-expert shifts, forces us to be cautious in making such attributions. Gelman's work documents impressive biases on the inferences preschoolers make about animates and inanimates and illustrates the central role that cause and mechanism can play in early concept development. It also illustrates the complexity of the problem and how much more work needs to be done in this area before we can unambiguously determine precisely what sorts of constraints are implicit in these biases.

**The Consequences of Contrasting Constraints**

Even a brief consideration of the research programs of Spelke, Markman, Newport, Brown and Gelman make it clear that many different senses of constraints are being invoked, senses invoking three cells of the matrix and possibly even the questionable fourth with domain general acquired constraints. Moreover, within each cell the constraints can be manifested in very different ways as was seen in the alternative construals of domain specificity, raising again the need for more formal consensuses on what constitutes a domain. Such distinctions may have important bearings on how we understand patterns of learning and in how we conduct our research. Depending on the sorts of constraints one thinks of at work, one may adopt very different research styles relying on different kinds of stimuli. Thus, a DGI theorist might look at process models of learning of completely artificial stimuli while a DS theorist will focus on structural descriptions on natural knowledge at all points in development, and possibly across cultures, for a domain that seems to be intuitively well bounded and formally different from others. Depending on our views, we often look only for main effects from our one perspective and, if we fail to take into account the interactions between the full variety of constraints, we may end up talking past each other. These papers nicely illustrate how several different types of constraints can be taken into account while trying to understand some phenomena. The critical issue is to see how all constraints fit together to form a coherent account of learning; for these papers do reveal how difficult it is to contain our accounts of any of these phenomena to just one cell in the matrix.

Virtually any instance of learning will be governed by DG constraints reflecting consequences of general processing patterns. But the influence of these patterns can range from total guidance for pure nonsense stimuli to little or none for natural domains in which DS constraints are heavily influential. There may only be one DG learning device; but when it cannot be supplemented by DS constraints it may be quite feeble, hence our consistently poor performance in learning tasks with pure nonsense stimuli. (Notice how hard we try to give meaning of any sort to such dismal stimuli, see for example Ericsson, Chase, & Faloon, 1980). The five papers in this issue represent a revolution that has occurred in cognitive development over the last several years as they demonstrate a marked shift away from DG constraints.
research and paradigms in which stimuli are made intentionally devoid of any naturalistic content. They have shown us the profound importance of looking at DS constraints of both the acquired and innate variety and the dangers of assuming too much for the DG ones. At the same time they show us the dangers of relying exclusively on domain specific constraints, realizing the need for more general learning mechanisms as well.

**SOME FINAL ISSUES**

In this article, I have tried to show how a simple partitioning of constraints on learning into four types allows one to see several alternative models of cognitive development for roughly the same class of phenomena. I have also emphasized that these alternative models are distinguishable by further empirical work that specifically explores how different sorts of constraints might have different effects on learning. Newport is perhaps the farthest along in this enterprise with her very explicit follow-up studies in which markedly different outcomes would result from different sorts of constraints. All the researchers in this issue, however, show a special sensitivity to these issues, which is why their work is among the most exciting in cognitive development and why it continues to yield exciting results and generate important new studies. It is also evident from all of these papers that several central questions still need a great deal more attention:

1. **What Forms Do Constraints Take Mentally and What Do They Yield?**

   We have seen across these papers a variety of ways of talking about constraints and how they might be manifested psychologically. As noted, the more domain general constraints tend to be described in terms of restrictions resulting from processes or heuristics that almost as a byproduct then result in certain knowledge structures being favored over others. The more domain specific constraints tend to be stated in terms of restrictions on the types of output structures that are most cognitively “natural” for the learning device. That is, the constraints state that among a set of possible endstate representations that could be induced from a set of a data, representations with certain structural properties are to be preferred. We have seen that this contrast is only approximate since domain general constraints can be structural, and domain specific ones more implicit in processes; in addition, the entire structure/process dichotomy turns out to be an elusive one. Beyond these different ways of manifesting constraints, there are differences in how they constrain. In some cases they seem like absolute prohibitions against certain outputs, while in others they seem like biases (again, the more absolute ones tend to be more common with the domain specific sort). I suspect that it is counterproductive to insist that they ever by completely in-

2. **Problems with Imposters**

   We have seen that without careful analyses designed to distinguish types of constraints, it can be difficult to tell which ones are predominant in any instance of learning. This was evident in Newport’s alternative account of the sorts of constraints responsible for critical periods but is perhaps most striking in cases where the achievement of a domain specific competence becomes useful to many different domains. Thus, if a child gains a new level of competency in spatial representations, he or she might then use that competence to aid cognitive performance in a wide range of domains; with the consequent of appearing to have gained a domain general acquired constraint. If however, the structure of the knowledge of spatial layout is different from other sorts of knowledge and if it is acquired only with the aid of constraints that are particularly tailored for that sort of knowledge and which make other sorts of logically possible spatial representations cognitively non-natural (cf. Gallistel, 1989), it is really a case of domain specific knowledge growth. Unfortunately, it may not always be easy to determine which is which. Thus, if a child gains a new competency in reasoning, one might wish to call it domain specific if the structure of that knowledge and the constraints on its acquisition are unique; but one might also wish to call it domain general if that reasoning is fundamentally involved in all cases of learning. My own bias is to say that if unique sets of constraints are required to learn that knowledge, it itself is domain specific; but that once it is acquired it may in turn exert domain general constraints on all future learning. There are many other types of imposters as well, but they can only be uncovered by explicit analyses of all the possible ways that constraints could model empirical phenomena.

3. **Connecting Constraints with Domains**

   There remains the difficult question of how a learner initially attaches input to a relevant domain versus to an irrelevant domain or to no specific domain at all. What is it about initial input that triggers the activation of one set of constraints and not others? This issue, though critically important, is far beyond the scope of this paper. At least two different mechanisms of domain selection appears feasible, however. One might work through a sort of parallel processing, where sets of constraints from many different domains act on the input, and subsequently get dropped as they fail to pass some threshold of inductive success. Alternatively, it may be that relatively low level sensory or perceptual clues trigger constraints in one domain and/or inhibit others. Thus, certain perceptual features might inform the learner that the input is language-like and should be acquired through a system sub-
ject to linguistically specific constraints. These two mechanisms might be distinguished empirically through different patterns of acquisition. Thus, in the parallel account, one might expect relatively slow initial learning that accelerates as candidate domains are eliminated. Conversely, in the triggering account one might expect initially rapid learning but also that occasionally the wrong domain would be triggered (perhaps by artificial stimuli that mimicked the surface features of a domain—for example, language-like elements in a nonlanguage structure), and the learner would make a false start and reach a dead end. Of course, these mechanisms are not mutually exclusive and might work in concert with each other as well. Other models are also possible but these two alone illustrate the complexity of making such accounts work. For example, how does one reliably link up perceptual triggers to constraints on such things as beliefs and how does one operationalize inductive success in a non-circular manner?

4. How Implicit Can Constraints on Knowledge Be?
   At some point we have to achieve a finer resolution of precisely what the constraints are constraining. The constraints that guide a spider in "learning" to weave its web are guiding something very different from those that guide a child learning a naive mechanics. Although both sets of constraints may well be implicit in that they are never consciously entertained, one set (the spider's), seems much more implicit than the other and may not really be learning at all. This again is a highly complex issue involving clarification of notions of representation, learning, and belief. It is however essential if we are to understand better why we don't wish to talk about constraints on the spider's "learning" about webs but do want to talk about constraints on the infant's learning about objects. Through such a discussion we may also be able to get a better handle on when to attribute the possession of implicit theories.

5. Is "Constraints" a Misleading Notion?
   It has been sometimes argued that the entire idea of constraints is somehow misleading because of its negative connotations. Organisms evolve to solve certain problems in a positive way and so it is more accurate to talk about adaptations than constraints. Perhaps this is true for purely biological traits, but learning is different; because far and away the most central issue of learning is how it ever works at all given the indefinitely large number of generalizations than can be induced from any pattern inputs. This central problem can only be understood in terms of constraints on the classes of representations that are generated by organism; and in this context, constraints are a very positive notion indeed.

REFERENCES


168

KEIL


Safe Takeoffs—Soft Landings

DOUGLAS L. MEDIN, WOO-KYOUNG AHN, JEFFREY BETTGER,
JUDY FLORIAN, ROBERT GOLDSTONE, MARY LASALINE,
ARTHUR MARKMAN, JOSHUA RUBINSTEIN, AND
EDWARD WISNIEWSKI

University of Illinois at Champaign-Urbana

1. INTRODUCTION

The richness of life's possibilities creates a paradox. It poses fundamental limitations on the ability of an organism to reason and act intelligently, as information in its environment is often ambiguous and consistent with an infinite number of interpretations. At the same time, interpretation of this information is required for intelligent action. To further complicate the situation, many problems that are solvable in principle are intractable in practice. As an example, the number of winning sequences of moves in a chess game is finite. Yet determining these sequences from all possible sequences is impractical for the chess player, as there are more possible sequences than atoms in the universe. Cognitive scientists are certainly familiar with underdetermination and computational complexity in association with work on language learnability theory (e.g., Gold, 1967; Wexler & Culicover, 1980), but complexity problems extend to virtually every area of cognition from perception to concept learning to decision making.

If organisms cannot, and therefore do not, examine all possibilities in some cognitive task, they must be "prepared" or biased to learn some things rather than others, to draw some plausible inferences rather than others, and in general to favor some possibilities at the expense of others. As evidenced by the present set of exemplary papers by Brown, Gelman, Keil, Markman, Newport, and Spelke, psychologists are increasingly drawing on analyses of constraints or "guiding principles" as a framework for their research. All of the present authors are interested in development, and it is no accident that developmental psychologists have been leaders in recognizing the importance of constraints. Without guiding principles of some form it simply is unclear how the learning process would ever get off the ground.

The purpose of this comment is to underline some of the important points and issues growing out of developmental work on constraints as reflected in the contributed papers. We first discuss a couple of distinctions and definitions associated with research on constraints, then shift attention to the per-

Correspondence and requests for reprints should be sent to Douglas L. Medin, University of Michigan, Perry Bldg., 330 Packard, Ann Arbor, MI 48104.