manuscript No.
(will be inserted by the editor)

Exhibiting interpretational and representational validity

Michael Baumgartner

the date of receipt and acceptance should be inserted later

Abstract A natural language argument may be valid in at least two nonequivalent senses: it may be interpretationally or representationally valid (cf. Etchemendy 1990). Interpretational and representational validity can both be formally exhibited by classical first-order logic. However, as these two notions of *informal* validity differ extensionally and first-order logic fixes one determinate extension for the notion of *formal* validity (or consequence), some arguments must be formalized by unrelated nonequivalent formalizations in order to formally account for their interpretational or representational validity, respectively. As a consequence, arguments must be formalized subject to different criteria of adequate formalization depending on which variant of informal validity is to be revealed. This paper develops different criteria that formalizations of an argument have to satisfy in order to exhibit the latter's interpretational or representational validity.

Keywords logical formalization; logical validity; interpretational validity; representational validity; argument reconstruction; Etchemendy

1 Introduction

Beall and Restall (2006) have provided novel support for logical pluralism, i.e. for the claim that there are at least "two different accounts of deductive logical consequence, two different senses of 'follows from'" (29). They argue that the pre-theoretic or informal notion of an argument's validity is genuinely ambiguous. Moreover, this ambiguity does not simply arise from different choices of logical constants. Rather, the claim is that relative to one particular specification of logical constants, say the constants provided by classical first-order logic, an argument can be valid in at least two different ways. To see this, consider the modal characterization of informal validity that can be found in virtually all textbooks: An argument $\langle \Gamma, \Psi \rangle$ consisting of a set of premises Γ and of a conclusion Ψ is valid if and only if it is impossible for all members of Γ to be true while Ψ is false. The ambiguity arises from the modality involved in this characterization. There are at least two ways in which it can be impossible for the premises in Γ to be true and the conclusion Ψ to be false:

M. Baumgartner
Institute of Cognitive Science, University of Osnabrück, Albrechtstrasse 28, 49076 Osnabrück, Germany
E-mail: michael.baumgartner@uni-osnabrueck.de

(I) there is *no permissible reinterpretation* of the non-logical (categorematic) terms contained in Γ and Ψ under which every member of Γ is true and Ψ false;
(II) there is *no possible world* in which every member of Γ is true and Ψ false.

Spelling out the modality informally associated with the validity of arguments in terms of (I) and (II) does not amount to a mere terminological difference, but yields two nonequivalent notions of validity. This is best illustrated by means of arguments featuring analyticity, as the following notorious exemplar:

(a) Clooney is a bachelor. Therefore, Clooney is an unmarried man.

Both the premise and the conclusion of (a) are subject-predicate statements that do not feature expressions corresponding to first-order logical constants. "Clooney" is a name that may be reinterpreted in terms of any other singular term, and "...is a bachelor" and "...is an unmarried man" may be reinterpreted in terms of any other (first-order extensional) 1-place predicate. Accordingly, a reinterpretation of the categorematic terms in (a) that renders (a)'s premise true and its conclusion false is easily found: simply substitute "...is an unmarried man" by "...is a woman" in (a)'s conclusion. That is, spelling out the informal notion of validity in terms of (I) entails that (a) is an invalid argument. By contrast, as the notions of a bachelor and of an unmarried man are interdefined, the transition from (a)'s premise to (a)'s conclusion is analytical, i.e. it is warranted by the meanings of the involved concepts only. There is no possible world where "Clooney is a bachelor" is true and "Clooney is an unmarried man" false. Therefore, if the validity of arguments is understood along the lines of (II), (a) is determined to be a valid argument.

According to Beall and Restall (2006), (I) and (II) are but two conceivable ways to cash out informal validity. In fact, they take the informal notion of validity to be a mere schema which is given by what they call the *Generalized Tarski Thesis*:

(GTT) An argument is valid$_x$ if, and only if, in every case$_x$ in which the premises are true, so is the conclusion.

While instantiating the variable *case*$_x$ by permissible reinterpretations and possible worlds yields two notions of validity that are accessible to classical first-order logic, i.e. that validate all classically valid inferences, taking cases$_x$ to be *situations* or *stages of proof constructions* yields two non-classical specifications of informal validity. According to Beall and Restall, there is no fact of the matter whether any one of those instantiations of (GTT) is *the correct* understanding of informal (pre-theoretic) validity. They hold that these understandings do not conflict. Rather, depending on a given reasoning context, it may be more suitable or fruitful to spell out validity in terms of one rather than another of the possible instantiations of (GTT).

Whereas the radical pluralism that treats classical and non-classical instantiations of (GTT) on a par has been criticized by several authors (cf. Field 2009; Bueno and Shalkowski 2009; Vecsey 2010), pluralism only with respect to the two classical instantiations (I) and (II) is less controversial. Plainly, spelling out informal validity in terms of (I) is much more common among (mathematical) logicians. Nonetheless, there is a long standing tradition, going back to Wittgenstein's *Tractatus*, according to which precisifications of informal validity along the lines of (II) are equally justifiable (cf. e.g. Etchemendy 1990; McFetridge 1990; Read 1994; Hale 1996; Jackson 2007; Baumgartner and Lampert 2008). According to Beall and Restall (2006, 40-43), (I) fares better when it comes to cashing out the idea that the validity of an argument exclusively hinges on the meanings of the logical constants, while (II) emphasizes that the truth of the premises of a valid argument necessitates the truth of its conclusion.

This paper neither addresses the correctness of logical pluralism nor the question which of (I) and (II) is better suited for spelling out what is informally meant by characterizing an argument as valid. Rather, I simply assume that both (I) and (II) are viable precisifications of informal validity. Against the background of this assumption, I then investigate the criteria that first-order formalizations have to satisfy when it comes to formally accounting for an argument's (I)- or (II)-validity, respectively. According to a frequently cited slogan, logic is the philosopher's tool to study correct reasoning and to exhibit the validity of arguments. Yet, before formalisms can be applied to render the validity of a natural language argument formally transparent, the latter's component statements must be transferred into the syntax of a pertaining formalism, i.e. they must be formalized. And here severe problems emerge, because the grammatical surface of natural language is often radically misleading with respect to underlying logical forms.

While standard textbooks content themselves with illustrating logical formalization with a handful of paradigmatic examples supplemented with commentaries to the effect that formalizing essentially is an artistic skill, Brun (2004) has presented the first book-length investigation exclusively dedicated to developing criteria that are intended to provide a systematic understanding of logical formalization and to assist in justifying and evaluating the adequacy of different formalization candidates. Further studies concerned with establishing criteria of adequate formalization include Epstein (1990, 1994), Sainsbury (2001), Baumgartner and Lampert (2008), and Brun (2012).[1]

The existing literature on criteria of adequate formalization does not distinguish between criteria that are suitable for exhibiting (I)- and (II)-validity. However, as these precisifications of informal validity differ extensionally and first-order logic fixes one determinate extension for the notion of *formal* validity (or consequence), some arguments must be formalized by unrelated[2] nonequivalent formalizations in order to account for their (I)- or (II)-validity. Accordingly, a first-order formalization $\langle \gamma, \psi \rangle$ of an argument $\langle \Gamma, \Psi \rangle$ must meet different criteria, depending on whether $\langle \gamma, \psi \rangle$ is intended to exhibit (I)- or (II)-validity. Neglecting the difference between (I)- and (II)-validity, therefore, constitutes a severe gap in the current literature on criteria of adequate formalization. This is the gap this paper intends to fill.

Etchemendy (1990) has famously distinguished between two different readings of standard models of logical formulas, to which he refers as *interpretational* and *representational semantics*. Section 2 locates the difference between (I)- and (II)-validity in this model-theoretic context and spells out (I)-validity as interpretational validity and (II)-validity as representational validity. In sections 3 and 4, I then develop two sets of criteria formalizations of arguments have to meet in order to exhibit interpretational and representational validity, respectively.

[1] The problem of systematizing logical formalization has also been addressed from a different angle in the literature. Instead of developing criteria of adequate formalization, authors as Davidson (1984), Chomsky (1977), Massey (1975), or Montague (1974), implicitly or explicitly subscribe to the ambitious project to define effective formalization procedures. This second thread in the formalization literature will be sidestepped in this paper.

[2] Two FOL formulas ϕ and ψ are said to be *unrelated* iff neither ϕ is a specification of ψ nor vice versa and there does not exist a formula χ that is a specification of both ϕ and ψ. For the relevant notion of specification cf. Brun (2004, 320) or Lampert and Baumgartner (2010, 89-90).

2 Interpretational view vs. representational view

Cashing out the modality involved in an argument's informal validity in terms of permissible reinterpretations or possible worlds, of course, simply amounts to replacing one modality by another. This maneuver immediately raises the follow-up question as to what counts as a *permissible* reinterpretation and a *possible* world, relative to which an informally valid argument is entailed to be truth-preserving. For instance, a predicate as "...is a bachelor" may be reinterpreted in terms of any other expression of the same semantic type. But what accounts for sameness of semantic type? While "Clooney" is certainly not of the same type as "...is a bachelor", what about "...is evenly distributed over the earth's surface" (cf. Sainsbury 2001, 50)? Analogous problems arise when it comes to delineating the realm of possible worlds. Possible worlds are often understood as complete sets or configurations of compossible atomic states of affairs. Yet, what are atomic states of affairs and which of them are compossible?

I shall not attempt to provide analyses of the modalities involved in (I)- and (II)-validity here. Rather, I join Shalkowski (2004) in doubting that (non-modal) analyses of these modalities are feasible. Hence, in what follows I just treat them as primitives. Whoever has no pre-theoretic conception of the realm of permissible reinterpretations of a given argument and possible worlds relative to which to evaluate the truth values of premises and conclusions, has no pre-theoretic conception of (I)- and (II)-validity either. Informally determining an argument to be valid presupposes clarity on what counts as permissible reinterpretations or possible worlds. That does not mean that somebody who informally judges an argument $\langle \Gamma, \Psi \rangle$ to be (I)- or (II)-valid can in fact construct the set of all permissible reinterpretations of $\langle \Gamma, \Psi \rangle$ or the set of all possible worlds relative to which $\langle \Gamma, \Psi \rangle$ is truth-preserving. It merely means that for any reinterpretation $\langle \Gamma', \Psi' \rangle$ of $\langle \Gamma, \Psi \rangle$ and any world w it is determinable whether $\langle \Gamma', \Psi' \rangle$ is permissible and whether w is possible. For the purposes of this paper, this kind of clarity with respect to the modalities involved in (I)- and (II)-validity shall be assumed to be given.

While the informal notions of validity give rise to numerous questions, first-order logic (FOL)—the tool designed to, among other things, formally exhibit the (I)- and (II)-validity of arguments—provides a completely unambiguous and straightforward notion of formal validity. Let me briefly recap the *model-theoretic* machinery by means of which FOL validity is commonly spelled out. A model $\mathfrak{M}$ of FOL is a structure consisting of a non-empty domain D and a function $\mathfrak{I}$ that assigns truth values $\{T, F\}$ to sentence letters, single elements of D to name letters, and sets of n-tuples of elements of D to predicate letters of arity n. Given an appropriate variable assignment g that assigns elements from D to the free variables in a formula ϕ, satisfaction of ϕ by g in $\mathfrak{M}$ can then be defined in the ordinary recursive way. This allows for defining truth of a formula ϕ in $\mathfrak{M}$ in terms of satisfaction of ϕ by the empty variable assignment $g_\emptyset$ in $\mathfrak{M}$. Finally, this yields the model-theoretic notion of FOL validity for an argument scheme $\langle \gamma, \psi \rangle$, where γ is a set of FOL sentences (i.e. FOL formulas without free variables) and ψ is an FOL sentence:[3]

FOL validity: An argument scheme $\langle \gamma, \psi \rangle$ is FOL-valid if and only if ψ is true in every model $\mathfrak{M}$ in which all elements of γ are true.

[3] For details on this model-theoretic background cf. any textbook, e.g. Chiswell and Hodges (2007, sect. 7.3); Barwise and Etchemendy (1999, sect. 18.2).

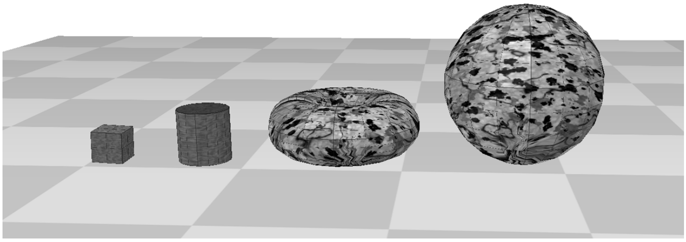

Fig. 1 Mary's world

Many have emphasized the usefulness of model theory when it comes to modeling (in the 'modeling-a-system' sense of the term) semantic aspects of natural language, for instance:[4]

> Truth in a model is interesting because it provides a transparent and mathematically tractable model—in the "ordinary" sense (...)—of the less tractable notion of truth. (Hodes 1984, 131)

The term "model" has two very different scientific uses: a model of a system or phenomenon is an idealized representation, a simulation, or even a theory of relevant aspects of the modeled system, whereas a model of a formal language is a mathematical structure, i.e. a mathematical object consisting of certain kinds of sets, and, as such, does not represent or simulate anything. In order to clearly keep these two notions apart, I subsequently speak of *models** whenever the notion is used in the former sense.

Even though FOL models in themselves are not representations, they suggest themselves as instruments to model* the construction of sets of permissible interpretations and possible worlds, respectively. To implement FOL validity as such a modeling* device for informal validity, FOL models must be read either as representations of permissible reinterpretations or of possible worlds. These are exactly the two readings Etchemendy (1990) has labeled the *interpretational* and the *representational* views of models. On the interpretational view, different models are understood as different meaning assignments to the terms of a corresponding formal language against the background of one particular (configuration of the) world, whereas on the representational view, different models are taken to stand for different possible worlds against the background of one particular (interpretation of the) language.

By way of example, consider Mary's world shown in figure 1. This world only consists of a cube (c), a cylinder (y), a torus (t), and a sphere (s). The former two are made of wood, the latter two of marble. To keep things simple, we talk about Mary's world by means of a very austere sublanguage of FOL, $\mathcal{L}_M$, which, apart from the FOL constants, only has the following categorematic terms: two 1-place predicates F and G, one 2-place predicate R, and two names a and b. Let us construct two models for $\mathcal{L}_M$ by defining two different functions $\Im_1$ and $\Im_2$ over a domain that maps onto Mary's world, i.e. over $D_M = \{\mathsf{c}, \mathsf{y}, \mathsf{t}, \mathsf{s}\}$:

$$\begin{aligned} \Im_1(a) = \mathsf{c}, \;\; \Im_1(b) = \mathsf{t}, \;\; \Im_1(F) = \{\mathsf{c}, \mathsf{y}\}, \;\; \Im_1(G) = \{\mathsf{t}, \mathsf{s}\} \\ \Im_1(R) = \{\langle \mathsf{c}, \mathsf{y}\rangle, \langle \mathsf{c}, \mathsf{t}\rangle, \langle \mathsf{c}, \mathsf{s}\rangle, \langle \mathsf{y}, \mathsf{t}\rangle, \langle \mathsf{y}, \mathsf{s}\rangle, \langle \mathsf{t}, \mathsf{s}\rangle\} \end{aligned} \tag{$\mathfrak{M}_1$}$$

$$\begin{aligned} \Im_2(a) = \mathsf{s}, \;\; \Im_2(b) = \mathsf{y}, \;\; \Im_2(F) = \{\mathsf{t}, \mathsf{s}\}, \;\; \Im_2(G) = \{\mathsf{c}, \mathsf{y}\} \\ \Im_2(R) = \{\langle \mathsf{y}, \mathsf{c}\rangle, \langle \mathsf{t}, \mathsf{c}\rangle, \langle \mathsf{s}, \mathsf{c}\rangle, \langle \mathsf{t}, \mathsf{y}\rangle, \langle \mathsf{s}, \mathsf{y}\rangle, \langle \mathsf{s}, \mathsf{t}\rangle\} \end{aligned} \tag{$\mathfrak{M}_2$}$$

[4] Similarly, Shapiro (1998, 137).

$\Im_1$ and $\Im_2$ assign different extensions to the categorematic terms of $\mathcal{L}_M$. Varying extensions of the categorematic terms of a language may have two distinct sources: they may be due (i) to variations of corresponding intensions or (ii) to variations in corresponding matters of fact. The interpretational view ascribes the extensional differences between $\mathfrak{M}_1$ and $\mathfrak{M}_2$ to (i) whereas the representational view ascribes them to (ii). More specifically, on the interpretational view, $\mathfrak{M}_1$ and $\mathfrak{M}_2$ are taken to reflect two different intensional interpretations of the language $\mathcal{L}_M$, which is applied to one and the same world, i.e. Mary's world, which is given independently of $\mathfrak{M}_1$ and $\mathfrak{M}_2$. On the representational view, the intensional interpretation of $\mathcal{L}_M$ is given independently of $\mathfrak{M}_1$ and $\mathfrak{M}_2$ (e.g. in a dictionary), and $\mathfrak{M}_1$ and $\mathfrak{M}_2$ stand for two different worlds to which $\mathcal{L}_M$ is applied—only one of which being Mary's world (the other being Helge's world in figure 2 below).

To bring this contrast out more vividly, it is helpful to consider translations of $\mathcal{L}_M$ into English. Such translations can be represented by functions that assign expressions of natural English to the terms of $\mathcal{L}_M$. Translations of the logical constants of a formal language are straightforward: logical constants have fixed translations into English that are independent of models and possible worlds. Translations of the categorematic terms, however, give rise to some intricacies. Moreover, as will become apparent shortly, translations of categorematic terms must be understood as very different sorts of functions in the interpretational and representational frameworks. Accordingly, I shall speak of *interpretational* and *representational translations*, respectively. An interpretational translation of the categorematic terms of $\mathcal{L}_M$, on the formal side, depends on which objects a given $\Im_k$ assigns to those terms, and on the informal side, it depends on the objects and the properties constituting the world w_k to which English is applied. More specifically, an interpretational translation assigns a singular term of English to a such that the reference of this singular term in w_k corresponds to the object of D_k assigned to a by $\Im_k$; moreover, it assigns an English predicate to F such that the extension of that predicate consists exactly in the objects of w_k which correspond to the elements of D_k assigned to F by $\Im_k$. For instance, relative to Mary's world and relative to $\Im_1(F) = \{\mathsf{c}, \mathsf{y}\}$, F can be translated in terms of "…is made of wood", because being made of wood is the property the cube and the cylinder (and only these two objects) share in Mary's world. That is, in order to interpretationally translate the categorematic terms of a formal language under a given $\Im_k$ of a model $\mathfrak{M}_k$ into English two things are required: (i) a specification of a possible world w_k and (ii) a one-to-one mapping σ_k of the elements of the domain D_k onto the objects of w_k. Overall, an appropriate interpretational translation of the categorematic terms of $\mathcal{L}_M$ as interpreted in $\mathfrak{M}_1$ relative to Mary's world w_M and relative to the mapping given above, call it σ_M, would be the following, where the objects of Mary's world are numbered from left to right:

$$\begin{aligned} \mathcal{T}_1(a) &= \text{"object \#1"} \\ \mathcal{T}_1(b) &= \text{"object \#3"} \\ \mathcal{T}_1(F) &= \text{"\ldots is made of wood"} \\ \mathcal{T}_1(G) &= \text{"\ldots is made of marble"} \\ \mathcal{T}_1(R) &= \text{"\ldots is smaller than\ldots"} \end{aligned} \qquad (\mathcal{T}_1)$$

Now, take an exemplary sentence of $\mathcal{L}_M$, say:

$$Fa \wedge Gb \wedge Rab \qquad (1)$$

(1) is true in $\mathfrak{M}_1$, and according to $\mathcal{T}_1$, (1) claims:

(b) Object #1 is made of wood, and object #3 is made of marble, and object #1 is smaller than object #3.

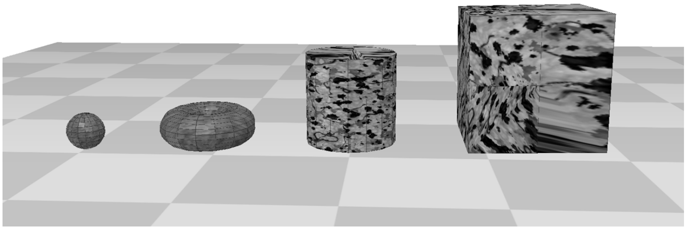

Fig. 2 Helge's world

In $\mathfrak{M}_2$, (1) comes out true as well. According to the interpretational view, (1) does not claim (b) relative to $\mathfrak{M}_2$, because $\mathfrak{M}_2$ supplies a new interpretation of $\mathcal{L}_M$. That is, an interpretational translation assigns different English expressions to $\mathcal{L}_M$ under $\Im_2$ relative to w_M and σ_M. An appropriate translation of the categorematic terms of $\mathcal{L}_M$ as interpreted in $\mathfrak{M}_2$ would be this:

$$\begin{aligned} \mathcal{T}_2(a) &= \text{“object \#4”} \\ \mathcal{T}_2(b) &= \text{“object \#2”} \\ \mathcal{T}_2(F) &= \text{“\ldots is made of marble”} \\ \mathcal{T}_2(G) &= \text{“\ldots is made of wood”} \\ \mathcal{T}_2(R) &= \text{“\ldots is larger than\ldots”} \end{aligned} \qquad (\mathcal{T}_2)$$

That is, against the background of $\Im_2$, (1) claims:

(c) Object #4 is made of marble, and object #2 is made of wood, and object #4 is larger than object #2.

According to the interpretational view, the transition from $\mathfrak{M}_1$ to $\mathfrak{M}_2$ is to be described as a swapping of the notions of being-made-of-wood and being-made-of-marble as well as of the smaller-than and larger-than relations. More generally put, in order to formally model* logical features of English from an interpretational perspective, the categorematic terms of the corresponding formal language under two different $\Im_n$ and $\Im_m$ must be translated into *different* English expressions.

In contrast, on the representational view, $\mathfrak{M}_2$ is not seen to reinterpret $\mathcal{L}_M$. Rather, it is taken to reconfigure Mary's world. $\mathfrak{M}_2$ represents another possible world, *viz.* Helge's world depicted in figure 2. In this world, the sphere takes the position, the size, and the makeup of the cube in Mary's world, and vice versa. Analogously, the cylinder and the torus switch position, size, and makeup. As $\mathfrak{M}_2$ does not change the meanings (intensions) assigned to the terms of $\mathcal{L}_M$, the translation for $\Im_2$ is the same as the translation for $\Im_1$. Accordingly, (1) makes exactly the same claim relative to $\Im_2$ as it does relative to $\Im_1$. On the representational view, a language is given one fixed interpretation independently of its models and independently of the world to which that language is applied. To properly model* this, we need a translation of $\mathcal{L}_M$ that is identical for $\Im_1$ and $\Im_2$. This can be generalized for any representational translation: all representational translations constantly assign *the same* English expressions to the categorematic terms of a formal language across all models of the latter. Section 4 will show that representational translations of formal expressions are normally simply given by a suitable correspondence scheme (or dictionary or key) (cf. Brun 2004, ch. 6.1). Depending on such a correspondence scheme, (1) is either constantly translated in terms of $\mathcal{T}_1$ or of $\mathcal{T}_2$, to the effect that (1) either constantly claims (b) or (c), which are both true in Mary's and in Helge's world.

Etchemendy (1990) introduces the distinction between the interpretational and representational views in the context of discussing Tarski's famous analyses of logical truth and consequence (cf. Tarski 1956), which Etchemendy reads as adopting an entirely interpretational perspective. As is well-known, his far-reaching criticism of Tarski's analyses has itself been criticized (cf. e.g. Ray 1996; Sher 1996; Gómez-Torrente 1996). While that controversy is not relevant for our present purposes, it is important to emphasize that contemporary model theory cannot be read in a purely interpretational way as originally conceived by Etchemendy. FOL models may not only differ with respect to the function that assigns elements and sets of elements from the domain D to the terms of FOL, but also with respect to D itself. Model theory systematically varies the cardinality of D up to infinity. If the elements of D, as is usual, are read as representing the objects FOL sentences quantify over, then two models that—unlike $\mathfrak{M}_1$ and $\mathfrak{M}_2$—not only differ in $\Im$ but also in D must be read as representing two different interpretations of FOL relative to two worlds that are constituted by different objects, that is, relative to two *different* worlds. Therefore, a strictly interpretational reading of models, according to which different models exclusively change the interpretation of FOL, does not yield an adequate understanding of the whole model-theoretic machinery.

In light of this, the difference between the interpretational and representational perspectives must not be understood in the broad way originally suggested by Etchemendy (1990). Rather, the difference only concerns the reading of one constituent of models, *viz.* of the function $\Im$. The interpretational view takes two different $\Im_n$ and $\Im_m$ that are defined for the same domain D as reflecting two different meaning assignments to the categorematic terms of FOL relative to D. By contrast, the representational view takes two different $\Im_n$ and $\Im_m$ that are defined for the same domain D as representations of two different FOL-describable possible worlds. Both views agree that two models $\mathfrak{M}_n$ and $\mathfrak{M}_m$ that involve two different domains D_n and D_m must be seen to represent different worlds. Still, while according to the representational account $\mathfrak{M}_n$ and $\mathfrak{M}_m$ represent two worlds in as much detail as is expressible in FOL (or a relevant sublanguage thereof), interpretationally read models $\mathfrak{M}_n$ and $\mathfrak{M}_m$ stand for any two worlds with the cardinality of D_n and D_m, respectively. On an interpretational reading, models never represent worlds in all their FOL-expressible particularity. Interpretationally read models only symbolize the cardinality of the worlds to which FOL is applied.[5]

It is evident from the above considerations that the interpretational view is best suited to model* (I)-validity, whereas the representational view suggests itself for modeling* (II)-validity. To make this more concrete, consider an argument $\langle \Gamma, \Psi \rangle$ and let $\mathfrak{M}^{\langle \gamma, \psi \rangle}$ designate a model of the FOL sublanguage constituted by the categorematic terms occurring in *FOL formalizations* γ of Γ and ψ of Ψ. Against this background, the set of permissible reinterpretations of $\langle \Gamma, \Psi \rangle$ relative to which $\langle \Gamma, \Psi \rangle$ must be truth-preserving in order to be (I)-valid can be modeled* on the basis of interpretationally read models $\mathfrak{M}^{\langle \gamma, \psi \rangle}$. Analogously, the realm of possible worlds in which $\langle \Gamma, \Psi \rangle$ must be truth-preserving in order to be (II)-valid can be modeled* by the set of representationally read models $\mathfrak{M}^{\langle \gamma, \psi \rangle}$. Accordingly, I subsequently also refer to (I)-validity as *interpretational validity* and to (II)-validity as *representational validity*.

In order to implement FOL as a means to model* or formally exhibit the interpretational or representational validity of an argument $\langle \Gamma, \Psi \rangle$, we not only need suitable readings of

[5] Essentially, this is what Shapiro (2005, 663) dubs the *blended* view of models. I prefer to stick to Etchemendy's original label "interpretational view", because it emphasizes the core difference from the representational view.

FOL models but also an adequate FOL formalization $\langle\gamma,\psi\rangle$ of $\langle\Gamma,\Psi\rangle$. Yet, since (I)- and (II)-validity differ extensionally and FOL fixes one determinate extension for the notion of formal validity, some arguments must be reproduced by different FOL formalizations depending on whether the goal is to exhibit interpretational or representational validity. Suppose that $\langle\Gamma,\Psi\rangle$ is (I)-invalid and (II)-valid. On the one hand, to exhibit its (I)-invalidity, $\langle\Gamma,\Psi\rangle$ must be formalized by an argument scheme $\langle\gamma,\psi\rangle$ that is FOL-invalid; on the other hand, to capture its (II)-validity $\langle\Gamma,\Psi\rangle$ must be formalized by an FOL-valid $\langle\gamma',\psi'\rangle$. That is, what counts as an adequate formalization of $\langle\Gamma,\Psi\rangle$ depends on whether the goal of formalizing $\langle\Gamma,\Psi\rangle$ is to account for interpretational or representational validity.

Accordingly, we need different criteria of adequate formalization, interpretational and representational criteria, which yield two different notions of adequate formalization, interpretational and representational adequacy. The next section develops interpretational criteria.

3 Exhibiting interpretational validity

Before discussing the details of exhibiting interpretational validity we need to set one question aside: what kind of entities are premises and conclusions of arguments? There are many candidates, for instance, beliefs, judgments, propositions, utterances, interpreted sentences, etc. (cf. Brun 2008; Russell 2008). Satisfactory answers to that question are hard to come by and I shall not attempt to address it here. Instead, I simply characterize premises and conclusions as being composed of declarative and de-contextualized *statements* (which in case of FOL formalizable arguments are moreover required to be extensional and bivalent), and invite the reader to view that label as a placeholder for whichever entity happens to be her preferred candidate category. Nonetheless, a crucial qualification is necessary at this point. While the constituents of representationally read arguments are statements with a fixed interpretation (or fully interpreted statements), the statements that make up interpretationally read arguments allow for reinterpretations. From the interpretational perspective, arguments are, from the outset, seen to be composed of statement schemata or representatives of logical forms rather than of statements with a fixed interpretation.

Conceiving of the statements constituting an argument's premises Γ and conclusion Ψ as representatives of logical forms suggests an approach to logical formalization according to which adequate formalizations of Γ and Ψ can, in one way or another, be read off the latter's natural language surface. For brevity, I shall speak of the *surface approach* in the following. As interpretational validity is the standard textbook precisification of informal validity, the surface approach is the standard textbook approach to logical formalization. In a nutshell, formalizing a statement Φ along its lines involves a stepwise procedure that gradually abstracts from descriptive content (cf. e.g. Ray 1996, sect. 2): (1) the expressions in Φ are partitioned into a set V of variable (reinterpretable) expressions and a set C of expressions that correspond to logical constants of the target formalism—FOL in our case; (2) the members of V are partitioned into semantic categories (names, predicates, relations etc.); (3) the elements of each semantic category are replaced by FOL terms of the corresponding semantic category and the elements of C are replaced by corresponding FOL constants. According to the surface approach, a formalization $\langle\gamma,\psi\rangle$ is said to be adequate for an argument $\langle\Gamma,\Psi\rangle$ iff γ and ψ reproduce the logical forms Γ and Ψ represent.

The surface approach, however, faces one major problem: the surface of natural language statements is often so misleading in regard to underlying logical forms that the latter cannot be reliably read off the former's surface. There are countless well-known examples illustrating this so-called *misleading form thesis* (cf. e.g. Brun 2004, 161-165; Sainsbury 2001, 44-53). For instance, the surface of “The whale is a mammal” misleadingly suggests that it is of the same form as “The king is a Frenchman”, which, as far as a Russellian analysis of definite descriptions is concerned, misleadingly suggests to be of subject-predicate form. Or “Humans are mortal”, which is commonly viewed to be of universal conditional form, fails to contain “if... then” or “all” as logical constants. In the end, only regimented non-natural idioms whose grammar is purposefully designed to reflect logical forms can be adequately formalized on the basis the surface approach.

To formalize non-regimented arguments, i.e. arguments outside of artificial textbook contexts, another approach is called for. The literature provides such another approach, which I shall dub the *correctness approach*. Rather than formalizing a statement Φ by taking its natural language form as a starting point, representatives of the correctness approach (cf. Blau 1977; Sainsbury 2001; Brun 2004; Baumgartner and Lampert 2008) formalize Φ on the basis of its truth conditions. They hold that an adequate formalization ϕ of Φ, first and foremost, must be *correct* for Φ, where ϕ is correct for Φ iff every model of ϕ in which ϕ is true models* a truth condition of Φ and every model of ϕ in which ϕ is false models* a falsehood condition of Φ.[6]

However, if truth conditions of a statement, as is usual, are understood as the *factual* conditions under which that statement is true, correct formalizations as defined by the correctness approach are only revealing with respect to whether the conclusion of an argument is true under all factual conditions under which the premises are true, i.e. with respect to whether that argument is representationally valid. Correct formalizations in that sense are of no help to determine whether the premises and conclusion are truth-preserving under all reinterpretations. That is, while the approach to logical formalization that is standardly used for exhibiting interpretational validity—the surface approach—is only applicable to regimented (non-natural) languages, the standard approach to formalizing non-regimented languages—the correctness approach—is custom-built for exhibiting representational validity.

In order to formalize non-regimented arguments in a way that suits interpretational purposes I subsequently adapt the correctness approach to the interpretational context. For clarity, I dub the representational (truth-conditional) variant of correctness sketched above (and mainly implemented in the next section) *r-correctness*, and to the interpretational variant developed in what follows I refer as *i-correctness*. The basic idea behind i-correctness can be easily stated: a formalization ϕ is i-correct for a statement Φ iff different models of ϕ model* different permissible reinterpretations of Φ. To make that idea more explicit and precise, some conceptual preliminaries are required.

The previous section has shown that the connection between a formalization ϕ and reinterpretations of a statement Φ is established by translations of the FOL-sublanguage $\mathcal{L}^{\phi}$ comprising the categorematic terms of ϕ relative to specific interpretations of $\mathcal{L}^{\phi}$ given in models $\mathfrak{M}_k^{\phi}$. We have also seen that interpretationally read models only specify the cardinality of corresponding worlds, which vastly underdetermines English translations of models. Therefore, to translate the categorematic terms of an FOL formula ϕ into English it is necessary that the domains of its models $\mathfrak{M}_k^{\phi}$ be mapped onto possible worlds. For a model $\mathfrak{M}_k^{\phi}$ such a domain-to-world mapping is guaranteed to exist if its domain D_k has the same cardinality as the possible world w_{Φ} in which Φ is intended to be evaluated. To the pair $\langle \Phi, w_{\Phi} \rangle$

[6] Cf. e.g. Brun (2004, 210); Baumgartner and Lampert (2008, 108). Correctness can also be defined syntactically. In this paper, I am only going to use the semantic variant of correctness. For details on the syntactic one cf. Baumgartner and Lampert (2008).

there corresponds the pair $\langle \phi, \mathfrak{M}_I^\phi \rangle$ constituted by a formalization ϕ of Φ and by the model $\mathfrak{M}_I^\phi$ featuring the so-called *intended interpretation* of ϕ. The intended interpretation $\Im_I$ of ϕ assigns to the names of ϕ the entities that correspond to the entities to which the names of Φ refer in w_Φ, and to the predicates of ϕ the sets of things that correspond to the extensions of the predicates of Φ in w_Φ (cf. Sainsbury 2001, 162). The relevant correspondence between $\langle \Phi, w_\Phi \rangle$ and $\langle \phi, \mathfrak{M}_I^\phi \rangle$ is established by the *correspondence scheme* that comes with each formalization. Correspondence schemes moreover identify the components of the statement Φ that are treated as reinterpretable units by a given formalization ϕ.[7]

That is, all $\mathfrak{M}_k^\phi$ over the same domain as $\mathfrak{M}_I^\phi$ have suitable domain-to-world mappings. For instance, in case of statement (b) about Mary's world, all models of the FOL sublanguage $\mathcal{L}_M$ with the domain $D_M = \{\mathsf{c}, \mathsf{y}, \mathsf{t}, \mathsf{s}\}$ can be mapped onto a possible world, *viz.* Mary's world. Also, all models with domains that result from D_M by removing objects can be mapped onto possible worlds that result from Mary's world by removing corresponding objects. By contrast, to translate $\mathcal{L}_M$ relative to models with a domain D_k that is a proper superset of D_M specifications of the objects contained in D_k but not in D_M must additionally be supplied. Such specifications can be given by simple conventions, for instance, to the effect that all D_k that are proper supersets of D_M correspond to worlds that result from Mary's world by adding wooden cubes of size t_1 or marble spheres of size t_2 etc; or they can be given by complex descriptions of resulting worlds.

Apart from a suitable domain-to-world mapping, the translatability of a formal language, relative to its models, into English also depends on the availability of English names that corefer with their formal correlates and English predicates that are co-extensional with their formal correlates; and for many models that availability is dubious, to say the least. The space of possible reinterpretations provided by the model-theoretic machinery far exceeds the expressive power of natural English. Yet, we do not need to require that translations may only involve English names and predicates that are listed in standard English dictionaries. For the purposes of logical formalization artificial English names and predicates may be introduced whenever dictionaries do not provide sufficient expressiveness. While artificial names can simply be generated by numbering the objects in a corresponding world, artificial predicates can be defined by means of (e.g. Boolean) functions of existing English predicates. As long as artificial predicates have explicit definitions in terms of existing predicates, the truth values of statements featuring those artificial predicates can be informally assessed, which, as we shall see below, is all that is of relevance for the i-correctness of logical formalizations.

Overall, we can hence stipulate that for a formalization ϕ of a statement Φ there exists a translation $\mathcal{T}_k$ of $\mathcal{L}^\phi$ relative to a model $\mathfrak{M}_k^\phi$ iff there exists a mapping of $\mathfrak{M}_k^\phi$'s domain D_k onto a possible world w_k. I subsequently call such a model $\mathfrak{M}_k^\phi$ *w_k-mappable*. A translation $\mathcal{T}_k$ allows for *verbalizing* a formalization ϕ of Φ (cf. Brun 2004, ch. 10). For brevity, I shall speak of *$\mathcal{T}_k$-verbalizations*. Against this conceptual background, we can now spell out i-correctness in more detail:

(i-COR) The formalization ϕ is i-correct for statement Φ iff for all w_k-mappable models $\mathfrak{M}_k^\phi$ and all corresponding worlds w_k: there exists a translation $\mathcal{T}_k$ of $\mathcal{L}^\phi$, as interpreted in $\mathfrak{M}_k^\phi$, such that $\mathcal{T}_k$-verbalizing ϕ yields a permissible reinterpretation Φ' of Φ, and ϕ has the same truth value in $\mathfrak{M}_k^\phi$ as Φ' in w_k.

[7] Correspondence schemes have a very different function in the representational framework. In that framework they fix the interpretation of formal languages. Cf. sect. 4 below.

(i-COR) is a necessary (but not a sufficient) condition for the interpretational adequacy of a formalization. Moreover, according to (i-COR), assessing the i-correctness of ϕ for Φ presupposes that it be informally determinable whether $\mathcal{T}_k$-verbalizations of ϕ amount to permissible reinterpretations of Φ. More generally, as anticipated in section 2, it is a precondition of the interpretational formalizability of Φ that it be clear for any statement Φ' whether Φ' permissibly reinterprets Φ or not. From the interpretational perspective, clarity on whether Φ' is a permissible reinterpretation of Φ is the minimal amount of *informal understanding* of Φ that must be available in order to formalize Φ. As indicated in section 2, I assume that the statements formalized in this paper are informally understood to that necessary degree.

Let me illustrate (i-COR) by way of example. First, consider statement (d), which is intended to be about the actual world $w_@$, and the formalization candidates (2), (3), and (4) with correspondence scheme (c_1).

(d) The whale is a mammal.

$$p \tag{2}$$

$$Ga \tag{3}$$

$$\forall x(Fx \rightarrow Gx) \tag{4}$$

$$\begin{array}{c} p : \text{The whale is a mammal} \ ; \quad a : \text{the whale} \ ; \\ F : \ \ldots \text{is a whale} \ ; \quad G : \ \ldots \text{is a mammal} \end{array} \tag{c_1}$$

That (2) is i-correct for (d) can be seen as follows. There are merely two models for the FOL sublanguage consisting of the sentence letter p only: $\mathfrak{M}_1^{(2)}$ comprising $\Im_1(p) = T$ and $\mathfrak{M}_2^{(2)}$ comprising $\Im_2(p) = F$. As (d) is true in $w_@$, $\Im_1(p) = T$ is the intended interpretation of (2). Verbalizing (2) based on the correspondence scheme (c_1) yields (d) itself, which trivially amounts to a permissible (re)interpretation of (d) with the same truth value in $w_@$ as (2) in $\mathfrak{M}_1^{(2)}$. Moreover, relative to $\mathfrak{M}_2^{(2)}$, (2) can be verbalized based on any translation of $\mathcal{L}^{(2)}$ that issues a statement that is false in $w_@$. Plainly, at least one of those translations generates a permissible reinterpretation of (d), for instance the translation that assigns "The elephant is a bird" to p. Accordingly, (2) is i-correct for (d).

By contrast, (3) is not i-correct for (d), for there are countless models $\mathfrak{M}_k^{(3)}$ whose domain D_k can be mapped onto $w_@$ but for which there does not exist a translation $\mathcal{T}_k$ of $\mathcal{L}^{(3)}$ such that $\mathcal{T}_k$-verbalizing (3) yields a permissible reinterpretation of (d). For example, take a model $\mathfrak{M}_k^{(3)}$ that assigns the object that corresponds to Obama to a and the singleton of that same object to G. Conceivable translations will generate verbalizations as "Obama is the 44th president of the US" or "Obama is the first African-American president of the US" etc., none of which are informally judged to be permissible reinterpretations of (d).

Finally, verbalizing (4) by virtue of (c_1) yields "For all objects, x, if x is a whale then x is a mammal". Even though this statement has a considerably different grammatical surface than (d), it is ordinarily judged to permissibly (re)interpret (d). The same holds for verbalizations of (4) based on other translations of $\mathcal{L}^{(4)}$, e.g. for "For all objects, x, if x is brown then x is a table" or "For all objects, x, if x is born in Dietramszell then x has blood type B^+". Just as (d), these statements claim that a first set is contained in a second set. This can be generalized: for all w_k-mappable models $\mathfrak{M}_k^{(4)}$ and all corresponding worlds w_k, there exists a translation $\mathcal{T}_k$ of $\mathcal{L}^{(4)}$ such that a $\mathcal{T}_k$-verbalization of (4) yields a permissible reinterpretation of (d) that has the same truth value in w_k as (4) in $\mathfrak{M}_k^{(4)}$. Thus, (4) is i-correct for (d).

In light of the fact that the notion of a permissible reinterpretation of a statement Φ is an essentially informal notion, clarity on whether a given Φ' counts as a permissible reinterpretation of Φ is clearly not an innocuous presupposition of i-correctness. Nonetheless, if it is indeterminate whether "Obama is the 44th president of the US" or "For all objects, x, if x is brown then x is a table" amount to permissible reinterpretations of (d), the i-correctness of (3) and (4) for (d) is indeterminate as well. Formalizing natural language statements presupposes that certain features of those statements are informally understood: while interpretational formalizations require clarity on permissible reinterpretations, representational formalizations, as we shall see in the next section, presuppose clarity on the (factual) truth conditions of statements. Relative to different informal understandings of a particular statement Φ, different formalizations will count as i-correct for Φ.

Furthermore, note that rigorously applying (i-COR) in the course of assessing the i-correctness of, say, (4) for (d) would amount to confronting an unmanageable amount of models of (4) with corresponding reinterpretations of (d), which is a task that obviously cannot be completed.[8] In practice, however, the i-correctness of formalization candidates of (d) can be reliably determined based on a comparison of the relevant structural features of the w_k-mappable models of a formalization and the relevant structural features of the resulting permissible reinterpretations of (d). For instance, (4) and all of its verbalizations state a subset relation between two sets. That is, (4) is true in a model $\mathfrak{M}_k^{(4)}$ iff $\Im_k(F) \subseteq \Im_k(G)$ holds in $\mathfrak{M}_k^{(4)}$; analogously, permissible reinterpretations of (d) are true in a corresponding world w_k iff the extension of the first predicate is contained in the extension of the second predicate. Hence, instead of laboriously confronting every model of (4) with every resulting reinterpretation of (d), the i-correctness of (4) for (d) can be established based on *structural descriptions* of the relevant features of the corresponding models and reinterpretations.

I now turn to formalizing (very simple) arguments on the basis of (i-COR). In accordance with the notational convention adopted so far in this paper, I formalize an argument $\langle \Gamma, \Psi \rangle$ in terms of a pair $\langle \gamma, \psi \rangle$ where γ represents a set of FOL sentences that correspond to the premises of the argument and ψ represents an FOL sentence that corresponds to its conclusion. For simplicity, I say that $\langle \gamma, \psi \rangle$ and $\langle \Gamma, \Psi \rangle$ are *true* in a model $\mathfrak{M}_k^{\langle \gamma, \psi \rangle}$ and a corresponding world w_k, respectively, if they are truth-preserving in $\mathfrak{M}_k^{\langle \gamma, \psi \rangle}$ and w_k, respectively, and *false* otherwise. Moreover, the domains of the intended interpretations of all subsequently discussed formalization candidates are (implicitly) assumed to be mapped onto the actual world $w_@$. First, consider the trivially interpretationally valid example (e) with formalization candidates (5) and (6) and correspondence scheme (c_2):

(e) Daryl is a mother. Therefore, Daryl is a mother.

$$\langle \{Fa\} \text{ , } Ga \rangle \tag{5}$$

$$\langle \{Ga\} \text{ , } Ga \rangle \tag{6}$$

$$a : \text{Daryl} \text{ ; } \quad F : \ldots \text{is a mother} \text{ ; } \; G : \ldots \text{is a mother} \qquad (c_2)$$

Even though both Fa and Ga are i-correct for the atomic statement "Daryl is a mother", (5) is not i-correct for (e), for verbalizing (5) based on most translations of $\mathcal{L}^{(5)}$ does not yield permissible reinterpretations of (e). In most models $\mathfrak{M}_k^{(5)}$, $\Im_k(F)$ differs from $\Im_k(G)$, whereas reinterpretations of the instance of "...is a mother" in the premise of (5) must not differ from reinterpretations of the instance of "...is a mother" in the conclusion. The two instances of "...is a mother" constitute one reinterpretable unit in the context of argument (e). Arguments are *complex* statements and their complexity constrains the reinterpretability of the atomic statements of which they are composed. Multiple instances of one and the same reinterpretable unit must always be reinterpreted jointly. This constraint is respected

[8] As we shall see in section 4, (r-COR) faces an analogous termination problem. Cf. also Baumgartner and Lampert (2008, 97).

in (6). Verbalizing (6) based on translations of $\mathcal{L}^{(6)}$ relative to models $\mathfrak{M}_k^{(6)}$ never yields different reinterpretations of the two instances of "…is a mother" in (e). (6) is i-correct for (e).

Next, consider the more interesting argument (f) with formalization candidates (7), (8), (9) and correspondence scheme (c_3):

(f) Daryl strolls slowly. Therefore, Daryl strolls.

$$\langle \{Fa\}\ ,\ Ga \rangle \tag{7}$$

$$\langle \{Ga \wedge La\}\ ,\ Ga \rangle \tag{8}$$

$$\langle \{\exists x(Jx \wedge Hxa \wedge Lx)\}\ ,\ \exists x(Jx \wedge Hxa) \rangle \tag{9}$$

$$\begin{array}{c} a : \text{Daryl}\ ;\ \ F : \ldots\text{strolls slowly}\ ;\ G : \ldots\text{strolls}\ ;\ J : \ldots\text{is a stroll;} \\ H : \ldots\text{is conducted by}\ \ldots\ ;\ \ L : \ldots\text{is slow} \end{array} \tag{c_3}$$

(f) is commonly considered to be interpretationally valid. That means "…strolls slowly" and "…strolls" are informally judged not to be two independently reinterpretable units. (7), however, imposes no constraints on the reinterpretability of "…strolls slowly" and "…strolls". Verbalizing (7) based on some of the translations of $\mathcal{L}^{(7)}$ yields arguments none of which pass as permissible reinterpretations of (f), e.g. "Clooney is a bachelor. Therefore, Clooney is a mother". Hence, (7) is not i-correct for (f). By contrast, as (8) is FOL-valid, verbalizations of (8) do not generate independent reinterpretations of the premise and conclusion of (f). Nonetheless, (8) fails to be i-correct for (f) because it does not i-correctly represent (f)'s premise. "Daryl strolls slowly" does not predicate two independent properties of Daryl, yet verbalizing $Ga \wedge La$ on the basis of many translations of $\mathcal{L}^{(8)}$ will result in statements all of which predicate independent properties of pertaining objects, e.g. "Obama is married and Obama is a politician". Finally, (9) draws on Davidson's (1967) celebrated analysis of action sentences. Verbalizing Davidson-style formalizations of statements involving adverbial predication is generally judged to generate permissible reinterpretations of the latter. Moreover, the FOL validity of (9) imposes constraints on the reinterpretability of the premise and conclusion of (9) that match the corresponding constraints informally ascribed to (f). As a consequence, the truth values of (9) in w_k-mappable models $\mathfrak{M}_k^{(9)}$ agree with the truth values of resulting reinterpretations of (f) in w_k. Overall, (9) is i-correct for (f).

This raises the question whether i-correctness and FOL validity is all that we need to require of a formalization $\langle \gamma, \psi \rangle$ in order to formally exhibit the interpretational validity of an argument $\langle \Gamma, \Psi \rangle$. That this question must be answered in the negative can be seen on the basis of the following example:

(g) Daryl is a mother. Therefore, Daryl is a woman.

$$\langle \{Ga\}\ ,\ Ga \rangle \tag{10}$$

$$\langle \{Fa\}\ ,\ Ga \rangle \tag{11}$$

$$a : \text{Daryl}\ ;\ F : \ldots\text{is a mother}\ ;\ G : \ldots\text{is a woman} \tag{c_4}$$

Verbalizations of (10) based on all conceivable translations of $\mathcal{L}^{(10)}$ generate co-intensional reinterpretations of "…is a mother" and "…is a woman". Plainly, co-intensional reinterpretations of these two predicates are permissible. Different instances of one reinterpretable unit must be reinterpreted jointly in arguments, but that does not entail that instances of different reinterpretable units must always be reinterpreted differently. That is, arguments as "Daryl is a woman. Therefore, Daryl is a woman" or "Clooney is from Dietramszell.

Therefore, Clooney is from Dietramszell" all pass as permissible reinterpretations of (g). Moreover, (10) has the same truth values in models $\mathfrak{M}_k^{(10)}$ as the resulting reinterpretations of (g) in w_k, *viz.* true. Hence, (10) is i-correct for (g).

The fact that the FOL-valid formalization (10) i-correctly reproduces (g), however, does not entail that (g) is interpretationally valid. Verbalizations of (10) only yield a very small subset of all permissible reinterpretations of (g), and obviously, that (g) preserves truth relative to that subset does not entail that it preserves truth relative to all permissible reinterpretations. And in fact, there are many permissible reinterpretations of (g) that do not preserve truth, e.g. "Clooney is a bachelor. Therefore, Clooney is from Dietramszell". Thus, interpretational validity of an argument $\langle \Gamma, \Psi \rangle$ cannot be inferred from the mere existence of an i-correct formalization $\langle \gamma, \psi \rangle$ that is FOL-valid. In order to establish the interpretational validity of $\langle \Gamma, \Psi \rangle$, an FOL-valid formalization $\langle \gamma, \psi \rangle$ must not only be i-correct but, in addition, it must be *exhaustive*, that is, it must model* the whole space of permissible reinterpretations of $\langle \Gamma, \Psi \rangle$. In the more general case of formalizations of statements, exhaustiveness amounts to this: The formalization ϕ of Φ is *exhaustive* for Φ iff verbalizing ϕ based on all translations of $\mathcal{L}^\phi$ yields all permissible reinterpretations of Φ.

Of course, if, as is normal, no overview over all permissible reinterpretations of a natural language statement Φ is on hand, the exhaustiveness of a given ϕ is not conclusively assessable. Nonetheless, analogously to approximatively establishing i-correctness of ϕ by virtue of balancing relevant structural features of models $\mathfrak{M}_k^\phi$ and corresponding reinterpretations of Φ, the exhaustiveness of ϕ can be approximated by establishing that all relevant structural features of the different permissible reinterpretations of Φ are represented among the verbalizations of ϕ. This condensed way of consolidating exhaustiveness suffices to rule out that (10) is exhaustive for argument (g). (10) does not yield any independent reinterpretations of the premise and conclusion of (g), even though such reinterpretations are informally judged to be permissible. Hence, (10) does not model* an important type of permissible reinterpretations of (g), and indeed a type relative to which (g) fails to preserve truth. (11), by contrast, is not only i-correct for (g) but moreover models* the whole space of permissible reinterpretations of argument (g), which is thus exhaustively formalized by (11).

In sum, in order to formally exhibit the interpretational validity of an argument $\langle \Gamma, \Psi \rangle$ an i-correct, exhaustive, and FOL-valid formalization $\langle \gamma, \psi \rangle$ is required. More specifically, as a consequence of the considerations of this section I propose the following principle for exhibiting interpretational validity:

(i-VP) An argument $\langle \Gamma, \Psi \rangle$ is exhibited to be interpretationally valid (I-valid) iff $\langle \Gamma, \Psi \rangle$ is formalized by an FOL-valid formalization $\langle \gamma, \psi \rangle$ that is i-correct and exhaustive for $\langle \Gamma, \Psi \rangle$.

For completeness, I shall say that a formalization ϕ is *interpretationally adequate*, i.e. *i-adequate*, for a statement Φ iff ϕ is both i-correct and exhaustive for Φ.

4 Exhibiting representational validity

As indicated in the introduction, even though interpretational validity is the more common precisification of informal validity, representational validity is an alternative precisification that has repeatedly been promoted in the pertinent literature. Moreover, the representational perspective has (implicitly or explicitly) driven a significant amount of the recent literature concerned with developing criteria of adequate logical formalization, e.g. Blau (1977), Sainsbury (2001), Brun (2004), Baumgartner and Lampert (2008).

In order to be revealing in regard to representational validity, logical formalizations of an argument $\langle \mathit{\Gamma}, \mathit{\Psi} \rangle$, first and foremost, must reproduce the (factual) truth conditions of $\mathit{\Gamma}$ and $\mathit{\Psi}$. For if a formalization $\langle \gamma, \psi \rangle$ that captures the truth conditions of the argument's premises and conclusion is FOL-valid, it follows that under all conditions that render $\mathit{\Gamma}$ true $\mathit{\Psi}$ is true as well, or in other words, that $\langle \mathit{\Gamma}, \mathit{\Psi} \rangle$ is truth-preserving in all possible worlds. As anticipated in the previous section, the central necessary condition for adequate representational formalization is correctness, or *r-correctness* as I shall call it. Roughly, a formalization ϕ is r-correct for a statement $\mathit{\Phi}$ iff ϕ and $\mathit{\Phi}$ have the same conditions of truth and falsehood. Before cashing that idea out in more detail, we again need to go through some preliminaries.

First, it must be emphasized that correspondence schemes play a very different role in the representational framework than in the interpretational one. From the representational perspective, a natural language statement $\mathit{\Phi}$ is viewed as a non-reinterpretable unit, i.e. as a unit with a fixed interpretation. Formally modeling* this perspective requires a formal language with an analogously fixed interpretation. It is the correspondence scheme, that comes with every formalization, which fixes the interpretation of the FOL sublanguage $\mathcal{L}^{\phi}$ that is used in a given formalization ϕ. To illustrate, consider the statement (h) with formalization (12) and correspondence scheme (c_5):

(h) Obama is married and Obama is a politician.

$$Fa \wedge Ga \tag{12}$$

$$a : \text{Obama} \ ; \ F : \ldots \text{is married} \ ; \ G : \ldots \text{is a politician} \tag{c_5}$$

Here are two exemplary models of the FOL sublanguage $\mathcal{L}^{(12)}$:

$$\begin{aligned} D_1 &= \{\mathsf{b}_1, \mathsf{b}_2, \mathsf{b}_3, \mathsf{b}_4\} \ , \\ \Im_1(a) = \mathsf{b}_1 \ , \ \ \Im_1(F) = \{\mathsf{b}_1, \mathsf{b}_3\} \ , \ \ \Im_1(G) &= \{\mathsf{b}_1, \mathsf{b}_4\} \end{aligned} \tag{$\mathfrak{M}_1^{(12)}$}$$

$$\begin{aligned} D_2 &= \{\mathsf{b}_1, \mathsf{b}_2, \mathsf{b}_3, \ldots, \mathsf{b}_{123}\} \ , \\ \Im_2(a) = \mathsf{b}_{13} \ , \ \ \Im_2(F) = \{\} \ , \ \ \Im_2(G) &= \{\mathsf{b}_1, \mathsf{b}_2, \mathsf{b}_3, \ldots, \mathsf{b}_{123}\} \end{aligned} \tag{$\mathfrak{M}_2^{(12)}$}$$

As we have seen in section 2, the representational view takes $\mathfrak{M}_1^{(12)}$ and $\mathfrak{M}_2^{(12)}$ to be specifications in $\mathcal{L}^{(12)}$ of two possible worlds. Interpreted in the vein of (c_5), $\mathfrak{M}_1^{(12)}$ describes a world constituted by 4 objects one of which being Obama, two of which being married and two of which being politicians. $\mathfrak{M}_2^{(12)}$ specifies a world constituted of 123 objects one of which being Obama, none of which being married and all of which being politicians. While $\mathfrak{M}_1^{(12)}$ and $\mathfrak{M}_2^{(12)}$ specify two different possible worlds, the interpretation of $\mathcal{L}^{(12)}$ remains the same, *viz.* the interpretation given in (c_5).

Second, note that $\mathfrak{M}_1^{(12)}$ and $\mathfrak{M}_2^{(12)}$ specify two possible worlds in as much detail as is expressible in $\mathcal{L}^{(12)}$. Moreover, the world-specifications provided by $\mathfrak{M}_1^{(12)}$ and $\mathfrak{M}_2^{(12)}$ suffice to determine whether statement (h) is true in a corresponding world: (h) is true in the world characterized by $\mathfrak{M}_1^{(12)}$ and false in the world characterized by $\mathfrak{M}_2^{(12)}$. Accordingly, every model $\mathfrak{M}_k^{\phi}$ of an r-correct formalization ϕ can be mapped onto a possible world w_k, *viz.* the world featuring the objects and properties specified in $\mathfrak{M}_k^{\phi}$, and thus models* a truth or falsehood condition of the statement $\mathit{\Phi}$. Requiring an r-correct formalization ϕ to agree with $\mathit{\Phi}$ in regard to truth and falsehood conditions, hence, amounts to requiring that the truth values of $\mathit{\Phi}$ and ϕ coincide relative to *all* models $\mathfrak{M}_k^{\phi}$ and corresponding worlds w_k.[9]

[9] In this respect the version of r-correctness presented here differs considerably from the versions developed by Blau (1977) and Brun (2004) who both only require correspondence of truth conditions relative to the subset of models $\mathfrak{M}_k^{\phi}$ that provide what Blau and Brun call *suitable interpretations* of ϕ. For a detailed criticism of this latter approach cf. Baumgartner and Lampert (2008).

With these characteristics of the representational framework in mind we can define r-correctness as follows:

(r-COR) The formalization ϕ is r-correct for statement Φ iff every model $\mathfrak{M}_k^\phi$ of ϕ specifies a possible world w_k in enough detail to determine the truth value of Φ in w_k, and ϕ has the same truth value in every model $\mathfrak{M}_k^\phi$ as Φ in w_k.

While the truth values of an FOL formula ϕ in different models are precisely defined on model-theoretic grounds, the truth values of a statement Φ in different possible worlds, i.e. Φ's truth conditions, must be assessed informally. That is, just as i-correctness presupposes that the permissibility of reinterpretations of Φ is informally determinate, r-correctness presupposes that the truth conditions of Φ are informally determinate. This is the sort of informal understanding of Φ that is required in order to representationally formalize Φ, and analogously to the case of interpretational formalizations, presupposing such an extensive informal understanding of Φ is not innocuous. However, if the truth conditions of Φ are informally dubious, Φ cannot be representationally formalized; or if the truth conditions of Φ are informally ambiguous, what counts as an r-correct formalization of Φ is equally ambiguous.

Let me illustrate (r-COR) by applying it to formalization candidates of statement (h). That (12) is r-correct for (h) can be established as follows: (12) is true in a model $\mathfrak{M}_k^{(12)}$ iff $\Im_k(a) \in \Im_k(F)$ and $\Im_k(a) \in \Im_k(G)$; analogously, (h) is true in a possible world w_k iff Obama has both the property of being married and the property of being a politician in w_k. Relative to (c_5), all models $\mathfrak{M}_k^{(12)}$ specify a possible world w_k in such a way that (h) has a determinate truth value in w_k, and (12) and (h) have the same truth values in all of these model-world pairs $\langle\mathfrak{M}_k^{(12)}, w_k\rangle$. Compare (12) with the formalizations (13), (14), and (15) of (h)—which are also to be interpreted in terms of (c_5):

$$Fa \vee Ga \tag{13}$$

$$Fa \rightarrow Ga \tag{14}$$

$$Fa \wedge Ga \wedge \forall x(Fx \rightarrow Gx) \tag{15}$$

None of these formalizations is r-correct for (h). In models $\mathfrak{M}_k^{(13)}$ and $\mathfrak{M}_k^{(14)}$ such that $\Im_k(a) \notin \Im_k(F)$ and $\Im_k(a) \in \Im_k(G)$, (13) and (14) are true, but (h) is false in worlds where Obama is an unmarried politician. Moreover, in a model $\mathfrak{M}_k^{(15)}$ such that $\Im_k(a) \in \Im_k(F)$, $\Im_k(a) \in \Im_k(G)$, and $\Im_k(F) \nsubseteq \Im_k(G)$, (15) is false, whereas (h) is informally judged to be true in a corresponding world w_k.

As in case of (i-COR), a comprehensive application of (r-COR) would amount to balancing the truth values of ϕ and Φ in an unmanageable amount of models and possible worlds, which is a task that cannot be completed. Yet, analogously to i-correctness, a complete application of (r-COR) can be abbreviated—as has been done above—by comparing structural descriptions of the models and possible worlds that render ϕ and Φ true and false, respectively.

Let us now turn to representationally formalizing arguments. First, reconsider example (f) with formalization candidates (7), (8), (9):

(f) Daryl strolls slowly. Therefore, Daryl strolls.

$$\langle\{Fa\}\ ,\ Ga\rangle \tag{7}$$

$$\langle\{Ga \wedge La\}\ ,\ Ga\rangle \tag{8}$$

$$\langle\{\exists x(Jx \wedge Hxa \wedge Lx)\}\ ,\ \exists x(Jx \wedge Hxa)\rangle \tag{9}$$

$$a: \text{Daryl}\ ;\ \ F: \ldots\text{strolls slowly}\ ;\ G: \ldots\text{strolls}\ ;\ J: \ldots\text{is a stroll};\quad H: \ldots\text{is conducted by}\ldots\ ;\ \ L: \ldots\text{is slow} \tag{c_3}$$

That (7) is not r-correct for (f) can be seen as follows: in models $\mathfrak{M}_k^{(7)}$ such that $\Im_k(a) \in \Im_k(F)$ and $\Im_k(a) \notin \Im_k(G)$, (7) is false[10], yet these models do not specify possible worlds relative to (c_3), for it is informally judged to be impossible for Daryl to stroll slowly without strolling. Hence, (7) does not model* the truth conditions of (f). By contrast, all models $\mathfrak{M}_k^{(8)}$—interpreted in the vein of (c_3)—can be said to specify possible worlds. Moreover, (f) is true in all of those worlds, for (f) is true in all possible worlds simpliciter. Correspondingly, (8) is true in all models $\mathfrak{M}_k^{(8)}$. Therefore, the FOL-valid (8) is r-correct for (f).

(9) is also r-correct for (f). Relative to (c_3), every model $\mathfrak{M}_k^{(9)}$ specifies a possible world w_k and (9) has the same truth value in every $\mathfrak{M}_k^{(9)}$ as (f) in the corresponding w_k, *viz*. true. That is, we have two unrelated r-correct FOL-valid formalizations of argument (f): (8) and (9). From the existence of any r-correct FOL-valid formalization $\langle \gamma, \psi \rangle$ of an argument $\langle \Gamma, \Psi \rangle$ it follows that $\langle \Gamma, \Psi \rangle$ is representationally valid. However, only one of those two formalizations actually exhibits the features of the premise and conclusion of (f) that are responsible for (f)'s representational validity, i.e. only one of those formalizations makes transparent why (f) is representationally valid, *viz*. (9). The reason is that only (9) r-correctly reproduces both the argument (f) as a whole, i.e. as a complex statement, as well as its premise and conclusion, i.e. its components. (8), on the other hand, does not r-correctly capture (f)'s premise, because not all models $\mathfrak{M}_k^{(8)}$, read in the vein of (c_3), specify possible worlds in enough detail to determine the truth value of (f)'s premise. For example, take a model $\mathfrak{M}_k^{(8)}$ such that $\Im_k(a) \in \Im_k(F)$ and $\Im_k(a) \in \Im_k(L)$. The corresponding world w_k is one where Daryl strolls and Daryl is slow. This specification is not sufficient to determine whether "Daryl strolls slowly" is true or false in w_k.

This shows that in order to exhibit the representational validity of an argument $\langle \Gamma, \Psi \rangle$, an FOL-valid formalization $\langle \gamma, \psi \rangle$ must not only be r-correct for $\langle \Gamma, \Psi \rangle$ as a whole, but all members of γ must moreover be r-correct for the corresponding members of Γ and ψ must be r-correct for Ψ. I shall speak of *recursive r-correctness* or *rr-correctness*, for short:

(rr-COR) The formalization $\langle \gamma, \psi \rangle$ is *recursively r-correct* (*rr-correct*) for the argument $\langle \Gamma, \Psi \rangle$ iff each member of γ is r-correct for the corresponding member of Γ, ψ is r-correct for Ψ, and $\langle \gamma, \psi \rangle$ is r-correct for $\langle \Gamma, \Psi \rangle$.

Note that γ and ψ each being r-correct for Γ and Ψ, respectively, is not sufficient for the r-correctness of $\langle \gamma, \psi \rangle$ for $\langle \Gamma, \Psi \rangle$—hence, the last conjunct of (rr-COR). To see this, consider the following formalization candidate of (f), which is likewise to be understood in the vein of (c_3):

$$\langle \{Fa\}\ ,\ \exists x(Jx \wedge Hxa) \rangle \tag{16}$$

(16) r-correctly accounts for both the premise and the conclusion of (f), but (16) as a whole is not r-correct for (f), because all of the models $\mathfrak{M}_k^{(16)}$ in which Fa is true and $\exists x(Jx \wedge Hxa)$ false do not represent possible worlds.

Next, reconsider argument (g) with two formalization candidates that have already been discussed in section 3—(10) and (11)—and three further candidates that are intended to illustrate peculiarities of representational formalizations:

(g) Daryl is a mother. Therefore, Daryl is a woman.

[10] According to the convention adopted in section 3, I say that arguments and their formalizations are *true* if they are truth-preserving in corresponding possible worlds and models, respectively, and *false* otherwise.

$$\langle\{Ga\}\ ,\ Ga\rangle \tag{10}$$

$$\langle\{Fa\}\ ,\ Ga\rangle \tag{11}$$

$$\langle\{Fa \wedge \forall x(Fx \rightarrow Gx)\}\ ,\ Ga\rangle \tag{17}$$

$$\langle\{Ka \wedge Ga\}\ ,\ Ga\rangle \tag{18}$$

$$\langle\{Ka \wedge Ga \wedge (p \vee \neg p)\}\ ,\ Ga\rangle \tag{19}$$

a : Daryl ; F : ...is a mother ; G : ...is a woman
K : ...has a child ; p : Clooney is a bachelor (c_6)

While (10) is i-correct for (g), it fails to capture (g) rr-correctly, because it does not r-correctly reproduce (g)'s premise. Read in terms of (c_6), the models $\mathfrak{M}_k^{(10)}$ such that $\Im_k(a) \in \Im_k(G)$ specify worlds in which Daryl is a woman. That does not suffice to determine whether (g)'s premise is true in those worlds, i.e. whether Daryl is a mother. Analogously, (11) is i-correct but not r-correct (and thus not rr-correct) for (g). The reason is that models $\mathfrak{M}_k^{(11)}$ such that $\Im_k(F) \nsubseteq \Im_k(G)$ do not specify possible worlds relative to (c_6). All of these models entail that there exist mothers that are no women, which is informally judged to be impossible. As (17) uses the same FOL sublanguage as (11), (17) is not r-correct for (g) either, because models $\mathfrak{M}_k^{(17)}$ such that $\Im_k(F) \nsubseteq \Im_k(G)$ do not specify possible worlds.

(11) and (17) are not r-correct for (g) because the correspondence scheme (c_6) interprets the predicate letters of the FOL sublanguage used in (11) and (17) in terms of predicates that are informally judged not to be independent, more specifically, the set of mothers is informally judged to be a proper subset of the set of women. Yet, the model-theoretic machinery treats all predicate letters of FOL as representing logically independent predicates. It thus generates models that cannot be read as representing possible worlds. Moreover, model theory systematically assigns every element of every domain to the name letters of a relevant FOL sublanguage and all truth-value combinations to the sentence letters. That is, in order for models $\mathfrak{M}_k^{\phi}$ of an FOL formalization ϕ to specify possible worlds relative to a correspondence scheme (c_i), (c_i) must meet the following requirement of *informal independence* (cf. Baumgartner and Lampert 2008, 107):

(IN) A correspondence scheme (c_i) accompanying a representational formalization ϕ satisfies (IN) iff all expressions assigned to the categorematic terms of $\mathcal{L}^{\phi}$ by (c_i) are mutually informally independent and (c_i) assigns neither informally tautologous nor contradictory expressions to any of the categorematic terms of $\mathcal{L}^{\phi}$.

As necessary condition of r-correctness, (IN) imposes severe constraints on the representational formalizability of a natural language statement Φ. If the truth conditions of Φ cannot be cashed out in terms of informally independent non-tautologous and non-contradictory expressions, Φ cannot be representationally formalized by means of FOL. FOL can only be used as a tool to model* truth conditions of statements that can be analyzed on the basis of components that conform to (IN). As a consequence, representational formalizations often need to be accompanied by semantic analyses.[11] In the case of argument (g), for instance, such an (IN)-satisfying semantic analysis is feasible. (18) and (c_6) analyze the predicate "...is a mother" contained in (g) in terms of "...is a woman and ...has a child". This analysis, in turn, yields that all models $\mathfrak{M}_k^{(18)}$ specify possible worlds w_k. Furthermore, in every $\mathfrak{M}_k^{(18)}$ the truth values of (18) and its components match the truth values of (g) and its components in the corresponding w_k. Hence, (18) is rr-correct for (g).

[11] In Baumgartner and Lampert (2008), we argue that, from the representational perspective, there is no clear distinction between logical formalization and semantic analysis.

The same holds for (19). The premise and conclusion of (19) are formally equivalent to the premise and conclusion of (18), and as (c_6) interprets the categorematic terms of (19) in compliance with (IN) all models $\mathfrak{M}_k^{(19)}$ represent worlds that are possible. Therefore, as (18) successfully models* the truth conditions of (g) and of its components, so does (19). All formalizations $\langle \gamma, \psi \rangle$ with components that are equivalent to the components of (18) and whose models specify possible worlds are rr-correct for (g). This finding can be generalized: for any two equivalent formalizations ϕ and ϕ' that are interpreted in such a way that (IN) is satisfied it holds that either both ϕ and ϕ' are r-correct for a statement Φ or none of them is. And analogously: for any two formalizations $\langle \gamma, \psi \rangle$ and $\langle \gamma', \psi' \rangle$, such that the members of γ are equivalent to the members of γ' and ψ is equivalent to ψ' and (IN) is satisfied, it holds that either both $\langle \gamma, \psi \rangle$ and $\langle \gamma', \psi' \rangle$ are rr-correct for an argument $\langle \Gamma, \Psi \rangle$ or none of them is.[12]

Clearly though, the tautologous subformula $p \vee \neg p$ of (19) adds nothing whatsoever to formally exhibiting the representational validity of (g). On the contrary, redundant syntactical surplus tends to obscure rather than render transparent the truth conditions of an argument's premises and conclusions. It is hence preferable to formalize (g) in terms of (18). A formalization $\langle \gamma, \psi \rangle$ that exhibits the representational validity of an argument $\langle \Gamma, \Psi \rangle$ must not only be rr-correct, but moreover it must not contain redundant elements. Relative to one (IN)-satisfying correspondence scheme (c_i), there exist infinitely many rr-correct formalizations of an argument $\langle \Gamma, \Psi \rangle$ such that for any pair of such formalizations $\langle \langle \gamma, \psi \rangle, \langle \gamma', \psi' \rangle \rangle$ it holds that each member of γ is equivalent to the corresponding member of γ' and ψ is equivalent to ψ' (and, hence, $\langle \gamma, \psi \rangle$ is equivalent to $\langle \gamma', \psi' \rangle$). Call the set ϱ of these formalizations an *rr-correct set* of $\langle \Gamma, \Psi \rangle$. Relative to different correspondence schemes and within different FOL sublanguages, an argument $\langle \Gamma, \Psi \rangle$ may have several rr-correct sets. The representational validity (or invalidity) of $\langle \Gamma, \Psi \rangle$ is best exhibited by the minimally complex formalizations in each rr-correct set.

There are several measures available in the literature that numerically reproduce the complexity of an FOL formula. For instance, Hodges (2001, 47) defines the *complexity* of a formula ϕ to be the number of its subformulas, where the subformulas of ϕ are the atomic expressions contained in ϕ, the molecular expressions in ϕ composed by means of logical connectives and quantifiers, and ϕ itself. According to this complexity measure, (18) has complexity 5, whereas (19) has complexity 9. Against this background, the *minimally complex* formalizations in a rr-correct set ϱ are simply definable as the formalizations with lowest complexity measure in ϱ. The minimally complex formalizations in different rr-correct sets exhibit representational validity equally well.

In light of these considerations, I advance the following principle for exhibiting representational validity:

(r-VP) An argument $\langle \Gamma, \Psi \rangle$ is exhibited to be representationally valid (II-valid) iff $\langle \Gamma, \Psi \rangle$ is formalized by an FOL-valid formalization $\langle \gamma, \psi \rangle$ that is both rr-correct for $\langle \Gamma, \Psi \rangle$ and minimally complex.

For completeness, a formalization ϕ can be said to be *representationally adequate*, i.e. *r-adequate*, for a statement Φ iff ϕ is both r-correct for Φ and minimally complex. Analogously, $\langle \gamma, \psi \rangle$ is *r-adequate* for an argument $\langle \Gamma, \Psi \rangle$ iff $\langle \gamma, \psi \rangle$ is both rr-correct for $\langle \Gamma, \Psi \rangle$ and minimally complex.

[12] In this regard, r-correctness differs significantly from i-correctness. While formalizations of a statement Φ that introduce arbitrary tautologous elements not contained in Φ cannot be seen to model* permissible reinterpretations of Φ, formalizations of Φ model* the truth conditions of Φ even if they feature an excessive tautologous surplus, provided that (IN) is respected.

5 Conclusion

By developing concrete formalization criteria that are suitable for exhibiting the interpretational and representational validity of arguments, this paper has shown that formalizing an argument from the interpretational and from the representational perspective are two very different tasks. Formalizing interpretationally presupposes clarity on the permissibility of reinterpretations and aims at modeling* the space of permissible reinterpretations relative to which a valid argument is truth-preserving. By contrast, formalizing representationally presupposes clarity on truth conditions and aims at modeling* the space of possible worlds relative to which a valid argument is truth-preserving. The first framework varies the interpretations of premises and conclusions and evaluates these reinterpretations relative to independently given specifications of possible worlds. The second framework varies the configurations of matters of fact to generate different possible worlds in which premises and conclusions are evaluated relative to independently given interpretations of pertinent languages. Formalizing an argument in an illuminating manner, hence, requires that the adopted framework be made explicit. Otherwise, confusion is inevitable.

Acknowledgements I am particularly indebted to Timm Lampert for countless discussions about the topics of this paper and for our common work on logical formalization. Moreover, I thank Wolfgang Spohn, Jochen Briesen, Alexandra Zinke, Brian Leahy and the anonymous referees of this journal for helpful comments on earlier drafts. Finally, the paper has profited from discussions with audiences at two research colloquia at the Humboldt University Berlin and the University of Konstanz.

References

Barwise, J. and J. Etchemendy (1999). *Language, Proof and Logic*. Stanford: CSLI Publications.

Baumgartner, M. and T. Lampert (2008). Adequate formalization. *Synthese 164*, 93–115.

Beall, J. C. and G. Restall (2006). *Logical Pluralism*. Oxford: Oxford University Press.

Blau, U. (1977). *Die dreiwertige Logik der Sprache*. Berlin: de Gruyter.

Brun, G. (2004). *Die richtige Formel. Philosophische Probleme der logischen Formalisierung*. Frankfurt a.M.: Ontos.

Brun, G. (2008). Formalization and the objects of logic. *Erkenntnis 69*, 1–30.

Brun, G. (2012). Adequate formalization and De Morgan's argument. *Grazer Philosophische Studien 85*, 325–335.

Bueno, O. and S. A. Shalkowski (2009). Modalism and logical pluralism. *Mind 118*(470), 295–321.

Chiswell, I. and W. Hodges (2007). *Mathematical Logic*. Oxford: Oxford University Press.

Chomsky, N. (1977). *Essays on Form and Interpretation*. Amsterdam: North-Holland.

Davidson, D. (1967). The logical form of action sentences. In N. Rescher (Ed.), *The Logic of Decision and Action*, pp. 81–95. Pittsburgh: University of Pittsburgh Press.

Davidson, D. (1984). *Inquiries Into Truth and Interpretation*. Oxford: Oxford University Press.

Epstein, R. L. (1990). *The Semantic Foundations of Logic: Propositional Logics*. Dordrecht: Kluwer.

Epstein, R. L. (1994). *The Semantic Foundations of Logic: Predicate Logic*. Oxford: Oxford University Press.

Etchemendy, J. (1990). *The Concept of Logical Consequence*. Cambridge: Harvard University Press.

Field, H. (2009). Pluralism in logic. *Review of Symbolic Logic 2*(2), 342–359.

Gómez-Torrente, M. (1996). Tarski on logical consequence. *Notre Dame Journal of Formal Logic 37*, 125–151.

Hale, B. (1996). Absolute necessities. *Philosophical Perspectives 10*, 93–117.

Hodes, H. T. (1984). Logicism and the ontological commitments of arithmetic. *Journal of Philosophy 81*(3), 123–149.

Hodges, W. (2001). Elementary predicate logic. In D. Gabbay and F. Guenthner (Eds.), *Handbook of Philosophical Logic* (2 ed.), Volume 1, pp. 1–130. Dordrecht: Kluwer Academic Publishers.

Jackson, B. (2007). Beyond logical form. *Philosophical Studies 132*, 347–380.

Lampert, T. and M. Baumgartner (2010). The problem of validity proofs. *Grazer Philosophische Studien 80*, 79–109.

Massey, G. J. (1975). Are there any good arguments that bad arguments are bad? *Philosophy in Context 4*, 61–77.

McFetridge, I. (1990). Logical necessity: Some issues. In J. Haldane and R. Scruton (Eds.), *Logical Necessity and Other Essays by Ian G. McFetridge*, pp. 135–154. London: Aristotelian Society.

Montague, R. (1974). *Formal Philosophy: Selected Papers of Richard Montague*. New Haven: Yale University Press.

Ray, G. (1996). Logical consequence: A defense of Tarski. *Journal of Philosophical Logic 25*(6), 617–677.

Read, S. (1994). Formal and material consequence. *Journal of Philosophical Logic 23*, 247–265.

Russell, G. (2008). One true logic? *Journal of Philosophical Logic 37*(6), 593–611.

Sainsbury, M. (2001). *Logical Forms. An Introduction to Philosophical Logic* (2 ed.). Malden, MA: Blackwell Publishing.

Shalkowski, S. A. (2004). Logic and absolute necessity. *Journal of Philosophy 101*, 55–82.

Shapiro, S. (1998). Logical consequence: Models and modality. In M. Schirn (Ed.), *The Philosophy of Mathematics Today*, pp. 131-156. Oxford: Oxford University Press.

Shapiro, S. (2005). Logical consequence, proof theory, and model theory. In S. Shapiro (Ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, pp. 651–671. Oxford: Oxford University Press.

Sher, G. (1996). Did Tarski commit 'Tarski's fallacy'? *Journal of Symbolic Logic 61*, 653–686.

Tarski, A. (1956). *Logic, Semantics, Metamathematics*. Oxford: Oxford, Clarendon Press.

Vecsey, Z. (2010). On the pluralistic conception of logic. *Pragmatics & Cognition 18*, 1–16.

Wittgenstein, L. (1995). *Tractatus logico-philosophicus*. Frankfurt a. M.: Suhrkamp.