

INTRINSICALLY/EXTRINSICALLY<sup>1</sup>  
CARRIE FIGDOR  
(Final draft; please contact me for citation)

Men of all sorts take a pride to gird at me: the brain of this foolish-compounded clay, man, is not able to invent anything that tends to laughter, more than I invent or is invented on me: I am not only witty in myself, but the cause that wit is in other men. (Sir John Falstaff, in *Henry IV* Part 2, Act I, Scene 2)

Shakespeare recognized a distinction between something's having a property in and of itself, and something else's having the same property only when certain external conditions hold. So do we all. Plato distinguished between a fire's being hot by its nature and another thing's being hot because of the fire in it.<sup>2</sup> Some basketball fans claim that Michael Jordan is a world-class player in and of himself, while Scottie Pippen plays world-class basketball only when he's on the court with Jordan. Many people would agree that Paris, France is intrinsically worth living in, but McMurdo Station, Antarctica is not. And Lawrence Summers, the former president of Harvard University, suggested (to his eventual dismay) that while men have an aptitude for science intrinsically, women do not have this aptitude in the same way (assuming they have it at all). In short, in both philosophical and everyday contexts we recognize that things have some of their

---

<sup>1</sup> I would like to thank audiences at the Long Island Philosophical Society, the Claremont Colleges, University of Iowa, College of William & Mary and Oklahoma State University for comments and discussion, in particular Andrea Borghini, James Cain, Stephen Davis, Richard Fumerton, Paul Hurley, Amy Kind, Peter Kung, Alex Raczji, Peter Thielke, Masahiro Yamada and Charles Young.

<sup>2</sup> Plato, *Phaedo* (102b8-105c7), trans. H. Tredennick, in E. Hamilton and H. Cairns, eds., *The Collected Dialogues of Plato* (Princeton, N.J.: Princeton, 1961), 83-87; see also Charles Young, "A Delicacy in Plato's *Phaedo*", *The Classical Quarterly*, New Series vol. 38, No. 1 (1988), pp. 250-251. Michael Jacovides, "Cambridge Changes of Color", *Pacific Philosophical Quarterly* 81 (2000): 142-163, interprets John Locke (*An Essay concerning Human Understanding*, Peter H. Nidditch, ed., (New York: Oxford, 1700/1975): Book II Ch. 8 §19) as arguing that the changing color of porphyry in different lighting conditions is not a "real alteration" in the porphyry because it has its color extrinsically.

properties intrinsically and others extrinsically, and also that the same property can be intrinsically by one thing and extrinsically by another. The primary goal of this paper is to outline an approach to distinguishing the ways in which things have properties along the “had intrinsically”/“had extrinsically” dimension.

I will call the target of my discussion the intrinsically/extrinsically (I-ly/E-ly) distinction, to contrast it with the intrinsic/extrinsic (I/E) distinction. The latter is a distinction between properties. The former is a distinction between ways in which particulars have properties; roughly, it registers the relative independence of a characteristic of an object from relevant changes in that object’s environment. The second goal of this paper is to show that we can and should explain the I-ly/E-ly distinction apart from how we classify properties under the I/E distinction. For this reason, I will avoid terms like “intrinsic properties” (which is ambiguous between the two distinctions) and even “intrinsic properties of objects” in favor of the precise, albeit sometimes awkward, “properties had intrinsically”. The third goal is to show the impact of this analysis of the I-ly/E-ly distinction on its relation to the I/E distinction. The result is a unified framework that illuminates the way we use the I-ly/E-ly distinction and explains the pervasive but philosophically puzzling phenomenon of cross-classification, in which properties classified as extrinsic can be had intrinsically and properties classified as intrinsic can be had extrinsically.

The structure of this paper is as follows. In section I, I motivate an explanation of the I-ly/E-ly distinction. In section II, I distinguish intrinsicness and extrinsicness from those notions captured in the essential/accidental, internal/external, and other nearby distinctions. In sections III, IV and V, I offer an analysis of the I-ly/E-ly distinction and a

basic framework that explains its relation to the I/E distinction. My discussion proceeds from well-known *a priori* analyses (Kim, Langton and Lewis) in section III to an *a posteriori* analysis (Ellis) in section IV to a hybrid analysis that draws on psychological research in counterfactual reasoning in section V. I also suggest how this model applies to other distinctions that have motivated philosophical interest in the I/E distinction: the distinctions between real and mere Cambridge change, intrinsic and instrumental value, qualitative and numerical identity, and strong and global supervenience. In section VI, I elaborate on the basic model in the light of some puzzling cases.

### Section I. Motivating An I-ly/E-ly Distinction

Philosophical discussion of the I/E distinction often begins with an intuitively compelling difference that the subsequent formal analysis is intended to capture. Here's a representative sample of the motivating intuitions:

Intuitively, a property is intrinsic just in case a thing's having it (at a time) depends only on what that thing is like (at that time), and not on what any wholly distinct contingent object (or wholly distinct time) is like.<sup>3</sup>

*You* know what an intrinsic property is: it's a property a thing has (or lacks) regardless of what may be going on outside of itself.<sup>4</sup>

A thing has its intrinsic properties in virtue of the way that thing itself, and nothing else, is. Not so for extrinsic properties, though a thing may well have these in virtue of the way some larger whole is. The intrinsic properties of something depend only on that thing; whereas the extrinsic properties of something may depend, wholly or partly, on something else.<sup>5</sup>

---

<sup>3</sup> Peter Vallentyne, "Intrinsic Properties Defined", *Philosophical Studies* vol. 88 (1997): 209-219.

<sup>4</sup> Stephen Yablo, "Intrinsicness", *Philosophical Topics*, vol. 26, No. 1 & 2 (Spring and Fall 1999): 479-505.

<sup>5</sup> David Lewis, "Extrinsic Properties" *Philosophical Studies* 44 (1983): 197-200.

Properties considered typical of each side of the distinction are then listed: for example, being an uncle or being six miles from a rhododendron, as opposed to having mass or being square.

It is not at all obvious, however, that the distinction motivated by these platitudes is the I/E distinction. A property that a thing has regardless of what's going on outside itself is simply a property that it has intrinsically. Whether the property itself is intrinsic is logically a separate question. For example, one might have a Platonic model of properties as universals that exist independently from particulars, in which the properties are classified *a priori* as intrinsic or extrinsic, and the distinct ways in which particulars may instantiate these properties is another matter. The idea that the I/E distinction simply reflects the distinct ways in which properties are had by particulars requires a theory that connects the two distinctions in this way, for it is a theory which, unlike the Platonic theory, makes the distinction between the properties dependent on the distinction between the ways in which properties are had.

This non-Platonic theory is implicit in the way we think of essential properties and innate properties: each property-level distinction is just reflection of a distinction in the ways properties are had.<sup>6</sup> One thing can be a mathematician essentially and another accidentally, and that is all; there is nothing further to be said in this regard about the property of being a mathematician. Because of this implicit policy, we are not tempted to think that something has a property essentially or innately because the *property* is

---

<sup>6</sup> Alternatively, one might say that there is no genuine property-level distinction in these cases, just a distinction in how properties are had.

essential or innate (*tout court*). For there's simply nothing more to being an essential property than being a property that is had essentially by some individual.<sup>7</sup>

If we treated the I/E distinction this way, an intrinsic property would just be a property that is had intrinsically by some individual and an extrinsic property just one that is had extrinsically by some individual. We would say that Falstaff is witty intrinsically and Prince Hal is witty extrinsically, and that Mr. Smith is an uncle extrinsically and has his mass intrinsically, and there would be nothing further to say about the properties of being witty, being an uncle or having mass. We could even explicitly reject the idea that properties are intrinsic or else extrinsic, since the distinction we would be drawing does not depend on the existence of a property-level distinction. And there would still be philosophical work to do, since we would have the I-ly/E-ly distinction to explain.

But we don't treat the I/E distinction this way. We think the definition of a property determines whether the property is intrinsic or extrinsic, whereas the definition plays no such role in whether a property is essential, accidental, innate or acquired. As a result, the intuitive distinction expressed in the platitudes and our policy for clarifying it are at odds. The platitudes reflect an *a posteriori* distinction, but the analyses that are supposed to illuminate this distinction are *a priori*.

We can diagnose the problem more precisely. There already exists widespread recognition of the I/E and I-ly/E-ly distinctions, called the "global" and "local"

---

<sup>7</sup> I do not mean to imply that essentialism is false; I am neutral on that issue. My point is simply that the locution "essential property" is not ambiguous in the way "intrinsic property" is; properties are essential (or not) relative to individuals that have them.

distinctions.<sup>8</sup> It is also widely recognized that the relation between these distinctions cannot be straightforward. For example, we can start from the I/E (“global”) distinction, and hold that if a property is intrinsic, then all of its instances have it intrinsically; or we might start from the I-ly/E-ly (“local”) distinction, and hold that if anything has a property intrinsically, the property is intrinsic. The same would hold for extrinsicness.<sup>9</sup> The usual cases used to illustrate the lack of a straightforward relation are disjunctive properties, such as being made of tin or next to something made of tin, or being square and accompanied or red and lonely. Such designer properties are typically introduced as counterexamples to particular analyses of the I/E distinction. They are used to show, for example, that properties whose definitions are such that they count as intrinsic by the analysis have instances that intuitively are had extrinsically, and so therefore the properties should be classified as extrinsic, *pace* the analysis in question.

But the possibility is not even considered that any classification of how properties are had that is based purely on property definitions is inadequate. This possibility will not become evident by focusing on disjunctive properties that are designed by definition to act as counterexamples. It becomes immediately evident when we look at the practices that give rise to the motivating platitudes. For suppose one claimed that being witty is extrinsic because it means, or one stipulates that it means, “makes other people laugh”. It doesn’t follow that Falstaff isn’t witty intrinsically. Shakespeare knew what “being witty”

---

<sup>8</sup> I.L. Humberstone, “Intrinsic/Extrinsic”, *Synthese* vol. 108 (1996): 205-267 (see p. 206). I avoid this well-known terminology because, as will become clear below, the I-ly/E-ly distinction is not just the I/E distinction applied to instances, nor is the I/E distinction just the I-ly/E-ly distinction applied to properties.

<sup>9</sup> The first option is called deriving a local notion from a global one, the second deriving a global one from a local one (Humberstone, *ibid.*; Brian Weatherson, “Intrinsic vs. Extrinsic Properties”, *The Stanford Encyclopedia of Philosophy* (Spring 2007 edition), Edward N. Zalta, ed., URL = <http://plato.stanford.edu/archives/spr2007/entries/intrinsic-extrinsic/>. Weatherson claims the first is “undoubtedly true” but doubts the second due to disjunctive properties. I think both are false, for reasons explained in the text.

meant when he distinguished Falstaff's wit from everyone else's in this way. Or suppose one claimed that shape is intrinsic because, given its definition, there's no logical inconsistency in something's having the shape it does in an otherwise empty world; it doesn't follow that a stretched rubber band does not have its shape extrinsically.<sup>10</sup> The obvious feature these cases share with many disjunctive properties is that, intuitively, some particulars have these properties intrinsically and others have them extrinsically. But what these cases reveal, and many disjunctive properties do not, is that property definitions are not invariably determining whether we think a property is had intrinsically or else extrinsically. Yet these are the practices that are used to motivate I/E analyses based on property definitions.

It is unlikely that we will understand the puzzling relation between any I/E distinction and the ways in which properties are had until we examine the I-ly/E-ly distinction more closely. We can start this task by expressing the original motivating intuitions as perhaps they always ought to have been:

*You know what it is to have a property intrinsically: it's for a thing to have (or lack) a property regardless of what may be going on outside of itself.*

This claim unambiguously motivates an explanation of the I-ly/E-ly distinction.

Representative examples of what we want to explain would include Mr. Smith's being an uncle and Mr. Smith's having mass, not being an uncle or having mass. It doesn't follow that the I-ly/E-ly distinction will be wholly *a posteriori*. For the definition of a property presumably plays *some* role in determining whether its instances are had intrinsically or else extrinsically. The burden of this paper is to explain what else must be going on.

---

<sup>10</sup> The example is from Brian Ellis, *The Philosophy of Nature* (Montreal & Kingston: McGill-Queen's, 2002): 51-54; see also Humberstone (op. cit.) and Bradford Skow, "Are shapes intrinsic?", *Philosophical Studies* 133 (2007): 111-130. I discuss Ellis below.

## Section II. A Brief Interlude: Other Distinctions.

Before analyzing the I-ly/E-ly distinction, it may be worthwhile to take a moment to clarify how having a property intrinsically or extrinsically differs from some of the other ways of having them: essentially or accidentally, internally or externally, or innately or by acquisition. Falstaff is not claiming that he was born witty, nor that he would cease to exist if he ceased being witty, nor is he locating his wit spatiotemporally – although he might have done any or all of these things too.<sup>11</sup> If we think of a property heuristically as a standard that all its instances meet, then ways of having a property are different ways of meeting that standard. We distinguish how an individual came to meet the standard to begin with (innately or via acquisition), whether it would still exist if it no longer met the standard (essentially or accidentally), whether it is met in virtue of something located inside the individual (internally or externally) and whether it would still be met in relevant counterfactual circumstances (intrinsically or extrinsically). A probable reason for highlighting when a thing has a property intrinsically is because properties had this way are considered dependable: they are not vulnerable to loss due to factors beyond the thing's control – often, factors literally beyond the thing, many of which tend not to be in its control.<sup>12</sup> Put the other way around, a property had intrinsically

---

<sup>11</sup>Many of those discussing the I/E distinction take care to distinguish it from the essential/accidental distinction, although they may hold that essential properties are a subset of the intrinsics (e.g., Ellis, p.50) – more precisely, that the properties a thing has essentially are a subset of the properties it has intrinsically. This is not universal; Lynne Rudder Baker, in *Persons and Bodies: A Constitution View* (New York: Cambridge, 2000), p. 30, fn. 10, remarks that she “departs from the tradition in holding that not all essential properties are intrinsic.”

<sup>12</sup>Christine M. Korsgaard (“Two Distinctions in Goodness”, *The Philosophical Review* vol. 92, no. 2 (April 1983): 169-195), while discussing G.E. Moore's view of intrinsic value (“The Conception of Intrinsic Value”, *Philosophical Studies* (London: Routledge and Kegan Paul, 1922); *Ethics* (Oxford, U.K.: Oxford, 1912), Ch. VII) explains on p. 175 that on that view intrinsic goodness “is dependent only on the thing's intrinsic nature and is just as constant: so long as the thing remains what it is it has the same value: and the



is one that is highly impervious to the impact of novel circumstances; from this point of view, having a property intrinsically implies an unfortunate inability to adapt.

Of course, in ordinary contexts we often use terms like “essentially”, “internally”, “innately” and “intrinsically” interchangeably (*mutatis mutandis* for “extrinsically” and the other contraries) even if the terms pick out different distinctions. For example, comedian Richard Pryor, who died on Dec. 10, 2005, was eulogized as follows:

Paul Mooney, a longtime friend and a frequent writing collaborator, said Mr. Pryor’s skills “came from God”. “It was innate,” said Mr. Mooney, who called himself Mr. Pryor’s “black writer” and lent the comic his first car, a 1952 Ford, during his early days. “He could have been born in Japan and it still would have been there. Geniuses just are.” (The New York Times, Dec. 13, 2005)

A property can be had innately in that one is born with it, and yet not had intrinsically in that in relevant counterfactual circumstances one might lack it. Otherwise there would be no wasted natural talents. But however the terms might be used or confused in ordinary conversation, we are asking different questions when we want to know whether or not a thing acquired a property, whether or not it would cease to be what it is if it lost a property, whether or not a thing has a property by virtue of what is currently located inside itself, and whether or not it would still have a property in relevantly different counterfactual circumstances.

That said, there is a sense of “in and of itself” in which it is used to denote an internal/external (or interior/exterior) distinction, rather than a claim about independence.

The two senses are distinct. The internal/external sense involves spatiotemporal claims,

---

value is the same, of course, for everyone and so also objective.” Young (p. 250) describes Socrates as seeking an explanation of when an object has a property “its title to which is insecure, in the sense (to judge from 102b8-c8), that the object’s having the property is not guaranteed by its being what it is.”

whereas the independence sense is essentially modal.<sup>13</sup> A part can be had intrinsically by a whole in the internal sense, but not had intrinsically in the independence sense; similarly, a sniper's bullet is had intrinsically in the internal sense, but not the independence sense, by her victim. The interior/exterior distinction is arguably prior to the independence sense, since many standard analyses of the I/E distinction essentially involve reference to objects that are distinct from, in that they share no parts with, the object that has the property in the way it does.<sup>14</sup>

In this paper, I will focus on the sense more widely associated with “in and of itself” – that of independence from relevant external circumstances. This is the sense that is of most concern in discussions of intrinsicness and extrinsicness. For example, when we want to know whether something has its value intrinsically, we are asking whether it would still be valuable in relevant counterfactual circumstances, not where its value is spatiotemporally located (assuming the latter question even makes sense).

### Section III. Explaining The I-ly/E-ly Distinction: Definitions.

An adequate account of the I-ly/E-ly distinction must explain what kinds of factors determine whether a property is had intrinsically or else extrinsically in any given

---

<sup>13</sup> There is also a sense of ‘internal’ that may be understood as ‘included in a thing’s nature’, which is roughly the same as the independence sense. I’ll discuss this sense of ‘internal’ in Section VI.

<sup>14</sup> Most also consider the intrinsic/extrinsic distinction distinct from the non-relational/relational distinction (but see Robert Francescotti, “How to Define Intrinsic Properties” *Nous* vol. 33 no. 4 (1999): 590-609), even though the two relations are often conflated in other contexts (e.g., Marc Lange, in *An Introduction to the Philosophy of Physics* (Oxford: Blackwell, 2002), on p.18 characterizes the properties a thing has intrinsically as those which “depend only on itself, not on its relations to other things”, while “a typical non-intrinsic (“extrinsic”) property is the property of standing in a certain relation to something else.”) However, the non-relational/relational distinction is distinct from both senses of intrinsic. Every object is identical to itself, but an object stands in this relation in and of itself; the property of containing a spinal column is one a vertebrate has in and of itself, yet it involves a relation. Humberstone (pp. 212-13) argues persuasively that the non-relational/relational distinction is best understood as a distinction between predicates: since we can change 1-place predicates into 2 (or n) -place predicates at will (and vice versa), it is simple to make a property “involve a relation” or not.

case, and how they do. It is helpful to consider an ordinary example to see what this task involves. When a basketball fan tells me that Jordan is a world-class basketball player in and of himself, I am not being told that this is so because he would still have that property at a world in which he exists in complete isolation. Such worlds are irrelevant.<sup>15</sup> I'm not told that Pippen is a world-class basketball player extrinsically because there are necessary existents; if there are any, they're irrelevant. Most importantly, the role of the nature of being a basketball player – which by definition involves a team of five people – in determining how this property is had is not at all clear, since Jordan and Pippen are indistinguishable in this regard. There are, in short, implicit constraints on the counterfactual possibilities we normally consider when making I-ly/E-ly classifications, both in terms of the relevant sorts of objects and the relevant class of possible worlds.<sup>16</sup> An adequate account must explain these relevance parameters and how they are applied.

Standard I/E analyses provide a useful point of departure, since they aim to capture the notion of independence involved in intrinsicness and extrinsicness, and the I-ly/E-ly distinction shares this core idea.<sup>17</sup> To fix ideas, we can begin with Kim's influential I/E distinction:

(KE) F is extrinsic = df. Necessarily any object x has F only if some contingent object wholly distinct from x exists.

---

<sup>15</sup> Experimental social psychologists have identified consistent constraints on the alternative possibilities that people normally entertain in counterfactual reasoning about events (Seelau et. al. 1995). I discuss these constraints in Section V; at this point, I ask the reader only to imagine herself in an ordinary context in which an I-ly/E-ly distinction, such as this one about Jordan, might be drawn.

<sup>16</sup> It is also likely that properties may be had I-ly or E-ly in degrees, as some recent analyses of the innate/acquired distinction also allow (e.g., Andre Ariew, "Innateness and Canalization", *Philosophy of Science*, vol. 63 supplement: proceedings of the 1996 Biennial Meetings of the Philosophy of Science Association, Part I: Contributed Papers (September 1996): S19-S27. I set this question aside here.

<sup>17</sup> I set aside definitions of intrinsicness in terms of duplication, whereby (roughly) an intrinsic property of a thing is one shared by all of its duplicates (e.g., Lewis "Extrinsic Properties"). Any mention of the I/E distinction in this paper can be analyzed in the idiom of duplication if the reader desires.

(KI) F is intrinsic = df. Possibly some object x has F although no contingent object wholly distinct from x exists.<sup>18</sup>

It is presupposed that the question of whether or not a property can be had by a thing that exists in an otherwise empty world depends entirely on the property's definition and what logically follows from it. This assumption ensures that the notion of independence at the heart of the I/E distinction is logical (or conceptual) independence. We can swiftly obtain an I-ly/E-ly distinction by modifying Kim's distinction as follows:

(KE-ly) x has F extrinsically = df. x has F, and necessarily x has F only if some contingent entity wholly distinct from x exists.

(KI-ly) x has F intrinsically = df. x has F, and it is possible both that x has F and no contingent entity y wholly distinct from x exists.<sup>19</sup>

(KI) and (KE) do an admirable job of classifying many properties, but Kim's analysis is widely considered vulnerable to a counterexample raised by Lewis involving the property of being lonely.<sup>20</sup> A thing is accompanied iff it coexists with some wholly distinct contingent object, and lonely iff it is not accompanied. Lewis objects that being lonely is just as extrinsic as being accompanied, but (KI) classifies it as intrinsic. Intuitively, a

---

<sup>18</sup> Jaegwon Kim, "Psychophysical Supervenience", *Philosophical Studies* 41 (1982): 51-70, reprinted in his *Supervenience and Mind: Selected Philosophical Essays* (New York: Cambridge, 1993), 175-193. I use Lewis's ("Extrinsic Properties") version of Kim's original distinction, which Kim expressed as follows:

(i) G is rooted outside the objects that have it = df. necessarily any object x has G only if some contingent object wholly distinct from x exists;

(ii) G is internal = df. G is neither rooted outside times at which it is had nor outside the objects that have it.

Vallentyne also provides a more concise version of Kim's definition:

P is intrinsic = df. Px is compatible with  $\sim Ax$ ,

where A is the property of being accompanied by at least one distinct contingent object; and what's not intrinsic is extrinsic. (See also Humberstone, p. 229). Vallentyne explains the intuition behind Kim's definition of intrinsic as follows: "The intuitive idea is that P can be had by an object even in a world with no other distinct objects." That may be true of Kim's or Vallentyne's intuitions; it is unlikely to be true of Shakespeare's or Summers'. At the very least, none of the intuitions expressed in the platitudes requires it to be true that a property I have intrinsically is one that I would have in an otherwise empty world.

<sup>19</sup> Derk Pereboom, in "Why a scientific realist cannot be a functionalist", *Synthese* vol. 88 (1991): 341-358, offers a local version of Kim's definition: F is an extrinsic property of x iff x's having F logically entails or rules out there being a contingent thing wholly distinct from x; otherwise a property is had intrinsically. However, this definition is closer to (LLE-ly), described below, rather than (KE-ly).

<sup>20</sup> "Extrinsic Properties", pp. 198-199.

thing's being lonely depends on what the world is like; since the core idea of intrinsicness is contextual independence, the property ought to be classified as extrinsic.

However, it is not clear that being lonely is a counterexample to (KI), as opposed to (KI-ly). The definition of being lonely is such that it gets classified as an intrinsic property by (KI). To consider this classification wrong and (KI) inadequate on the basis of how things have their properties is to assume, inconsistently, that we do not apply (KI) and (KE) based entirely on property definitions. It may not be a mark in (KI)'s favor that it yields intrinsic properties that are had extrinsically, but (KI) is not contradicted by this fact given how the I/E distinction is drawn. (KI-ly), in contrast, classifies a thing's being lonely as a property it has intrinsically, and this does seem incorrect. (By analogy: S's being an only child does not just depend on S.) (KE-ly) is also in trouble. (KE-ly) says that if a thing has a property extrinsically, it cannot be the only contingently existing thing. But anything that is lonely can be the only contingently existing thing. We can consistently take Lewis's counterexample as an objection to (KI-ly) and (KE-ly) because differences in how things have their properties are precisely what these analyses are meant to capture.

In response to the alleged counterexample to (KI) and (KE), however, Langton and Lewis (1998) propose an influential alternative I/E analysis based on Kim's:

(LLI) F is intrinsic = df. Fx is compatible with Ax and with  $\sim$ Ax, and so is  $\sim$ Fx, where A is the property of being accompanied and  $\sim$ A is the property of being lonely, defined as above; and what's not intrinsic is classified as extrinsic.<sup>21</sup> By this definition,

---

<sup>21</sup> Rae Langton and David Lewis, "Defining 'Intrinsic'", *Philosophy and Phenomenological Research* vol. 58 no. 2 (June 1998): 333-345; see also David Lewis, "Redefining 'Intrinsic'", *Philosophy and Phenomenological Research* vol. 63 no. 2 (2001), Special Symposium: Defining Intrinsic: 381-398. In the text I use Vallentyne's version of (LLI); Langton and Lewis express this conception, but do not display the

an intrinsic property is one that is independent of accompaniment or loneliness, and to be independent of accompaniment or loneliness is for these four cases to be logically possible as determined by the property's definition: there can be an accompanied individual that has F, a lonely individual that has F, an accompanied individual that lacks F and a lonely individual that lacks F. Being lonely comes out extrinsic by this analysis, since its definition entails that there cannot be an accompanied individual that has it or a lonely individual that lacks it.

Unfortunately, the success of this I/E distinction does not carry over entirely to the corresponding I-ly/E-ly distinction:

(LLI-ly) x has F intrinsically = df. x has F, and x's having F is compatible with x's having A and with x's having  $\sim$ A.

(LLE-ly) x has F extrinsically = df. x has F, and x's having F is not compatible with either x's having A or else x's having  $\sim$ A.

(These analyses are simplified because in I-ly/E-ly judgments there is no need to consider how *other* objects might have the property.) By these definitions, a thing's being lonely is correctly classified as a property it has extrinsically. (LLE-ly) also avoids the problem that beset (KE-ly), for an object's being the only existing thing is plainly incompatible with its also not being the only contingently existing thing. Nevertheless, these analyses fail to capture many of the I-ly/E-ly classifications we make. If Falstaff is witty in and of himself and Prince Hal is not, this is not because we judge that Falstaff would have this property in an otherwise empty world and Prince Hal wouldn't: by incorporating A and

---

definition in this (or any) form. Their analysis is subject to two restrictions: it does not apply to impure, or non-qualitative, properties (properties that contain reference to particular objects, places or times) nor to disjunctive properties. Langton and Lewis also distinguish between basic intrinsics, which are natural properties, and non-basic intrinsics, which are disjunctive properties or their negations. These are property-level restrictions on the application of (LLI) or (LLE) which do not affect my analysis of the I-ly/E-ly distinction; disjunctive properties are (often) just properties that can be had either way.

$\sim A$  in the definitions, (LLI-ly) and (LLE-ly) require us to consider worlds that are irrelevant to the distinction Falstaff drew between himself and Prince Hal. Nor does it explain how he did draw the distinction given what is required by definition to be witty, since Falstaff and Prince Hal are alike in this regard. In short, using (LLE) we may (correctly) classify the property of being witty as extrinsic based on its definition, but (LLE-ly) gives us no clue as to how this extrinsic property might coherently be had intrinsically. The cross-classification problem remains as puzzling as ever.

But we can begin to address the problem with slight revisions of the definitions of being lonely and being accompanied: one is lonely if one is the only relevant contingently existing thing at a relevant world, and one is accompanied if one co-exists with the relevant contingently existing object(s) at a relevant world. In nearly all cases (we will see degenerate cases below), relevantly lonely worlds are those that lack relevant external objects. These revisions explicitly raise the issues of how we determine which objects must go missing for something to be relevantly lonely (as opposed to lonely *tout court*, which we may call Lewis-lonely) and how we determine which worlds are relevant if we do not (or almost never) consider the logically possible worlds where an object is Lewis-lonely. We can make these relevance parameters explicit in a set of revised definitions:

(I-ly)  $x$  has  $F$  intrinsically = df.  $x$  has  $F$ , and  $x$ 's having  $F$  is compatible, in the relevant set of possible worlds, with  $x$ 's having  $A_R$  and with  $x$ 's having  $\sim A_R$ ;

(E-ly)  $x$  has  $F$  extrinsically = df.  $x$  has  $F$ , and  $x$ 's having  $F$  is not compatible, in the relevant set of possible worlds, with  $x$ 's having  $A_R$  or else  $x$ 's having  $\sim A_R$ ,

where  $A_R$  is the property of being relevantly accompanied and  $\sim A_R$  that of being relevantly lonely.<sup>22</sup> This analysis does not yet explain, but only makes room for

---

<sup>22</sup> Some recent analyses of the I/E distinction build relevance implicitly into their I/E definitions (Vallentyne; David Denby, "The Distinction between Intrinsic and Extrinsic Properties", *Mind* vol. 115

explaining, how Falstaff can be intrinsically witty and Prince Hal extrinsically witty, and how a stretched rubber band can have its shape extrinsically and a hardened steel band can have the same shape intrinsically. But what determines the relevance parameters in (I-ly) and (E-ly)?

#### Section IV. Explaining the I-ly/E-ly Distinction: Relevance, Step One

Consider again the claim that Michael Jordan is a world-class basketball player in and of himself, while Scottie Pippen is a world-class player only when he's on the court with Jordan. An ordinary fan may make coherently make this claim even though both men are alike (at times) in being world-class basketball players; even though she knows that no one can play basketball without four other people; even though what's needed for Jordan and Pippen to exist implies that many other things exist at every possible world she considers; and even though she does not even consider a Lewis-lonely world as relevant to her distinguishing them.<sup>23</sup> How does she do this? How do *we*? In what follows, I will argue that the relevance parameters in (I-ly) and (E-ly) follow the same heuristics that constrain the relevant alternatives considered in ordinary counterfactual reasoning. This is because the I-ly/E-ly distinction serves an explanatory purpose that depends on such reasoning for its success. Since these constraints reflect our empirical knowledge, empirical facts, not just property definitions, play an essential role in drawing this distinction.

---

(2006): 1-17), but do not explain what determines relevance. Such analyses are roughly analogous to (LLI) and (LLE) when "relevantly-lonely" and "relevantly-accompanied" are substituted into the original formulas and left unexplicated.

<sup>23</sup> The latter claim is supported by empirical research on counterfactual reasoning discussed in section V.



Before explaining the Jordan/Pippen (or Falstaff/Hal) examples, consider the superficially simpler case of Mr. Smith's being an uncle. We judge that Mr. Smith has this property extrinsically; his being an uncle is not compatible with his being relevantly lonely. In this case it seems clear that the definition of being an uncle determines the relevant external objects. If offspring of Mr. Smith's sibling or siblings did not exist, he would not have the property. Although this case seems straightforward, if not trivial, there are two unexplained features that bear emphasis. First, when making our judgment we don't consider those worlds where Mr. Smith is female, even though these are also worlds where he wouldn't have the property. But the definition didn't require us to restrict our counterfactual reasoning to those worlds where Mr. Smith is just as he is (and everything else is just as it is) except the relevant external objects are absent. So what determines the class of relevant worlds? Second, our judgment about Mr. Smith's being an uncle may seem trivial because it is not logically possible for *any* male who does not have siblings with offspring to be an uncle. Anyone who knows the definition knows this immediately – it just *follows* from the definition that he has this property extrinsically, one might say. But when considering the I-ly/E-ly distinction, our concern is only with whether Mr. Smith would be an uncle in the relevant counterfactual circumstances; it is not with what we think we can infer about how other objects have a property just from the definition of the property. If such inferences were always valid, there would be no distinction between Falstaff and Hal to explain. So why does the inference fail generally? And why does it go through in this case? I will return to these questions below.

Now consider a physical object's shape. Shape is usually classified as an intrinsic property, but most physical objects (unlike the shapes of some abstract objects) have their

shapes extrinsically.<sup>24</sup> The actual shapes of physical objects partly depend on the external physical forces acting on them. A stretched rubber band illustrates this clearly, but even unstretched bands are subject to external stresses. So if shape is an intrinsic property, it is an intrinsic property that is had extrinsically in many actual cases.<sup>25</sup> Resistance to this claim does not stem from lack of basic familiarity with physical theory, since it is common to classify mass as intrinsic and weight as extrinsic, presumably because the latter depends in part on gravitational forces. Instead, it seems symptomatic of the assumption that property definitions determine not just whether a property is intrinsic or else extrinsic, but also how properties are had. It is quite true that (e.g.) being a square is being a plane figure with four equal sides and four equal angles, and that based on this definition being a square will be classified as an intrinsic property by standard I/E definitions. But the present discussion is about actual squares, and the problem with actual squares having their shapes the way we know they do is that it confounds a purely *a priori* approach to the I-ly/E-ly distinction.<sup>26</sup>

---

<sup>24</sup> It is possible that in classifying shape as intrinsic the intended interpretation of “intrinsic” is the “internal” sense. For example, Lewis (“Extrinsic Properties”) declares that “some properties of things are entirely *intrinsic*, or *internal*, to the things that have them: shape, charge, internal structure”. However, this would at best be an equivocation, since the subsequent definition motivated by the case of shape explicates intrinsicness in terms of logical independence of external circumstances (or duplication), not spatiotemporal location.

<sup>25</sup> So apparently is molecular structure, in a molecular model consistent with both quantum mechanics and observation (R. Wooley, “Must a Molecule Have a Shape?”, *Journal of the American Chemical Society* 100 (1978): 1073-1078 and “The Molecular Structure Conundrum” *Journal of Chemical Education* 62 (1985): 1082-1084; Pierre Claverie and Simon Diner, “The Concept of Molecular Structure in Quantum Theory: Interpretation Problems”, *Israel Journal of Chemistry* vol. 19 (1980): 54-81; Steven Weininger, “The Molecular Structure Conundrum: Can Classical Chemistry be Reduced to Quantum Chemistry?” *Journal of Chemical Education* 61 (1984): 939-944; Robin LePoidevin, “Space and the Chiral Molecule”, in Nalini Bhushan and S. Rosenfeld, eds., *Of Minds and Molecules: New Philosophical Perspectives on Chemistry* (Oxford: Oxford, 2000): 129-142.

<sup>26</sup> This may explain Ellis’s complaint (p. 51) that philosophers “have not succeeded in explicating a concept [of intrinsic] that is of much relevance to the theory or practice of science.”

To address this problem, Ellis proposes an I/E distinction based on what he calls causal independence, rather than logical (definition-based) independence. Stated in terms of an I-ly/E-ly distinction, the essence of his proposal is as follows:

(CI-ly) x has F causally-intrinsically iff x has F, F is posited by physical theory T, and Fx is compatible with the absence of relevant properties and laws posited by T;

and what's not had CI-ly is had CE-ly.<sup>27</sup> The idea is that while a physical object typically has its actual shape in a way that depends in part on external forces acting on it, the way in which they have these shapes is distorted from the point of view of our theoretical models. Within these models, the theoretical analogues of actual things have their shapes in an undistorted way – that is, in a way not dependent on the relevant forces. This is the way actual things *would* have their shapes if they were not subject to the forces posited by the model. In short, to have a shape causally-intrinsically is to have a shape in an undistorted way, where what's undistorted is determined by the model.

Actual physical objects *can* have their shapes causally-intrinsically, although simply having a shape like that of the relevant model-theoretic analogue is not sufficient; it must have that shape in an undistorted way. To use Ellis's example, a strip of hardened steel shaped the way a given rubber band is shaped when stretched "would have this shape naturally, independently of any accidental forces. Therefore, in the causal sense of 'intrinsicality', the actual shape of a body may or may not be the same as its intrinsic shape." That is, the steel band, but not the rubber band, has its shape in an undistorted way, because if the relevant forces were removed the steel band would still have its shape

---

<sup>27</sup> Ellis does not display a definition, so (CI-ly) is based on his remarks. This I-ly/E-ly form may best express Ellis's intent, since he claims: "Intrinsicality in the causal sense is therefore not a property of properties, as many of those who have tried to explicate a logical concept of intrinsicality have supposed. It has more to do with the role that the property has in physical theory." He also applies his model to rest mass (the mass the object would have were it not in motion relative to us) and a particle's spin.

but the rubber band would not.<sup>28</sup> Again, we need not consider the steel band's shape in a Lewis-lonely world, only in a relevantly lonely world.

The importance of Ellis's proposal in this context is that it provides a model of how the relevance parameters in (I-ly) and (E-ly) can be determined other than by definitions. In other words, (CI-ly) is just a special case of (I-ly) in which the relevant external objects are fixed by physical theory. It is still the case that the definition of being square determines what is necessary and sufficient for being square; by this definition, and (LLI), being square counts as an intrinsic property. But the definition does not determine what's relevant to a thing's being square intrinsically or else extrinsically. In the case of physical objects, physical theory tells us what's relevant to whether a thing has its shape in one way or the other. In general, the relevant external objects may be determined by a theoretical model of the system's behavior, including but not limited to the models of basic physical theory.<sup>29</sup>

This insight will not yet get us out of the woods. For consider again the case of Michael Jordan. Being a world-class basketball player, unlike having a particular shape, is not a property that is posited by physical theory or any other theoretical model, so Ellis's proposal seems inapt. Yet the definition of being a world-class basketball player, unlike that of being an uncle, also does not determine which objects or worlds are relevant to drawing the distinction between Jordan and Pippen. For the fan who draws this distinction can surely infer from its meaning, as we do, that no one could be a

---

<sup>28</sup> Its having its shape in an undistorted way also depends on the relevant forces *not* being present; so the steel band would have to have the shape its theoretical analogue has both when the forces are present and when they are not in order to have it causally-intrinsically. (CI-ly) would need to be modified for this.

<sup>29</sup> It is worth emphasizing the fact that Ellis's proposal can be incorporated naturally into my account. In the literature on the I/E distinction, his contribution is either ignored or mentioned briefly only to be set aside.

basketball player in a world that lacks at least four other people. But she doesn't thereby infer that no one can have this property intrinsically. She – like Shakespeare, like Summers, like all of us in many ordinary contexts – draws the distinction *in spite of* this logical fact. But if what's relevant to drawing her distinction is not determined by a definition nor by a theoretical model, we still lack an adequate explanation of how the I-ly/E-ly distinctions are drawn.

Since the Jordan/Pippen case was presented as a typical example of the I-ly/E-ly distinction, one may be tempted to think that that the I-ly/E-ly distinctions we make are either too vague or unsystematic or even incoherent to theorize about. If so, then the fact that the I/E distinction is *a priori* is a relief – although it would still follow that we should not necessarily judge a candidate I/E distinction by whether it captures intuitions about how properties are had. But this pessimistic conclusion about the I-ly/E-ly distinction is unwarranted.

#### Section V: Explaining the I-ly/E-ly Distinction: Relevance, Step Two

Turn again to the case of Falstaff and Hal, which is parallel to the Jordan/Pippen case. Suppose being witty is defined as “makes other people laugh”. How can Falstaff be intrinsically witty and Prince Hal not? How can Falstaff be intrinsically witty at all, given what is necessary for being witty?

There is a very simple explanation. What would distinguish their wittiness is a quality of Falstaff – a disposition – that is the source of his wittiness and the cause of wittiness in others. Falstaff's being witty would be compatible with his being relevantly

lonely, because Falstaff would have this disposition even if other people did not exist.<sup>30</sup> In the relevantly lonely worlds he would still be witty, but his wittiness would not be manifested. We may not accept his boast, but we certainly understand it: we understand the distinction that Shakespeare is drawing and how it is possible for Falstaff to be witty even at the relevantly lonely worlds. Similarly, the sports fan might say Jordan is a world-class basketball player intrinsically because he would still have the disposition to play as he does at a relevantly lonely world – where, for example, he’s alone on the court – even though it would not be manifested. In short, having a property intrinsically is, in some cases, having another property intrinsically – in this case, a disposition – that explains why the object has the first property at all.<sup>31</sup>

But is this simple explanation the right one, in these ordinary cases? Plausibly, yes. It is standard procedure to seek explanations of observable phenomena in terms of other, usually more basic, phenomena. This explanatory tendency is not limited to scientific pursuits. The same approach is reflected in the thought that Jordan has some underlying cause of his talent; the fan merely adds that this source is something he possesses in and of himself. It is also revealed in Summers’ distinction between those who have an intrinsic aptitude for science from those who do not: he drew the distinction to explain why there are relatively few women in the upper reaches of science. We needn’t posit dispositions. We don’t think Paris, France has a disposition that makes it worth living in intrinsically. But to someone who distinguishes Paris from McMurdo

---

<sup>30</sup> In effect, Falstaff himself is the relevant object for being or not being witty. He can both co-exist with the relevant object, and so be relevantly accompanied, and be the only relevant object, and so be relevantly lonely. In this degenerate case, the distinction between being relevantly accompanied and being relevantly lonely collapses, since being accompanied just by oneself is equivalent to being lonely. Relative to Hal, Falstaff is the relevant *external* object, and the usual accompanied/lonely distinction revives. Since this is the usual case, I will typically write of “relevant external objects” in the text.

<sup>31</sup> This is not circular: dispositions can also be had I-ly or E-ly. See also section VI.

Station in this way, some ineffable *je ne sais quoi* explains why Paris is worthwhile being in independently of any further reasons we might have for being there.

The important point is that I-ly/E-ly distinctions are typically drawn in order to provide explanations, however rudimentary or proleptic, of how a thing has its properties. By an explanation, I mean nothing more fancy than an answer to a why-question.<sup>32</sup> Such explanations may involve positing an additional property (or properties), of which dispositions are just a well-known type, that are not required by definition for having a property but which serve the purpose of explaining how the property is had. In such cases it is the way in which the explanans property is had, or are thought to be had, that determines whether an explanandum property is had intrinsically or else extrinsically, whether the explanandum property itself is classified as intrinsic or else extrinsic. In these cases, the inference from what is required by definition for having a property to the way in which *any* individual has that property can fail. It is also plausible that when no further explanation can be given – for example, when we have explained how the disposition that explains Falstaff’s wittiness is had – the ultimate explanans property is had intrinsically or else extrinsically as determined by its definition or by a theoretical model, likely in physics. So in principle there is no need to list further ways of determining the relevant objects in (I-ly) and (E-ly). However, that analysis does require the following revision:

(I-ly\*) x has F intrinsically = df. x has F, and (i) x’s having F is compatible, in the relevant set of possible worlds, with x’s having  $A_R$  and with x’s having  $\sim A_R$ , or (ii) x has G intrinsically, and x’s having G explains x’s having F.

---

<sup>32</sup> In particular, the explanations I have in mind need not satisfy models of scientific explanation, and they are typically incomplete relative to scientific or philosophical standards of completeness. For example, Summers’s suggestion regarding the scientific aptitude of women was an answer to the following question: Why are there so few women in the highest ranks of science? The why-questions that motivate the other I-ly/E-ly examples (or answers) I’ve used throughout the text are similarly straightforward.

(E-Iy\*) x has F extrinsically = df. x has F, and (i) x's having F is not compatible, in the relevant set of possible worlds, with x's having  $A_R$  or else x's having  $\sim A_R$ ; or (ii) x has G extrinsically, and x's having G explains x's having F.

With this analysis, plus the claim that the I-Iy/E-Iy distinction plays an explanatory role, we are in a position to explain the I-Iy/E-Iy distinction and its relation to the I/E distinction in a way that solves the puzzle of cross-classification.

First, the idea that the I-Iy/E-Iy distinction is drawn for explanatory purposes allows us to explain how the relevant sets of possible worlds are determined. As with the I/E distinction, the core of the I-Iy/E-Iy distinction is the idea of independence from the environment, cashed out in counterfactual terms. But because of its explanatory function, the counterfactuals that support the I-Iy/E-Iy distinction will be constrained to those alternatives that are relevant to the proposed explanation. Moreover, we can explain these constraints if we recognize this reasoning as just a special case of ordinary counterfactual reasoning, which follows robust constraints on which counterfactual possibilities are entertained. These general constraints can explain how we limit the set of worlds to a contextually relevant class in the special case of drawing I-Iy/E-Iy distinctions.

Following Kahneman and Tversky, psychologists have sought to isolate the parameters that govern and constrain everyday counterfactual reasoning and develop cognitive models that explain these regularities.<sup>33</sup> The constraints concern which

---

<sup>33</sup> D. Kahneman and A. Tversky, "The simulation heuristic", in D. Kahneman, P. Slovic and A. Tversky, eds., *Judgment under uncertainty: Heuristics and biases* (New York: Cambridge): 201-208; Daniel Kahneman and Dale T. Miller, "Norm Theory: Comparing Reality to Its Alternatives", *Psychological Review* vol. 93 no. 2 (1986): 136-153; Ruth M.J. Byrne, "Mental models and counterfactual thoughts about what might have been", *Trends in Cognitive Sciences* vol. 6 no. 10 (October 2002): 426-431; Eric P. Seelau, Sheila M. Seelau, Gary L. Wells and Paul D. Windschitl, "Counterfactual Constraints", in Neal J. Roesch and James M. Olson, eds., *What Might Have Been: The Social Psychology of Counterfactual Thinking*, (Mahwah, N.J.: Lawrence Erlbaum Associates, 1995); Ahogni N'gbala and Nyla R. Branscombe, Nyla R., "Mental Simulation and Causal Attribution: When Simulating an Event Does Not Affect Fault Assignment", *Journal of Experimental Social Psychology* 31 (1995): 139-162; Gary L. Wells and Igor Gavanski, "Mental Simulation of Causality", *Journal of Personality and Social Psychology* vol. 56, no. 2



alternative possibilities are considered, which facts are more easily mutable, and which are more likely to be mutated. One source of constraint appears to be the purpose the reasoner has for reasoning counterfactually.<sup>34</sup> A primary purpose is to explain how and why something occurred. By entertaining alternative possibilities, we may seek to isolate necessary causes (John's father's stopping to chat with neighbors on the way to picking John up from school) and sufficient causes (a drunken driver's hitting John while he was waiting to be picked up). Other goals include assigning responsibility and blame, providing consolation to others and oneself, amplifying negative and positive affective responses, and guiding future behavior. (These need not be mutually exclusive.)

Besides the conscious purposes we may have, there are constraints that operate automatically in that without conscious effort they sharply restrict the infinite range of alternative possibilities that could be entertained to a very small set. The natural-law constraint includes ("folk") knowledge of basic physical, chemical and biological laws and facts. Possibilities with mutated natural facts are almost never considered in normal discourse. For example, when mutating the event of a plane crash, we never consider "if only it had fallen up". (This possibility would also violate the purpose constraint, on the assumption that we are trying to isolate a sufficient cause of the crash or assign fault.) On the other hand, when we seek to explain the 1986 Challenger disaster, subjects commonly consider "if only the O-rings had been checked".

---

(1989): 161-69. The experimental social psychological research focuses on counterfactual reasoning about events (often those that evoke regret or other affective responses); quite plausibly, the same constraints apply to counterfactual reasoning about objects. Kahneman and Miller explicitly link their concept of mutability – a modal notion referring to that feature of some aspects of reality under which these aspects are more or less easily altered in counterfactual reasoning than others – to David Lewis's analysis of counterfactuals (*Counterfactuals* (Blackwell and Harvard, 1983)), in which some possible worlds are "closer" or involve "smaller" changes.

<sup>34</sup> Here I largely follow Seelau et al., "Counterfactual Constraints".

The availability constraint involves factors that make certain mutations more likely than others. For example, we are more likely to mutate exceptions rather than normal events (e.g., Tom's taking a new route home when he got into a crash, rather than his usual route), actions rather than inactions (e.g., subjects judge that Dick, who is unhappy after having transferred to a new college, feels more regret than Harry, who is unhappy after not having transferred), aspects of reality over which we may have control rather than those that we do not (e.g., the O-rings), and more recent events rather than those earlier in a series (e.g., the last missed field goal in a football game that was lost by 2 points). In general, an event or aspect that is contextually salient is more available for mutation.

The natural-law, availability and purpose constraints can always be violated, and regularly are in science fiction and philosophical discussion. But we see clear evidence of them in the ordinary I-ly/E-ly distinctions discussed above. In the case of Mr. Smith's being an uncle, we do not bother to consider the world where he is female, as this would violate the natural-law constraint; our omission also makes sense if we are consciously focusing on the definition in order to identify the external objects that might affect his having the property. We also do not consider being an uncle the sort of property that calls for further explanation. Jordan's talent, on the other hand, does cry out for explanation: why is his level of play so extraordinary? Yet when the sports fan distinguishes Jordan from Pippen, she does so in a way that also follows the constraints. For example, she does not consider worlds where everyone else is 10 feet tall. Such worlds would not only violate the natural-law constraint; they would obviate her attempted explanation of Jordan's talent and her attempt to assign him full credit for his

success. The same goes for Falstaff when he seeks to credit himself for his wittiness: worlds where he couldn't be witty, for whatever reason, are simply not relevant. In all these cases, the person drawing the distinction can be wrong; the I-ly/E-ly distinction is not subjective (at least, that claim goes beyond what I have argued for here). But the distinction that is drawn reflects the purposes of the person drawing it, even if it can be challenged.

Second, we have a natural explanation of the puzzle of cross-classification, when a property classified as extrinsic (such as being witty) can be had intrinsically, and a property classified as intrinsic (such as being square) can be had extrinsically. Cross-classification arises when something's having a property that counts as intrinsic by definition is explained by its having another property extrinsically; *mutatis mutandis* for extrinsic properties that are had intrinsically. The following diagram (fig.1) illustrates this relation between the I/E and I-ly/E-ly distinctions:

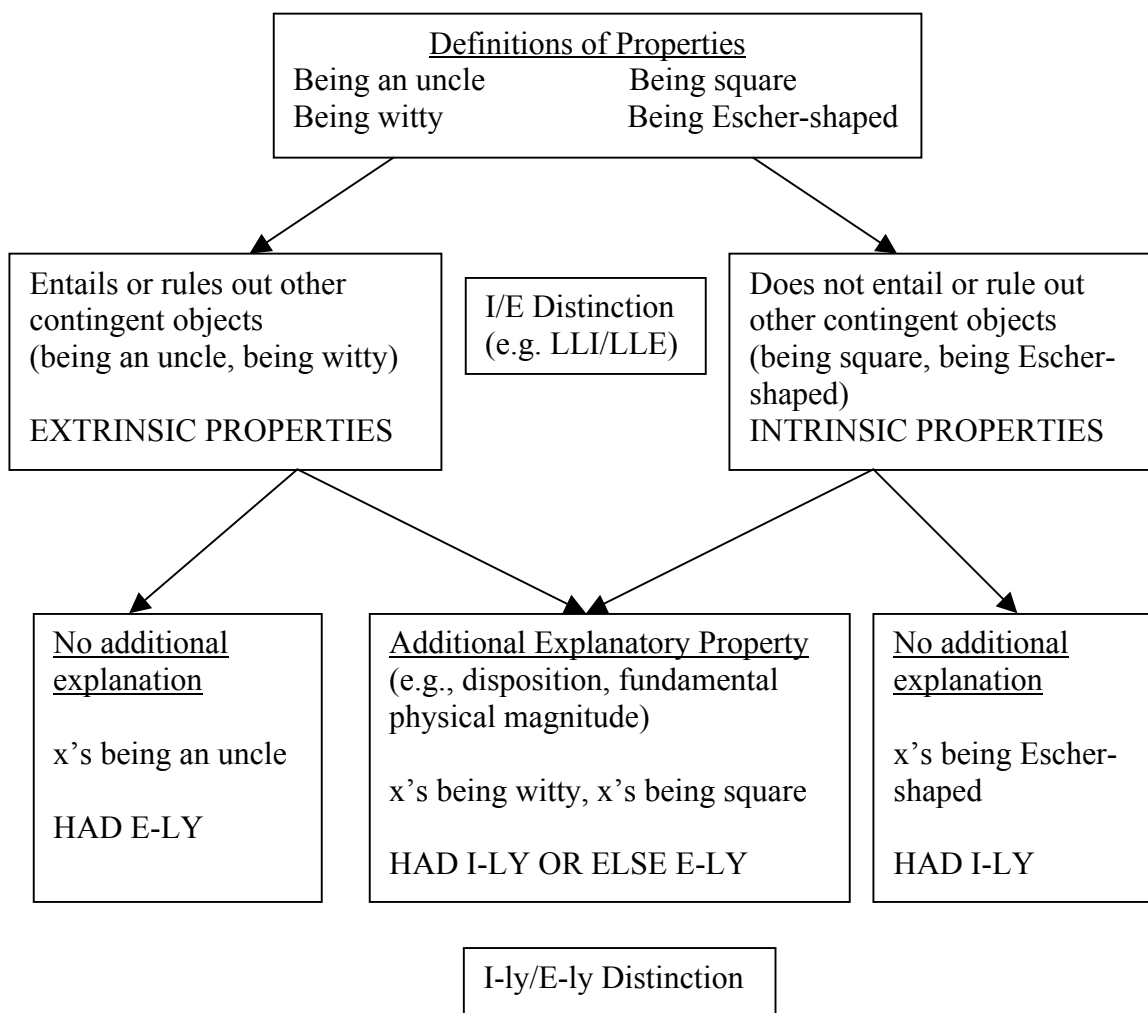


Fig. 1: A Framework for the I/E and I-ly/E-ly Distinctions<sup>35</sup>

The overall account of the I-ly/E-ly distinction remains fairly simple, however: (i) definitions and empirical theory each can determine the relevant external objects in (I-ly\*) and (E-ly\*); (ii) empirically-determined constraints on ordinary counterfactual reasoning reliably determine the relevant worlds in (I-ly\*) and (E-ly\*); (iii) cross-

<sup>35</sup> Being Escher-shaped is having the shape of an Escher-drawn object that cannot be had by an actual physical object. Although I use Langton and Lewis's I/E distinction to illustrate, the reader may plug in her favorite analysis. For example, Gene D. Witmer, William Butchard and Kelly Trogdon ("Intrinsicity without Naturalness", *Philosophy and Phenomenological Research* vol. 70, no. 2 (March 2005): 326-350) propose a two-tiered account of the I/E distinction – superficially similar to this element of the present I-ly/E-ly account -- to eliminate Langton and Lewis's reliance on the natural/unnatural property distinction.

classification between the I/E and I-ly/E-ly distinctions can arise when something's having one property is explained by its having another; (iv) because of (iii), we cannot invariably infer from what is required by definition to have a property to how it is had in any particular case.<sup>36</sup> I should note that despite the emphasis on explanation in this model, the I-ly/E-ly distinction itself is metaphysical. The distinction matters to us because of its explanatory utility. Merely ascribing a property to a thing to explain its behavior is often insufficient to provide a satisfactory explanation; one reason for this is because other things can have that same property but not behave the same way.

This account does not require us to change how we draw the I/E distinction. We may still classify properties as intrinsic or extrinsic depending on their definitions, as current I/E analyses do, although in this case single-case cross-classifications are not necessarily counterexamples to a proposed I/E analysis. Nevertheless, we might consider a better analysis to be one that yields fewer cross-classifications, or that best distinguishes those cases where cross-classification is very rare or non-existent. Note that the cross-classification issue is just the problem (not always recognized as such) of how to classify what have been called "mixed" cases in the literature. Kim's I/E analysis classifies a property as intrinsic if it is logically possible for it to be had intrinsically (i.e., it can be had in a Lewis-lonely world). Most other analyses classify a property as extrinsic if it is logically possible for it to be had extrinsically; disjunctive properties with "mixed" instances, such as being square or (Lewis-) accompanied, are routinely classified as

---

<sup>36</sup>If divine power were invoked to explain Jordan's prowess, the distinction between him and Pippen would not be an I-ly/E-ly distinction, but a divinely-inspired/not divinely-inspired distinction. But while Jordan would have his prowess extrinsically, it would not be because of the *definition* of being a world-class basketball player, but because of his dependence on God (Pippen might still have this property extrinsically because of his dependence on Jordan). So two individuals can have the same property extrinsically, and both because of an explanatory factor had extrinsically, but the explanatory factors need not be the same.

extrinsic. But in the absence of arguments that the I/E distinction must satisfy a kind of law of excluded middle, and that there is a non-arbitrary reason to classify properties with “mixed” instances on one side of the distinction rather than the other, it is hard to see which side, if either, is correct. An alternative solution would be to allow that the I/E distinction comes in degrees, such that a property is intrinsic (extrinsic) if most of its instances are had intrinsically (extrinsically), assuming some way of counting instances. But this solution would tie the I/E distinction to the I-ly/E-ly distinction in a way that would mark a sharp departure from the *a priori* methods used so far.

Before elaborating on the empirical element of this model in the following section, I should note briefly how it applies to disjunctive and conjunctive properties and to the philosophical tasks for which the I/E distinction has been used.

In the case of disjunctive properties, the relevance parameters are those that hold for each disjunct when it is had individually (not as a disjunct). For example, a relevantly-lonely red square can have the property of being red or relevantly-accompanied either intrinsically or extrinsically, depending on whether colors can be had intrinsically or extrinsically (which in turn depends ultimately on our best explanation of color). A relevantly-accompanied blue square presumably would have this disjunctive property extrinsically, depending just on the property’s definition (assuming, reasonably, that no further explanation is wanted). For conjunctive properties, the relevance parameters for how a conjunctive property is had are not so reducible. The property of being red and relevantly-accompanied would be had extrinsically by a relevantly-accompanied red square (on the assumption that no further explanation of how it has this property is wanted). What is crucial to keep in mind, in both cases, is whether there is a

reason to posit an additional explanatory property, and, if so, how that property is had. As noted, it is a consequence of the model that we cannot invariably infer from a property's definition whether it is had intrinsically or else extrinsically in particular cases.

As for how this model affects the philosophical tasks to which the I/E distinction has been put, its use in clarifying the real vs. mere Cambridge change distinction will serve as an illustration. According to Geach, Cambridge changes are changes in the predicates true of a thing.<sup>37</sup> All changes are Cambridge changes, but some are such that (a) there is a change of predicates true of a thing and (b) there is a real change in the thing. A change that satisfies (a) and (b) is a real change, while a change that satisfies only (a) is a mere Cambridge change. For example, when Socrates dies, there is a real change in Socrates but a mere Cambridge change in Xanthippe, who becomes a widow.

The I/E distinction has been used to explain this difference: real changes are changes in intrinsic properties, while mere Cambridge changes are changes in extrinsic properties. However, this explanation is too quick: some properties that may be classified as intrinsic by their definition, such as shape, may be had extrinsically. The above model provides the means for a more precise understanding of the real/mere Cambridge change distinction. The model allows us to distinguish between two claims: real changes are those that involve changes in properties had intrinsically, or real changes are those that involve physical changes in a thing. The first option seems unjustified except insofar as it is implicitly assumed to be a variant of the latter. The latter option makes the real/mere Cambridge change distinction compatible with what physical theory tells us about how things have their physical properties. For whether physical changes are changes in properties a thing has intrinsically or else extrinsically depends on physical theory, not on

---

<sup>37</sup> Peter Geach, *God and the Soul* (London: Routledge and Kegan Paul, 1969): 70-72.

property definitions.<sup>38</sup> The upshot is that we can no longer strictly identify real change with change in intrinsic properties, but we can say that real change is typically (but not by definition) a change in the properties a thing has intrinsically. The I/E distinction would remain a reliable indicator of the real/mere Cambridge change distinction.

## VI. Elaborating The Basic Model: Causes and Conditions, Constancy and Mutability

It may seem that by using dispositions as an illustrative case of explanatory properties, I have made the I-ly/E-ly distinction dependent on the I/E distinction, since dispositions are widely considered to be intrinsic properties. But there is no circularity here, because dispositions – and any other explanatory property – are subject to the I-ly/E-ly distinction too. A disposition can be had intrinsically or else extrinsically even if they are all classified as intrinsic properties.<sup>39</sup> One way a sports fan could be wrong about Michael Jordan is if it turns out that he has his disposition, or the physical properties that explain it, extrinsically. The basic model could be elaborated to distinguish between the ways in which properties are had according to our “folk” intuitions about the natural world and the ways in which they are had according to our best physical theories. I have not done so for two reasons: first, “folk” intuitions are updated in the face of physical

---

<sup>38</sup> Jacovides suggests that real change involves a change in fundamental explanatory properties. Carol E. Cleland, “The Difference Between Real Change and *Mere* Cambridge Change”, *Philosophical Studies* 60 (1990): 257-280, argues that they involve changes in dispositions posited by physics.

<sup>39</sup> There may well be extrinsic dispositions (e.g., see Jennifer McKittrick, “A Case for Extrinsic Dispositions”, *Australasian Journal of Philosophy* 81 (2003): 155-174), although that is an I/E claim about which I remain neutral. But well-known counterexamples to a simple conditional analysis of dispositions also may be considered cases where a disposition that is classified as intrinsic is had extrinsically (e.g., the finked wire in C.B. Martin, “Dispositions and Conditionals”, *The Philosophical Quarterly* 44 (1994): 1-8, or the gold chalice in Mark Johnston, “How to Speak of the Colors”, *Philosophical Studies* 68 (1992): 221-263. According to a simple conditional analysis, an object is disposed to M when C iff it would M if it were the case that C (Michael Fara, “Dispositions”, *The Stanford Encyclopedia of Philosophy* (Fall 2006 Edition), Edward N. Zalta, ed., URL = <http://plato.stanford.edu/archives/fall2006/entries/dispositions/>). Further discussion of these possibilities must be left for another occasion.



theory, although not always or immediately; second, the role of explanatory properties in the I-ly/E-ly distinction would not change.

It also seems that we are willing to allow some slippage when drawing I-ly/E-ly distinctions. For example, if we are willing to say that some physical objects are square, then we also ought to be willing to say that some physical objects have their squareness intrinsically. That is, if being square is a property that can be had by physical objects, then a physical object's being square intrinsically will be acceptable to a similar extent. Ellis seems to take this for granted in his steel band example. This vagueness does not undermine the I-ly/E-ly account any more than it forces us to deny that some physical objects are squares.

However, bringing empirical theory into the picture raises other problems, nicely illustrated by hydrangeas.<sup>40</sup> A difference in soil pH determines the flowers' color by affecting the aluminum content in the plant. If the soil is acidic, the flowers are blue, and if it is basic they are pink – the opposite of litmus paper. So do a hydrangea's flowers have their color extrinsically, or are they just caused to have their color by the soil pH? In general, given that empirical theory can determine whether a property is had intrinsically or else extrinsically, how do we distinguish an *external cause* of a thing's having a property and an *external condition* that determines whether it has the property intrinsically or extrinsically? I'll address this problem in steps.

---

<sup>40</sup> See <http://hgic.clemson.edu/factsheets/HGIC1067.htm>, downloaded on June 2, 2008. Many thanks to Amy Hurshman for this example.

To begin with, the problem is *not* that of the well-known context-relativity or vagueness between a cause and a condition.<sup>41</sup> It is relative to our explanatory purposes which factor in a total cause we pick out as a sufficient cause (e.g., the match's being lit) and which as a necessary cause or condition (e.g., the presence of oxygen). These differences between causal factors highlighted as causes and those treated as background conditions are reflected in ordinary counterfactual reasoning.<sup>42</sup> A chemist seeking to discover the nature of a hydrangea's color might try to isolate a sufficient external cause. If when the soil pH is changed from its original state, the hydrangea's *new* flowers bloomed in a different color, he might conclude that the soil pH is a sufficient external cause of their color, relegating the other factors to background conditions. But this is not sufficient for the hydrangea to have its color extrinsically. That depends on whether the *same* flowers would change color whenever soil pH changes – for example, from pink to blue when switched from basic to acidic soil, and back again to pink if replaced in basic soil. It is this sort of dependence on external conditions that matters in the I-ly/E-ly distinction and that must be distinguished from external causes.

We can begin to illuminate the nature of this dependence with the following case. Suppose we inject a hydrangea with permanent pink dye such that all its flowers would remain pink through any change in soil pH. Does it have its color intrinsically? We might say, with Plato, that the dye is pink intrinsically but the hydrangea is pink because it contains the dye, and hence it is pink extrinsically. On this view, although the dye is constantly present and spatiotemporally internal, it counts as external to the plant in terms

---

<sup>41</sup> John L. Mackie, "Causes and Conditions", in E. Sosa and M. Tooley, eds., *Causation* (New York: Oxford, 1993): 33-55, reprinted from *American Philosophical Quarterly* 2/4 (October 1965): 245-55, 261-4.

<sup>42</sup> See N'gbala and Branscombe, "Mental Simulation and Causal Attribution: When Simulating an Event Does Not Affect Fault Assignment" and Wells and Gavanski, "Mental Simulation of Causality".

of its nature. (This is the other sense of internal/external I mention in fn.9.) But there is another option: we could say the dye *is* now part of the hydrangea's nature because of its permanence. Its constancy in the face of relevant external change might justify the claim that it has its color intrinsically, even though its color had an external cause and it is the only hydrangea to have its color in this way.

Our background knowledge about plants would probably lead us to side with Plato in this case; more precisely, our modal intuitions about plants and their properties would follow the natural-law constraint. But there may be no principled way to determine when a constant condition that had an external cause becomes part of a thing's nature, and hence makes it or any property explained by that condition a property it has intrinsically, or whether the constant condition forever remains external to the thing's nature (wherever it might be spatiotemporally located) and makes it or any property explained by that condition a property it has extrinsically. In the case of weight, we automatically choose the latter route, perhaps because gravity or other physical forces are not divisible and so are not even candidates for becoming part of an individual's nature the way a dram of dye can be. This does not undermine the model so much as prompt further inquiry into any general principles under which a constant condition either remains or ceases to be external to a thing's nature. Acquired dispositions are good test cases.

But there is an important distinction that this discussion has elided so far, brought out by the following case. Assume that pain is always had intrinsically, on the assumption (granted for the sake of argument) that we identify pain with pain qualia. If I always wear a cilice that causes pain (like the character Silas in The Da Vinci Code), there is a

constant external condition that is sufficient for my pain. Do I have my pain extrinsically? No, because what matters in the I-ly/E-ly distinction is mutability, and mutability and constancy lie on distinct continua. Mutability is a modal notion; it is that feature of some aspects of reality by which they are altered in counterfactual reasoning more or less easily or often. The natural-law and availability constraints are constraints on mutability. Constancy is a temporal notion. The cilice is a constant external condition, but it is highly mutable: I could remove it at any time if I so chose. Dependency on an condition that is constant, rather than immutable, does not determine that a property explained by the presence of that condition is had extrinsically.

The same principle holds for the dyed hydrangea. By assumption the dye is immutable, and not merely constant, once injected; its immutability, not its constancy, makes it a candidate for becoming part of the hydrangea's nature. The term "permanent" is ambiguous between these notions. The cilice case also reflects the difference between mutability and constancy, in a more subtle way. In that case, our automatic response is to think of it as remaining external to my nature no matter how long and constantly I wear it. For without this response, the cilice case would not present a problem to considering the pain as something that I have intrinsically. And it is the fact that it is highly mutable, despite its constancy, that makes it ineligible for being considered part of my nature – in sharp contrast to the psychological disposition I have that explains why I always wear it. Thus, the cilice remains a constant external condition (or sufficient cause) of my pain; it is neither an immutable external condition (which would make my pain something I have extrinsically) nor an immutable part of my nature (which would make my pain something

I have intrinsically in an atypical way). So I do have my pain intrinsically in the intuition-sanctioned way, given the identification of pain and pain-qualia.

The Falstaff/Hal case also shows how mutability and constancy come apart. Falstaff need not be constantly in Hal's presence for Hal's wittiness to depend on Falstaff's in the relevant sense. What matters is whether Hal's wittiness is mutable in the relevant counterfactual situations, given Falstaff's claim that he himself is the relevant external object (external to Hal). From Falstaff's perspective – which is the one that matters, given that it is he who is drawing the distinction – the answer is clearly *yea*.

### Concluding Remarks.

I have proposed an analysis of the distinction immediately motivated by our intuitions that properties can be had intrinsically or else extrinsically: the I-ly/E-ly distinction. The basic difference between this distinction and the I/E distinction is that the former depends on *a posteriori* as well as *a priori* ways of determining which objects and worlds are relevant for classifying the ways in which properties are had. I also argue that the I-ly/E-ly distinction is used for explanatory purposes, and that the counterfactual reasoning that supports it falls within the purview of empirical studies on the nature of and constraints guiding ordinary counterfactual reasoning in general. I further propose a model for understanding the relation between the I/E and I-ly/E-ly distinctions that explains the puzzle of cross-classification: how intrinsic properties can be had extrinsically and vice versa. Finally, I suggest how the model affects the philosophical tasks for which the I/E distinction has been used, and elaborate on it in response to some puzzling cases.