Contents

Organon F 32 (1) 2025: 2–28 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32101

RESEARCH ARTICLE

A Temporal Relationship Theory for the Justification of Love

Muk Yan Wong*

Received: 15 June 2020 / Revised: 10 April 2022 / Accepted: 30 May 2022

Abstract: We believe that love happens for a variety of reasons, yet the alteration of these reasons need not alter love. Philosophers call the former the selectivity of love and the latter the constancy of love. In this paper, I critically review quality theory and person theory and argue that neither can explain both phenomena. In light of Kolodny (2003) and Bagley (2015), I propose a temporal relationship theory (TRT) and argue that love between two people is justified if they are in a relationship for which a similar interpretation of the history, understanding of the present, and anticipation of the future of some identify-shaping ends of the relationship is shared. TRT can explain both the constancy and selectivity of love and the role of the beloved's qualities and identity in the justification of love.

Keywords: Love; temporality; qualities; identities; justsification.

1. Introduction

In this paper, I review quality theory and person theory in terms of their justification of love and argue that neither can explain both the constancy

* The Hang Seng University of Hong Kong
https://orcid.org/0000-0003-0891-8660
Department of Social Science, The Hang Seng University of Hong Kong, Hang Shin Link, Siu Lek Yuen, Shatin, N.T., Hong Kong
mywong@hsu.edu.hk

and selectivity of love. In light of Kolodny (2003) and Bagley (2015), I propose a temporal relationship theory (TRT) and argue that the love between two people is justified by a relationship that shares a similar temporality about some identity-shaping ends of their relationship. TRT explains both the constancy and selectivity of love and the role of the beloved's qualities and identity in the justification of love.

2. Justification for Love

First, let me clarify the kind of justification to which I am referring. This is not a physical or biological explanation of what exactly causes one loving another but rather a normative justification explaining why one's love towards another is appropriate. Such appropriateness is not dictated by reasons like a moral or inferential requirement, of which one is wrong not to behave in certain way or hold a certain belief given the corresponding moral or logical reason. Instead, such appropriateness is more like an aesthetic reason which provides grounds for holding certain beliefs or experiencing certain feelings free of the potential charge of being unreasonable if one fails to believe or feel so. Abramson and Leite (2010) refer to such justification as 'warranting but not requiring reasons'. Kolodny (2003) refers to it as 'non-insistent reason'. Such a weak sense of reason is commonly used in the justification of emotion. For instance, cognitive/appraisal theories have suggested that an emotional response is justified when its object constitutes the emotion's core relational theme (See Lazarus & Smith 1993). Anger is about "a demeaning offence against me and mine." So, your anger responses towards X are appropriate when X insults you or unfairly violates your rights. Despite what X did, you are not being irrational not to get angry because you may hold certain non-emotional beliefs suggesting that you should not do so. Understanding justification in this way, this paper aims to explain the proper grounds of love without implying that these grounds are necessary or sufficient conditions of love.

Among the common forms of love towards human and non-human beings, and towards living things and non-living things, it is doubtful that all of them share the same kind of justification. This paper targets the kind of love that typically develops between family members, lovers and friends.

Thus, the justification I give does not deny that other kinds of love are not love. I choose this specific kind of love to be my *explanandum* because of three of its profound yet puzzling features. First, such love is non-egocentric, i.e., a concern about the well-being of another for their own sake. Its justification is particularly interesting because most emotions are egocentric and can be justified for their function of enhancing our well-being, e.g., a fear of snakes is justified because it helps us avoid a threat to our life. However, we actively care about the growth and prosperity of our loved ones not because their well-being affects our well-being, but because their well-being matters to us just like ours do (even if it does not matter to them.) Badhwar (1987) refers to this as 'end friendship', which is differentiated from loving someone instrumentally. For instance, even though parting ways with her child can make a mother sad, she still wants them to become independent because that is what they want, she believes, and what is good for them. We care about our beloved so much that we basically share their emotions. Kolodny (2003) describes this particular characteristic of love as emotional vulnerability:

> To say that A is emotionally vulnerable to B (or r) is to say, in part, that A is disposed to have a range of favorable emotions in response to A's beliefs that B (or r) has fared or will fare well, and a range of unfavorable emotion is in response to As beliefs that B (or r) has fared or will fare poorly (Kolodny 2003, 152).[1]

Thus, we feel happy when our lovers fare well, sad about their losses, angry when they are wronged and fearful of the threats they face because although our well-being is not affected theirs is. Helm (2017) calls such a non-egocentric feature the robust concern of love (see also LaFollette 1996; Frankfurt 1999).

Non-egocentric love, however, is not unconditional like *agape* or the love of a neighbour, according to Luther's interpretation of the Bible. Instead, non-egocentric love is highly selective, i.e., there are always reasons why we start, continue and stop loving a specific person (but not another) non-egocentrically. Helm refers to this feature as the discernment of love and asks, 'What, if anything, justifies my coming to love this particular person

[1] In p.151, Kolodny defines "r" as the reasons for A's emotional vulnerability to B.

rather than someone else given limited time, energy, and other resources?' (Helm 2010, 20) Not only do we need a reason to love, we also want our lover to love us for the right reason. In Delaney's words, we have 'a desire to be loved for such properties where these properties are appreciated in a way not too different from the way that you appreciate them' (Delancy 1996, 345). For instance, if X, even though she looks gorgeous, identifies herself as a poet, she would want her partner to love her for her writing talent rather than for her beautiful face. I call this feature the selectivity of love.[2]

The third feature of non-egocentric love is its constancy. As Shakespeare famously said, 'Love is not love which alters when it alteration finds'. The constancy of love can be interpreted as its power to survive change or sustain the loss of the lover's qualities that initially justified the love, and to prevent a replacement of a lover even when someone better appears. Such constancy is sometimes referred as the non-fungibility of love (see Helm 2017). Unlike a ten-dollar note that can be replaced by two five-dollar notes without losing any value, our lover cannot be replaced by a qualitatively indistinguishable duplicate suggested by Parfit's thought experiment (1984). Even if we fall in love with the duplicate, our love towards her is fundamentally different from our original love. A puzzle about such constancy is that it is seemingly incompatible with the selectivity of love. Taking our intelligence as the reason of love, Delancy (1996) suggests that we hope that love remains constant enough to prevail over any change in our intelligence, yet not so blind that it regards a diminished level of intelligence as something negligible. Badhwar characterises such tension between constancy and selectivity of love in a different way:

[2] One of the reviewers questioned whether parental love is a kind of unconditional love which is not selective by nature. I agree that parental love is very close to unconditional love for there is rarely any change of quality of one's son or daughter which will terminate our love to them. Yet, parental love is still selective if we understand the qualities of one's child in a relational or historical sense. For instance, we will have a stronger reason to love our child if we have established a long-term intimate relationship with them than if we are only biologically related to them but do not spend much time with each other. Further elaboration of such qualities can be found in section 5.

> On the one hand, my love for you, who are my friend, is not love if it alters *whenever* it alteration finds…But neither is my love for you, the unique person, love for *you* if it remains unaltered through *all* alterations of your qualities (as if 'you' = 'bare particular') – as it is in unconditional love (Badhwar 1987, p. 6).

With these three features of love in mind, my question is this: What are the proper grounds for starting and continuing a non-egocentric love that selectively directs towards a specific person, yet remains constant when this person changes, like the kind of love that typically instantiates among lovers, friends and family members?[3]

3. Quality Theory

One answer to the question above is provided by quality theory[4], which claims that X's love towards Y is justified if Y possesses qualities that X regards as valuable. Plato, who was arguably the earliest proponent of such an approach, suggested that it is the eternal beauty of our beloved that justifies our love towards them. It is more justified to love someone with a beautiful soul than a beautiful body because only the former can be pregnant with immortal beautiful offspring, e.g. great ideas, that offer perpetual glory to oneself. Such an instrumental account of love, despite its strength in explaining the selectivity of love, is incompatible with the non-egocentric love we aim to explain (See Badhwar). Aristotle suggested three qualities as reasons for friendship, namely, profit, pleasure and character, and

[3] Frankfurt (1999) may object that such a question is ill-formed as it assumes that there is a reason for love. He proposes a no-reason view, which suggests that 'loving is circumscribed by a necessity of that kind: what we love and what we fail to love is not up to us'. (p. 46). As we cannot choose who to love, it is redundant to ask whether our choice is appropriate or not. I am not sure whether love is involuntary as suggested by Frankfurt. It seems counterintuitive to suggest that we never choose who to love, or can never judge whether someone's love is reasonable or not. Not denying the validity of Frankfurt's argument, I propose that a more rational way is to review the strengths and weaknesses of different proposed reasons of love before we draw the conclusion that there is no reason for love at all.

[4] Named by Kolodny (2003).

acknowledged only the last one as a reason for real friendship. For Aristotle, the love between friends must be non-instrumental. He says, 'Let "loving" [*to philein*] be wishing for someone the things that he deems good, for the sake of that person and not oneself, and the accomplishment of these things to the best of one's ability'. (*Nicomachean Ethics*, 1380b36-81a1) To appreciate the character of one's friend for the sake of its contribution to her but not one's happiness therefore serves as a *prima facie* valid reason for non-egocentric love.[5]

Regarding beauty or character as justification, however, faces the problem of replaceability. As these qualities can be instantiated in someone else, they serve as proper grounds for us to replace our lovers with someone with the same qualities, or to trade our lovers with someone with better qualities, thus violating the constancy of love. Modern proponents of quality theories have tried to explain such constancy of love by elaborating the effect of valuing the qualities of our beloved as a reason for love.[6] Jollimore suggests that 'Loving someone is, in large part, a kind of positive, appreciative response to her in virtue of her attractive, desirable, or otherwise valuable properties' (Jollimore 2011, 25), yet these qualities are immune from comparison with others' because of love's 'blinding' effect:

---

[5] Irwin (1988) provided an alternative interpretation of Aristotle's theory of friendship. He thought that Aristotle's account is instrumental in the sense that A is willing to care about B's character for her own sake only because A regards B as a propagation of himself, and therefore the good character of B is simultaneously contributing to the happiness of B and A. Whiting (1991) criticised such an interpretation for its assumption of 'the colonizing ego' (See Whiting 2013 for a more elaborate criticism).

[6] One influential quality theory that gives strong explanation of the constancy of love is given by Velleman (1999). Velleman proposes that the quality that justifies love is the dignity of the beloved. As the nature of human beings (in a Kantian sense), it will never change. A similar account is given by Setiya (2014), who argues that the quality that justifies love is the humanity of the beloved. Such accounts, however, cannot explain the selectivity of love because the quality of dignity is shared by all human beings. Velleman responds that one loves a particular person because of the 'contingent fit' between them and us. As Helm (2017) points out, such a 'contingent fit' only explain why one falls in love with a particular person, but not why one's love is justified.

> The beloved's attractive properties 'fill the mind' of the lover, leaving no room for her to appreciate similar properties possessed by others. An infatuated person becomes blind to the attractions of everyone but his beloved (Jollimore 2011, p. 42).

Not only is a lover blind to the good qualities in people other than his beloved, he is also blind to the adverse qualities of the beloved that he usually finds intolerable in others, and sees his beloved 'in the best possible light (Jollimore 2011, 48).' By such double-blindness, Jollimore's account claims to explicate the constancy of love by eliminating the chance of replacing beloveds through a comparison of their qualities. Nevertheless, other than the limited experimental results cited by Jollimore, our being psychologically wired to overlook qualities other than our lover's does not prove that it is reasonable to overlook similar qualities in others or unreasonable to pay attention to them. Besides, if our attention is so predominantly focused on the good qualities of our beloved that the bad ones become invisible, it is hard to imagine that we are to remain ignorant or indifferent if these qualities, as the foundation of our love, are gone. If so, would not such heightened sensitivity to the good qualities of a beloved weaken rather than strengthen the constancy of love?

Rorty (1986) suggests that constancy of love is ensured by the power of love to change the lover such that they learn to value the new qualities of the beloved when the old qualities disappear. Characterising love as 'dynamic permeable', she states the following:

> It is permeable in that the lover is changed by loving and changed by truthful perception of the friend. Permeability rejects being obtuse to change as an easy way of assuring constancy. It is dynamic in that every change generates new changes, both in the lover and in interactions with the friend. Having been transformed by loving, the lover perceives the friend in a new way and loves in a new way (Rorty 1986, 402).

In her example, Louis loved Ella for her 'crisp way of playing Scarlatti'. Through their continuous interaction, Louis was changed by his love towards Ella in that he learned to appreciate some other traits of hers that previously had not seemed admirable, such as 'the sequence of her moods, the particular way she had of sitting still, head bent when she listened to

music'. Louis thus realised that he would not transfer his love even if someone played Scarlatti more brilliantly than Ella. In this way, the constancy of Louis' love is not guaranteed by Ella's unchanging qualities but rather by the correspondence between Ella's changing qualities and his appreciation of them. Rorty argued that such dynamic constancy is superior to the person theory that suggests that one's love never changes, no matter how the character of our beloved changes, for 'he'd be lunatic to love her at 60 in just exactly the same way as he had at 20; and he'd be cruel to love her way of playing Scarlatti if her hands had been mangled in an accident' (Rorty 1986, 403).

I find Rorty's account problematic in two ways. First, it is uncertain whether the dynamic interaction between lovers contributes more to the constancy or inconstancy of love. It is a beautiful coincidence that a lover happens to develop a new appreciation of the qualities that exist in his beloved. However, as such changes are not directed by either party, there is no reason why consonant changes are more likely to happen than inconsonant ones. Worse still, compared with typical quality theories, a termination of love can be justified in Rorty's account even when the beloved's qualities do not change if the lover develops new interests and no longer appreciates those old qualities. Secondly, assuming that consonant changes sometimes happen coincidentally, the problem of replaceability persists. In Rorty's example, even if Louis does not replace Ella with someone who plays Scarlatti better, because he starts appreciating the sequence moods of Ella, it is not clear why Louis does not replace Ella with someone who better expresses this new quality. The problem is not eased even if Ella is likely to produce changes that are consonant with Louis's changes because such capacity is a second-order quality that is by itself replicable in others. Despite these problems, Rorty offers valuable insight with the idea that the justification of love should be considered in a *temporal* manner because both the reasons for love and the appreciation of those reasons are continuously changing through the interaction of lovers.

Some proponents of quality theory have tried to explain the constancy of love by elaborating on how we value the qualities of the beloved. Badhwar (1987) suggests that such valuing is mediated through the idiosyncratic expression of the beloved's qualities. The qualities of our beloved

are not comparable with those of others because qualities are expressed in a style that is unique to everyone. Style, however, although arguably more unique than qualities, is not immune from comparison with others either. In fact, Badhwar admits that the distinction between style and qualities is just a matter of degree: 'the style in which one expresses certain qualities can itself be described as a set of qualities, and the qualities expressed can be described as a style of facing life' (Badhwar 1987, 20). If so, individualising qualities with style seems to be a futile move in protecting the constancy of love.

Abramson and Leite provide a more plausible explanation concerning the expression of the beloved's qualities: 'the laudable qualities of character are displayed in ways directed towards the lover that makes those qualities the proper grounds for the lover's response' (Abramson — Leite 2011, 681). For instance, John being a considerate person is not enough to justify Mary's love towards him; such character must be expressed in terms of the behaviours or attitude directed towards Mary for it to become proper grounds for her love. Such relational context, although not a reason for love by itself, is what makes the beloved's quality incomparable to other's qualities. In response to Kolodny's (2012) thought experiment of an imagined intrinsic duplicate of Jane, Abramson and Leite argue that one has no grounds for replacing one's beloved with such a person because the relevant relational context is missing. I find their interpretation of an appropriate relational context as the direct contact between lover and beloved questionable in two ways. First, it is not clear why John's good character, being expressed directly towards Mary, is a necessary condition of such character being a reason of love. Let us say that Mary is a wealthy woman who needs no help from her friend John. Mary finds out about John's generous acts of regularly helping the poor. Even if she knows that such goodwill is not to be expressed towards her, she may still appreciate such character, which gives her more reason to love him (as a friend). Secondly, it is not clear why John's quality, as expressed towards Mary, cannot be compared with, say, Jones' similar quality as expressed towards Mary, given that both qualities are expressed towards the same person. The claim that such comparison is impossible because the relational context that helps justify Mary's love towards John must involve John, sounds no different from saying that Mary

has no reason to love Jones because Jones is not John no matter what qualities Jones possesses. Despite these problems, Abramson and Leite's insight that the relationship between lovers plays an essential role in the justification of love should not be overlooked.

4. Person Theory

While quality theory in general struggles to explain the constancy of love, person theory provides a much simpler answer. If the reason of love is the personhood of one's beloveds, a change in their qualities, or a stronger instantiation of such qualities in others, does not justify a shift of love because the personhood of the beloved is unaffected by these changes and comparisons. Arguably, the earliest version of person theory was provided by Aristophanes in Plato's *Symposium*, which told a myth about how human beings had been split into halves by Zeus and then spent the rest of their lives longing to find their lost half. We love and continue to love this one half of us, i.e., our soulmate, not because they carry any qualities we appreciate but because they are who they are, the lost half of ourselves as determined by fate. Putting aside this preposterous transcendental assumption, modern philosophers agree that what justifies our love is the beloved's personhood rather than their qualities. Nozick says, 'An adult may come to love another because of the other's characteristics; but it is the other person, and not the characteristics, that is loved' (Nozick 1947, 167). One notorious problem of person theory is its weakness in explaining the selectivity of love. To say that one loves A but not B because A is A but B is not A is clearly no explanation at all. Logically, everyone is themselves. This does not give one a reason to love everyone. Besides, as our beloved always remains who they are, the justification for our love is never invalidated, which suggests that we should never terminate our love. Even if we accept a more liberal view of identity, e.g., Parfit's idea of psychological continuity, the situation is not much better. Only if there is a significant change in our lover's mental states, including their beliefs, desires and emotions, can they be regarded as a different person, which then justifies terminating our love. Such occasions, e.g., suffering from permanent vegetative state (PVS) or Alzheimer's

Disease, are quite rare. Our beloved losing some of our favourite characteristics, e.g., loyalty, faithfulness or kindness, need not become a reason for terminating our love if those changes are not drastically enough to change who they are (rather than what we want them to be).

Inspired by Kripke's theory of rigid designation, Kraut proposes a historical person theory of love by drawing an analogy as follows: 'a proper name is committed to its bearer, in much the way that a lover is historically committed to the object of his love' (Kraut 1986, 424). Just as a proper name rigidly designates an object by a casual history with it but not a definite description of its properties, one's love towards the beloved is justified by the causal history with that particular person but not by their characteristics. This explains why it is the person, rather than their qualities, that justifies love, and why changes of the beloved's qualities do not justify a change of love. Also, the problem of replaceability does not emerge because the historical relationship between two lovers is unique and non-replicable. Kraut's historical account is superior to the metaphysical account in explaining the selectivity of love. Even if both John and Jones possess the same kindness that Mary appreciates, Mary is justified to love John but not Jones only if John's kindness (but not Jones') is expressed in a historical connection between him and Mary that defines who he is. In other words, Mary loves John not because John is John, but because John is the person to whom she is historically related. Kraut emphasizes that such historical relation does not prohibit a justified termination of love: 'Historicity does not entail permanence; analogously, proper names are not eternally bound up with their referents. Every proper name can lose its use. Every love has it limits' (Kraut 1986, 425). However, it is not clear under what conditions the personhood defined in terms of historicity stops being a justification of love given that history is physically unchangeable. Grau (2010) suggests that Kraut may stick to his analogy by following Kripke's way to explain how a proper name loses its use when it acquires a new use through a new causal history, e.g., 'Santa Claus' being switched from a historical saint to a fictional entity. Thus, Mary is justified to stop loving John if she establishes a historical relation with Jones that justifies her love towards him. Here, an obvious question is this: Given that both John and Jones are historically connected to Mary, why might the history between

Jones and Mary override the history between John and Mary? How is a comparison between historical relationships possible?

A deeper question behind the larger historical approach is this: Why does history matter? As Helm (2017) states: 'The mere fact that I have loved her in the past does not seems to justify my continuing to love her in the future'. Badwhar (1987) suggests that a historical relationship with the beloved has two values. First, it reveals their lovable characteristics and provides us with a better knowledge about each other. Secondly, it creates new appreciable characteristics of us that strengthen love when history is shared emotionally and cognitively. Admitting these epistemological and creative values of shared history as justification of love, however, violates the basic insight of the person theory, i.e., we love our beloved as the person who has a history with us but not for the value that such history provides because these replaceable values cannot be the factors defining who our lover is. Grau raises a similar concern by stating that 'this approach grants the past a type of instrumental value, one which derives solely from the importance of its future predictive benefits' (Grau 2011, 17). In response to the question of why history matters in love, Grau suggests that the importance of history is so fundamental that it 'can be defended as in an important sense ungrounded, natural, and best construed as neither rational nor irrational, but a fundamental non-rational (or arational) feature of our lives that need not be revised away out of fear of irrationality' (Grau 2011, 19). I believe that there is a reason why history matters, and the solution lies in how we understand the notion of history, which is the bedrock of the account I propose in section 5.

Nozick's person theory offers another solution to the problem of selectivity. He interprets love as follows:

> Love, romantic love, is wanting to form a *we* with that particular person, feeling, or perhaps wanting, that particular person to be the right one for you to form a *we* with, and also wanting the other to feel the same way about you. (Nozick 1989, 78)

Given that the purpose of love is to form such we, one's love is justified if the beloved is the right person to form a *we* with. Constituted by the well-being and autonomy of each party, whether someone is the right person to form such *we* is determined by how their corresponding senses of well-being and autonomy fit together. Thus, what determines the selectivity of love is

not the identity of the beloved per se, but the part of their identity that contributes to the formation of a joint *we*. Similarly, one is justified to stop loving the beloved if the latter becomes unsuitable to be part of the *we*. As Nozick says: 'though no longer dependent upon the particular characteristics that set it (love) off, it can be overcome over time by new and sufficiently negative other characteristics' (Nozick 1989, 76). One problem of Nozick's account is that the requirement of being the right person in a *we* seems highly demanding. As Nozick explains: 'People who form a we pool not only their well-being but also their autonomy. They limit or curtail their own decision-making power and rights; some decisions can no longer be made alone' (Nozick 1989, 71). So, to be the right person not only requires that one to surrender (part of) autonomy; the autonomy surrendered must also be combined nicely with the other's autonomy to serve the purpose of promoting the joint well-being of the we. If selectivity of love is justified only by such miraculous formation of joint autonomy, then it may lead to the unwelcome consequence that justified love is rare and most everyday love choices are unjustified.

Friedman argues for a less demanding bond of the lover, namely, an interpersonal federation, which 'does not erase the existence of the two lovers as separable and separate agents with continuing possibilities for the exercise of their own respective agency' (Friedman 1998, 165). Lovers may merge partially in their subjectivity, agency or objecthood and yet remain independent in other aspects. These merged aspects may change overtime without affecting the legitimacy of the federation. The only requirement is that the merged aspects must pertain to the identity of the lovers. By cutting loose the commitment of such federation, Friedman claims that it can actually promote our autonomy 'by promoting our self-understanding, self-esteem, and capacities to act effectively in concert with others' (Friedman 1998, 170). Contrary to Nozick's highly demanding account of we, Friedman's account of federation seems so liberal that it is not clear how the selectivity of love can be explained. Given that no merging identity-shaping aspect is more important than others, and that no specific amount of autonomy from each party contributing to the federation is required, we have no idea why someone is a better candidate to form a federation than another, or whether we are justified to shift love to someone who can merge

with us in aspects other than those of our beloved, or whether we are justified to terminate love when our beloved stops merging with us in specific aspects. In short, while I find the idea of justifying the constancy and selectivity of love by a specific union of lovers that defines who they are to be insightful, the solution seems to lie between Nozick's narrow account and Friedman's liberal account.

5. Relationship Theory

Let us see what we can learn from the last two sections. Generally, the qualities of a loved one are specific and alterable, which make them good candidates to explain the selectivity but not constancy of love. Depending on the perception and expression of qualities, however, the constancy of love can be explained to a certain extent. Of particular interest is Rorty's idea that love may continuously change the lover's perception of the beloved's qualities, along with Abramson and Leite's idea that qualities expressed in a relational context are more stable than non-relational qualities. However, the personhood of a lover is relatively general and unalterable, which make it a good candidate to explain the constancy but not the selectivity of love. Depending on what constitutes the identity of a person, the selectivity of love can be somehow explained. Of particular interest is Kraut's idea that what makes a person a justified object of love is her history with oneself, and Nozick and Friedman's idea that some kinds of union constituted by the personhood of lovers justify love. Drawing from such a review, it seems that we need something more stable than qualities and less rigid than personal identity to explain both the selectivity and constancy of love. Such a thing is jointly constituted by both lovers as a person (including their shared history), which may affect how they interpret other qualities and how these qualities are expressed. Some philosophers have proposed that a relationship is that thing.

Kolodny's (2003) relationship theory argues that what justifies our love towards another is our relationship: 'My reason for loving Jane, I suggest, is my relationship to her: that she is my daughter, or my mother, or my sister, or my friend, or the woman with whom I have made my life' (Kolodny 2003, 146). According to Kolodny, a relationship that justifies love

must consist of an ongoing and historical pattern of mutual concern arising between particular people. The beauty of such a theory is that it simultaneously captures the strength of the quality theory and the person theory.

Concerning the constancy of love, as what justifies love is not the qualities of one's beloved but the less alterable qualities of the relationship, namely an ongoing and historical pattern of mutual concern, changes in the beloved's qualities or someone who appears to possess better qualities; these need not justify termination or a shift of love. Let us say that Mary is in a romantic relationship with John in which John cares about Mary because he is caring person. John stops being a caring person because of some trauma. This does not mean that Mary is therefore justified in ceasing to love John. John may acquire other reasons to care for Mary, e.g., his belief that Mary needs him. As long as the quality of mutual concern in their relationship does not change, the justification for their love remains unchanged even if John's qualities change. Even if someone else cares about Mary more than John does, as long as Mary's care for John is not based on John being the person who cares about her most, Mary's care for John need not decrease. Thus, their mutual concern, as the corresponding justification for love, need not change accordingly. Concerning the selectivity of love, a termination of love can be justified by a change of the less rigid beloved's identity *in a relationship* but not her personal identity per se. Kolodny discusses such a possibility with a story of a reputable historian who comes to devote himself to denying the Holocaust. He still expresses a deep concern for his wife, as does she to him. His mental states do not undergo the drastic change that may lead people to see him as a different person. However, Kolodny argues that his wife is justified in not loving him anymore because he is no longer the same man *in their relationship*. Being forced to withdraw her respect for him, she can no longer see him as someone with equal standing, without which his identity in the relationship cannot be maintained.

Appealing as relationship theory in general may be, I find Kolodny's definition of relationship in terms of mutual concern problematic in two ways. First, it may lead to the counterintuitive conclusion that love is always unjustified at the beginning. Given that love is a non-egocentric concern of the beloved, if X is justified to love Y only if they are in a relationship in which Y loves X, it is puzzling how their love can start without

either party loving another unjustifiably at the beginning. Under Kolodny's account, it seems that love is justified only after an unjustified love exists first, which leads to mutual concern in a relationship that in turn justifies love. Kolodny is well aware of this problem. He states that critics might argue that 'relationship cannot be reasons for falling in love because they do not exist until one has fallen in love' (Kolodny 2003, 169). In response to this criticism, he suggests that before the establishment of a relationship, a shared history of friendly interaction 'gradually give(s) rise to noninstrumental concern' for one's potential lover, on the assumption that she is 'disposed to reciprocate one's emerging concern'. Here, Kolodny lowers the requirement for the justification of love from a relationship of mutual concern to a friendly interaction of mutual disposition to reciprocate other's concern. The problem is, firstly, that friendly interaction seems not to be enough to justify love because it can occur among people who are clearly not in love with each other, e.g., between colleagues, or between salespeople and customers. Secondly, if what is needed to justify our love is not the actual concern of our beloved *to us*, but their general character, e.g., 'she isn't antecedently hostile…or divisively competitive, or self-absorbed, or sociopathic' (Kolodny 2003, 169), that dispose them to reciprocate *anyone's love*, the uniqueness of one's lover is lost and the problem of replaceability emerges again given that these non-relational characteristics or qualities are non-idiosyncratic and can be instantiated in many people.

Secondly, Kolodny's definition of relationship as mutual concern is both too wide and too narrow. It is too wide in the sense that an ongoing and historical pattern of mutual concern can exist among people who are not justified to love each other. Bagley suggests that the pedagogical relationship between teachers and students may include a concern of other's interest 'in pedagogically appropriate ways and for pedagogically appropriate forms of emotional vulnerability' (Bagley 2015, 11). The teacher may have concern for all of her students but is not justified to fall in love with any of them because what she really values is the pedagogical relationship as such, not any particular pedagogical relationship with any particular student. These mutual concern relationships that cannot justify love may also exist among athletes on the same team, colleagues working for the same goal, or even soldiers in the same army. However, Kolodny's definition of relationship

is too narrow in the sense that a justified love can occur among people *not* in a relationship with mutual concern. Price questions whether Kolodny's account 'leaves sufficient scope for unrequited love' (Price 2012, 224). In his example, even if Edith does not believe that she shares any relationship with Laurie as she does not return her feelings to his love, we may still believe that Edith has better grounds to love Laurie if she is kind and charming rather than mean and tiresome. Denying any justifying role of the non-relational qualities of the beloved, Kolodny's definition of relationship renders unrequited love always unjustified. Another counterexample is the love among family members. Parents are not always the concern of their children as much as children are the concern of the parents, but it does not render parents' love towards their children unjustified or serve as a reason not to love them. Kolodny responds by defining family relationship as an attitude-independent relationship that may justify love without mutual concern (See Kolodny 2003, 149). However, it is not clear what still constitutes the power to justify love in a family relationship when we take away mutual concerns from the latter. Is it because family members are tied by blood, by social institutions, or by moral norms? None of these seems strong enough if we have never related personally to our family members. It seems preposterous to argue that one must love one's father if one has never met him once in one's life.

Kolodny's relationship theory is thick in the sense that a relationship must meet a specific condition to qualify as a justification of love, namely, mutual concern. Bagley (2015) provides a thin relationship theory that may avoid Kolodny's problems. Drawing from the analogy of musical improvisation, Bagley argues that what justifies love is an improvisational partnership, which is defined as 'a type of ongoing relationship grounded in the partners' mutual recognition of one another as sharing an end with respect to a given activity' (Bagley 2015, 25), for which 'the ends lovers share constitute fundamental values with which they identify' (Bagley 2015, 26). That is, a relationship that justifies love must involve some shared ends between lovers whose values they identify with. However, there is no specific requirement for what these ends must be. Each relationship may pursue its own ends, just as each band may strive for a different style of music. Besides, these ends need not have been previously agreed upon by both parties

but are always in the process of being determined through their interactions. In such relationship, one should recognise another as an authority in judgement with respect to the value of the shared ends and the ways to achieve such ends, in the sense that one assumes that their actions are always working towards those shared ends, even though one has never witnessed such actions, just as one may acknowledge the decision of one's bandmates to play a riff that expresses a shared musical idea and play along. Bagley's account not only allows different relationships to have different ends, but lovers in a relationship do not even need to be able to specifically spell out their shared ends. Such a thin relationship theory provides us with the necessary flexibility to explain a great variety of love, including love without mutual concern.

Moreover, Bagley's account offers a deep explanation of the constancy of love. Mary cannot trade John for a better partner not because her beloved John possesses unique qualities, but because the standard to judge whether John deserves Mary's love, i.e., the value of their relationship, can only be provided by the interaction between Mary and John in the process of improvisation. As Bagley says, 'If you had a different partner, you'd have different standards: there's no common basis of comparison' (Bagley 2015, 28). Nor can Mary terminate her love when John appears not to be sharing the same goal or value with her anymore, for she has to assume John is still working towards the same goal or achieving the same value in a legitimate yet alternative way in an improvised relationship. The specific ends and functioning in such a relationship may rule out those counterexamples faced by Kolodny's theory, e.g., relationships between teacher and students, and among teammates or comrades. These relationships do not justify love because even though the parties involved may share the same end within a specific period, they may not be ready to acknowledge the other's authority or judgement when working towards the same undetermined end.

Despite these advantages over Kolodny's account, I find Bagley's historical explanation of the selectivity of love to be inadequate. Just as there are hardly any objective qualities for determining who is a good partner for musical improvisation, there seems to be no objective standard for determining who may be a good partner in a deep improvised relationship. Bagley suggests that we may appeal to the lover's joint history of particular

interactions as a 'common evaluative currency' (Bagley 2015, 29) to determine the qualification of a deep improvisor. This is a reasonable answer. Just as a musician may have a rough idea of how far their band can go after a few performances, two lovers may see how far their relationship can go after several interactions involving the clash and harmony of their fundamental values. Nonetheless, it is not clear in what way such joint history informs and determines the justification level of love in a relationship. It is not uncommon that people with a history of conflict may turn out to be life-long partners, or that people with a harmonious history turn out to be each other's most hateful enemies. Besides, such joint history may not explain why we choose someone as our lover at the beginning when no history is yet available. Neither does it explain why people may be justified to love someone who completely forgets the history of the relationship, e.g., a patient of Alzheimer's disease, nor someone who is not even capable of constructing or understanding any history, e.g. a new-born baby. The historical dimension of a love relationship seems more complicated than the function of 'learning-from-history' as suggested by Bagley's account.

6. A Temporal Relationship Theory

In light of the reviews above, I propose a temporal relationship theory to explain the justification of love and argue that it offers a better explanation of the constancy and selectivity of love. Such a theory can be regarded as a complement to Bagley's account with major modifications of his conception of the 'history' of a relationship.

The importance of the joint history of a relationship for the justification of love has been widely acknowledged by quality theories, person theories and relationship theories. In response to the problems faced by Bagley's account (and by the traditional historical account discussed in section 3 and 4), I propose that history of a relationship should not be understood as independent of its present and future; instead, it is an interpretation that is affected by the anticipation of the future, just as the latter is an interpretation determined by the recalling of a joint history. Defining the relation between the past and future of a relationship, we can say that the recalling of history and the anticipation of the future both determine the

present identity of a relationship. For instance, if a couple interprets their past as a history of material satisfaction, they are less likely to expect a poor yet mind-fulfilling life in their future. Thus, they may regard themselves as profit-maximising partners in the present relationship. If a couple expects to share a rich moral life with each other in the long run, they may interpret their sexually fulfilling history as a waste of time, and may regard themselves as sinners and moral practitioners in their present relationship. I call such an interactive relation between the interpretation of the past, the understanding of the present, and the anticipation of the future of a relationship the temporality of a relationship.[7]

In light of the revised idea of history above, I propose that love between two people is justified if they are in a relationship in which both parties are sharing a similar temporality of some identity-shaping goals or ends of the relationship. By sharing, I take the broad sense proposed by Bagley, which suggests that lovers who recognise each other as an authority in the judgement of an indeterminate end, either factually or counterfactually, can be regarded as sharing an end even if they cannot specify what the end is. By identity-shaping ends, I refer to ends of a relationship that at least partially determine the identities of the lovers. Only then can we explain the intuition of person theory that what justifies our love is the beloved's personhood rather than their qualities. I call this temporal relationship theory (TRT). TRT suggests that the level of justification of love is determined by the similarity between lovers' interpretation of the history, understanding of the present, and anticipation of the future of some specific identity-shaping ends of their relationship. Not only can TRT explain both the constancy and selectivity of love, it can also explain the roles played by the beloved's qualities and identity in the justification of love.

[7] The idea of the interlock between past, present, and future is enlightened by Heidegger's description of the authentic temporality of *Dasein*. He says, "The future is not later than having been, and having-been is not earlier than the Present. Temporality temporalizes itself as a future which makes present in a process of having been." (*Being and Time* 68: 401) My use of the term 'temporality', however, does not presuppose any metaphysical or phenomenological assumption of Heidegger's account. All it means is that the anticipation of the future, understanding of the present, and the interpretation of the past may affect each other.

Concerning the selectivity of love, one has a better reason to love someone who shares a more similar temporality of the relationship with oneself than to love those who do not. Before a relationship starts, there is no history of the relationship to which to appeal.[8] Thus, lovers can only focus on the similarity of their anticipation of the future of the relationship. If Mary foresees herself as a future parent, she has a stronger reason to love John, who also foresees himself as a parent, than Mark, who foresees himself without children in his life, because it is more likely for Mary and John to both identify their relationship as a family-establishing relationship than for Mary and Mark. The qualities of John that attract Mary's attention, e.g., piety and kindness towards young children, plays the role of reason of love by serving as an indicator reflecting how much John shares the ends Mary regards as essential to their relationship. Thus, some qualities, e.g., characters and virtues, are usually regarded as better reasons for love than others, e.g., appearance and social status, not because they are deeper or more sophisticated but because they are better indicators of whether someone is and will continue to share an important end of a relationship.

Another example to which no history can be appealed is the love of parents towards their infant. Even though a history between them exists, the infant is cognitively incapable of recalling any of it. Nor is the infant capable of anticipating any shared ends in the future. I argue that parents are justified to love them by appealing to a counterfactual similarity of temporality of their relationship. That is, parents are not unreasonable to believe that if their baby is cognitively mature enough, they will agree to the ends their parents plan to share with them. Such counterfactual shared anticipation of the future may then provide grounds for the parents to believe that they were currently sharing the same end and the time they spend together were a history of the shared end if the baby was cognitively mature, which in turn constitutes grounds as a justification of their love towards their baby. Thus, if a couple anticipate a similar future with their adopted baby or foster child, their love's towards them is no less justified than their baby related by blood. Along with the development of the

[8] They certainly have a shared history of acquaintance or friendship relation, and such relationship may affect how they anticipate the future of the current romantic relationship.

infant's cognitive capacity of interpreting history and anticipating future, such justification will normally become less and less counterfactual.

For those friends and lovers whose relationship has started, their everyday co-experience begins to accumulate as shared history and develop into a joint future. The interaction between the writing of history and the planning of the future leads to a dynamic interaction between the rewriting of history and the anticipation of the future, which, in turn, lead to a continuous modification of the understanding of the relationship's present identity. While harmony in the interpretation of one end of the timeline may strengthen the harmony at the other end, which may confirm the harmony of the present relationship, a discrepancy at either end may weaken the harmony of the other end, which may bring doubt or even crisis to the relationship. During such interaction, the qualities of the beloved that used to be good indicators of their shared ends may lose their function. New qualities (or new interpretations of the old qualities) are needed to match the new future or the new interpretation of the past, without which a termination of love could be justified. Let us say that John continuously demonstrates his caring and sweetness to Mary, e.g., when Mary is sick and when they have a disagreement. Mary may start to see their relationship as heading towards marriage in the foreseeable future. If John does not see their future in the same way, he may interpret their history as a pleasurable *carpe diem* rather than as preparation for marriage. Such discrepancy in the temporal understanding of their relationship may justify a termination of love. Even if John shares the same anticipation of the future with Mary, this expectation may affect how Mary interprets their past in a way different from John. She used to feel content about John's caring and kind character when she saw their relationship as romantic, but now she wants John to further demonstrate his sense of responsibility and commitment as she foresees them as potential wife and husband. If John fails to demonstrate these new qualities, their love may run into a crisis. Due to the dynamic interaction of the interpretation of history and the anticipation of future in a relationship, the reasons for love require continuous revision, renewal and reconfirmation.

Concerning the constancy of love, TRT suggests that one is not justified to stop loving someone even if their character or qualities change as long as

they share a similar temporality of the relationship. Changes in qualities only matter when they affect how the beloved interprets the past or/and anticipates the future. Even in such cases, a stabilising force can be found within a relationship. Specifically, even if they anticipate a different future, their shared history may provide resources, e.g., shared values, ways of communication, valuable memory, to help adapt to, accommodate or fix such discrepancies. For instance, if Mary wants three children and John declares that he only wants one after witnessing the chaos in his friend's triplet home, they may reach the conclusion that having one child rather than three may best fit what they both want, based on their shared experience of spending quality time alone with each other and on their consensus that social life, in addition to family life, is important. However, even if they interpret the past differently, their shared future plan may provide motivations for them to adapt to, accommodate or fix such discrepancies. For instance, let us say that Mary regards their past as a bitter adaptation to each other's personality and John learns to see it as a sweet interaction with someone different from himself after a few classes of meditation. If they both anticipate their relationship as running towards a marriage, such shared anticipation of the future may make both of them reinterpret their past as a collective bittersweet effort to become harmonious with each other. In this way, the constancy of love is preserved over change in qualities.

The personal identity of the beloved may contribute to the constancy of love if we understand it as the temporality of the beloved which is, by itself, a partial constitutive component and a cause of the temporality of the relationship. If John firmly regards himself as a father-figure, he is likely to turn his relationship with Mary into a family-establishing relationship, e.g., show her his sense of commitment and responsibility. Once their relationship acquires such a nature, John not only identifies himself as a father but also as a father to the children with Mary. In this way, his identity is constituted by the identity of the relationship. Such casual-cum-constitutive relation[9] between the identity of the beloved and the identity of the relationship may prevent a comparison of qualities between the beloved and others that threatens the constancy of love. For instance, Mary is not justified in replacing John with a qualitatively identical person, Jones, because

[9] I borrow the idea of a causal-cum-constitutive relation from Whiting (2013).

Jones' identity is different from John's in the sense that the former is not constituted by a relationship with Mary as was John. Without any shared history as a resource to overcome changes in qualities or discrepancy in anticipation of future, it is highly unlikely that Jones and Mary share a similar temporality of their relationship as much as John and Mary. Even if the qualities of our beloved have been changed so drastically that it renders impossible the recalling of a joint history or the anticipation of a joint future, as long as the shared history or future is sufficiently rich, one may still be justified to continue loving them. Even if John suffers from Alzheimer's disease that causes amnesia or is in a permanent vegetative state (PVS), thus preventing him from making any future plans, Mary's love towards him may still be justified if their shared future (in the case of Alzheimer's disease) or shared history (in the case of PVS) is rich enough to let Mary continue to believe that John is still who he is and that their relationship is still what it is.

Defining the role of the beloved's identity in this way, we can also explain why love can be justifiably terminated even if the beloved's identity does not change. The relationship is just part of the beloved's identity. Other parts may interfere and create discrepancies in the temporality of their relationship. Back to our example, being the father of Mary's child is only part of John's identity. He may also identify himself as a soldier who fights for his country. If John was being promoted and required to spend time in overseas training and missions, he may adjust his role of fatherhood and anticipate only one child with Mary. If Mary expects more than one child in their relationship, such a discrepancy may serve as a reason for the termination of her love. In this case, we can say that such termination of love is justified even though John's identity does not change. In this sense, the constancy of love ensured by TRT is less rigid than by the person theories.

7. Conclusion

Despite the theoretical advantages of TRT in explaining the selectivity and constancy of love, one may still ask, why does temporality matter to love? Other than giving a logical reason or a scientific explanation of the

importance of temporality, I can only outline what it means to take the temporality of a relationship seriously by comparing it with other theories. For quality theory, our beloved possesses qualities that satisfy our desire. Thus, love is a matter of desire satisfaction. For person theory, our beloved is the person who deserves our unselfish concern. Thus, love is a matter of giving and devotion. For relationship theory, our beloved is our partner to achieve a specific end. Thus, love is a matter of collective effort to achieve an end. For the temporal relationship theory, our beloved is our partner to work for some ends from the beginning till the end. Thus, what matters is not the achievement of the end, but the accompaniment along the way towards the development, achievement, or even abandonment of such end. It is about satisfying desires *with someone*. It is about devoting one's care and concern *with someone*. It is all about accompaniment. If every justification must stop somewhere, that's where TRT grounds.

I must also admit that the temporal relationship theory I propose is not a parsimonious theory. The dynamic interaction between interpretation of past, the understanding of the present and the anticipation of future, concerning various important identity-shaping ends of a relationship, can be so complex that it almost renders it impossible to judge whether love is justified or not at any given moment. Taking the complexity of love into consideration, the corresponding complexity of a theory that justifies different kinds of love should not be surprising.

Acknowledgements

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. UGC/FDS14/H11/20).

References

Abramson, Kate & Leite, Adam. (2011). "Love as a Reactive Emotion", *Philosophical Quarterly*, 61(245):673–99. https://doi.org/10.1111/j.1467-9213.2011.716.x

Badhwar, Neera Kapur. (1987). "Friends as Ends in Themselves", *Philosophy & Phenomenological Research*, 48: 1–23. https://doi.org/10.2307/2107703

Bagley, Benjamin. (2015), "Loving Someone in Particular", *Ethics*, 125: 477–507. https://doi.org/10.1086/678481

Barnes, Jonathan, ed. (1984). *The Complete Works of Aristotle, Volumes I and II*, Princeton: Princeton University Press.

Delaney, Neil. (1996). "Romantic Love and Loving Commitment: Articulating a Modern Ideal", *American Philosophical Quarterly*, 33: 375–405

Frankfurt, Harry. (1999). "Autonomy, Necessity, and Love", in *Necessity, Volition, and Love*, Cambridge: Cambridge University Press, 129–41. https://doi.org/10.1017/cbo9780511624643.012

Friedman, Marilyn. (1998). "Romantic Love and Personal Autonomy", *Midwest Studies in Philosophy*, 22: 162–81. https://doi.org/10.1111/j.1475-4975.1998.tb00336.x

Grau, Christopher. (2010). "Love and History". *Southern Journal of Philosophy*, 48(3):246-271. https://doi.org/10.1111/j.2041-6962.2010.00030.x

Heidegger, Martin (1927). *Being and Time*, translated by J. Macquarrie and E. Robinson. Oxford: Basil Blackwell, 1962

Helm, Bennett. (2010). *Love, Friendship, and the Self: Intimacy, Identification, and the Social Nature of Persons*, Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199567898.001.0001

Helm, Bennett. (2017). "Love", The Stanford Encyclopedia of Philosophy (Fall 2017 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/fall2017/entries/love/

Irwin, Terence. (1988), *Aristotle's First Principles*. Oxford: Oxford University Press.

Jollimore, Troy. (2011). *Love's Vision*, Princeton, NJ: Princeton University Press.

Kolodny, Niko. (2003). "Love as Valuing a Relationship". *Philosophical Review*, 112(2):135–89. https://doi.org/10.1215/00318108-112-2-135

Kraut, Robert. (1986). "Love De Re". *Midwest Studies in Philosophy*, vol. 10(1):413–30. https://doi.org/10.1111/j.1475-4975.1987.tb00549.x

LaFollette, Hugh. (1996). *Personal Relationships*. Cambridge: Blackwell.

Lazarus, Richard & Smith, Craig. (1993). "Appraisal Components, Core Relational Themes, and the Emotions". In N. Frijda (Ed.), *Appraisal and Beyond* (pp. 233-270). Hillsdale, NJ: Erlbaum. https://doi.org/10.1080/02699939308409189

Nozick, Robert. (1989). "Love's Bond", in T*he Examined Life: Philosophical Meditations*, New York: Simon & Schuster, 68–6.

Parfit, Derek. (1984). *Reasons and Persons*. New York: Oxford University Press.

Price, Carolyn. (2012). "What is the Point of Love?", *International Journal of Philosophical Studies*, Vol. 20(2), 217–37. https://doi.org/10.1080/09672559.2011.629367

Rorty, Amelie. (1986). "The Historicity of Psychological Attitudes: Love is Not Love Which Alters Not When It Alteration Finds". *Midwest Studies in Philosophy*, vol. X., ed. by Peter A. French, Theodore E. Uehling, and Howard K.

Wettstein, 399–412. Minneapolis: University of Minnesota Press. https://doi.org/10.1111/j.1475-4975.1987.tb00548.x

Setiya, Kieran. (2014). "Love and the Value of a Life", *Philosophical Review*, 123: 251–80. https://doi.org/10.1215/00318108-2683522

Velleman, David. (1999). "Love as a Moral Emotion", *Ethics*, 109: 338–74. https://doi.org/10.1086/233898

Whiting, Jennifer. (1991). "Impersonal Friends." *The Monist*, 74(1): 3–29. https://doi.org/10.5840/monist19917414

Whiting, Jennifer. (2013). "Love: Self-Propagation, Self-Preservation, or Ekstasis?", *Canadian Journal of Philosophy*, 43: 403–29. https://doi.org/10.1080/00455091.2013.857131

Organon F 32 (1) 2025: 29–41 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32102

RESEARCH ARTICLE

An Objection to Railton's Full-Information Analysis of Non-Moral Value

Xuanpu Zhuang*

Received: 22 November 2022 / Revised: 30 August 2024 / Accepted: 11 February 2025

Abstract: This paper presents an objection to Peter Railton's full-information account of non-moral value. According to this account, if an idealized individual A who is fully rational and has full information wants the non-idealized A to desire X, then X is good for A. Those desires like X are called objective interests. Railton's analysis holds that non-moral values are constituted by natural facts that are independent of subjective opinions. I argue that it is hard for the full-information analysis to achieve all its goals. My discussion focuses on intrinsic interests—those good for an individual without reference to any other objective interests. I attempt to show that either it is hard for the full-information account to give a normative force on individuals, or the account is circular. The conclusion I reach is moderate: in its current version, the full-information analysis of non-moral value cannot explain the link between the normative and the empirical.

Keywords: Intrinsic interests; moral realism; normativity; the full-information analysis of non-moral value.

* Zhejiang University
https://orcid.org/0000-0003-4635-1925
School of Philosophy, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, 310058, P.R. China
xuanpuzhuang@outlook.com

1. Introduction

Many philosophers believe that morality is objective. Some of them claim that moral statements are essentially beliefs, and there exist moral facts that make those beliefs true or false. These philosophers are called moral realists (Finlay 2007). Among moral realists, some think that moral facts are constituted or realized by natural facts and therefore endorse ethical naturalism (Suikkanen 2016). Although there are different forms of naturalist moral realism, Peter Railton's realist account might be one of the most successful (Railton 1986a, 1986b, 2003). According to his account, moral judgments could bear truth values, and moral facts are constituted by natural facts and may be reducible to them. Holding that moral properties are objective in this sense, Railton (1986b, 164-165) thinks such a theory could establish the objectivity of morality.

Many people, especially anti-realists, however, believe that there is an intuitive gap between the empirical and the normative. It seems we cannot derive normative force from pure empirical facts.[1] To refute such a view, Railton attempts to provide a realist analysis of ethics.[2] His strategy is to develop the realist notion of moral rightness, and thus he begins with a realist account of non-moral value (Railton 1986b, 171-189). Many called Railton's account as a *full-information analysis* of non-moral value, because according to this account, what is good for a person is defined by reference to an idealized notion with full information (Rosati 1995a, 1995b, Sobel 1994, and Baker 2016).[3] Specifically, if an idealized person A who is fully rational and has full information wants non-idealized A (i.e., the actual one) to desire X, then X is good for A. Those desires like X are called objective interests, because it seems they are not dependent on what non-idealized A

[1] This statement could be called "Hume's Law," as people usually believe that Hume in *A Treatise of Human Nature* says that no ought-judgment may be correctly inferred from a set of premises expressed only in terms of "is" (Hume 2007).

[2] Railton gives some reasons to reject such an intuitive distinction between the empirical and the normative as well, which I believe is closely related to his positive arguments in constructing the link (Railton 1986a, 5-31; 1986b, 166-171).

[3] Some such as Lin (2017, 2019), however, argue that there is no need for such an idealized analysis when conceiving welfare.

conceives. Railton thinks that non-moral values supervene on or even are reducible to natural facts, which are independent of subjective opinions. Those values could also provide a normative force on individual A. If so, the full-information analysis seems to indicate that there is a powerful link between the empirical and the normative. In this paper, I argue that it is difficult for Railton's account to achieve all his goals.

My discussion focuses on intrinsic interests, i.e., those good for an individual without reference to any other objective interests. In Section 2, I reconstruct Railton's full-information analysis. Section 3 introduces an example of intrinsic interests, in which Railton's account may be on the horns of a trilemma about how we should understand the choice of the idealized individual. If the theory implies that we as observers could not know the individual's choice, such an account may end up in a circular argument. If we as observers do know it, possible explanations may be focused on present states or idealized states. The former one may not be independent of the individual's subjective states anymore and thus could not be seen as a successful realist explanation. The latter one seems hard to produce a normative force on the individual. After that, I consider some objections and respond to them in Section 4. In the end, I hold that the full-information analysis needs more work to construct the link between the empirical and the normative.

2. Railton's Full-Information Analysis of Non-Moral Value

To argue for his realist account of non-moral value, Railton (1986b, 172-186) has at least four premises. First, he defines non-moral value as something desirable or good for someone. In this sense, Railton sees subjective interests—individuals' wants or desires—as a secondary quality. That is, they are like tastes, e.g., sweetness. They supervene upon the primary qualities of the individual, the object, and the circumstance, even though they will excite a subjective inclination or sensation. Railton calls primary qualities the *reduction basis* of the secondary qualities. Second, Railton believes that if this account of non-moral value could help explain our experience and give a normative force, then this account is plausible. Third, a realist account could succeed in explaining our experience, only if the non-moral

values postulated in it have two characteristics: independence and feedback. Independence means that those facts exist, and their features are independent of our opinions. Feedback means that we can interact with those facts, and they are able to affect our thoughts and actions. Fourth, Railton introduces norms of individual rationality as a so-called criterial explanation. An individual's rationality is primarily defined in terms of relative efficiency given the agent's beliefs and desires.[4] In other words, we see an individual rational when he could adopt efficient means to certain ends he believes or desires. Based on this, we could evaluate someone more rational than others. Their rationality is assessed relative to their beliefs and desires.

Railton proceeds to define the non-moral value, i.e., what is good for someone, with an idealized notion of the individual. Specifically, if an idealized individual A who is fully rational and owns full information wants the non-idealized A to desire X, then X is desirable or good for A. In Railton's example, Lonnie is a person with malaise and idealized Lonnie knows clear liquids instead of milk will help him. In this case, idealized Lonnie will want the non-idealized self to desire clear liquids instead of milk. If so, clear liquids are a good thing for Lonnie. Railton calls these desires objective interests. Now that idealized Lonnie with full information about his conditions and circumstances will decide on the non-moral values, we may admit that whether X is good for someone depends on facts about his or her circumstances and constitution. In this way, such facts in the so-called *reduction basis* are independent of one's opinions. Besides, individuals could also learn and modify their interests through experience to lead to satisfactory results for themselves. According to Railton, we may utilize a wants/interests mechanism to explain such evolutions of desires. For example, in the case of Lonnie, he may happen to drink some clear liquids that help him. And because of it, he may consciously or unconsciously choose to drink clearer liquids instead of milk when having a malaise again. Gradually he develops such a desire to drink clear liquids in similar circumstances. In other words, the individual may change his desires and interests in trials to lead to his satisfaction. Therefore, Railton thinks that this account satisfies the requirements of independence and feedback.

[4] The understanding of rationality is intuitive here. Accordingly, I do not consider disagreements on how to understand rationality properly in this paper.

Given that individual rationality can be evaluated with efficiency relative to present desires and beliefs, it could be also assessed relative to objective interests (Railton 1986b, 184-189). After all, with the wants/interests mechanism individuals need to take actions that reflect new interests. Railton seems to believe that it is a psychological fact that if one's desire is not supported by the idealized self, this will count against acting upon this desire.[5] If so, this account seems to give a normative force to individuals. Consider the case of Lonnie again, the desire to drink milk is not supported by the idealized Lonnie, as the latter will want non-idealized Lonnie to drink clear liquids. Thus, such an action is not efficient relative to objective interests. Drinking clear liquids is more efficient and thus a more rational action. So, there is a reason for non-idealized Lonnie to drink clear liquids, even though it goes against his own motivation. To sum up, it provides a normative force on non-idealized Lonnie, which finally shows that drinking clear liquids is good for Lonnie.

So far, Railton's full-information analysis seems to explain values, display the process of feedback, and indicate the key link between the empirical facts and the normative force. Railton concludes that this full-information analysis of non-moral value not only succeeds in explaining our experience but gives a normative force as well. If so, it would provide a stable foundation for moral realism.

3. The Trilemma of the Full-Information Analysis

In this section, I discuss an example of intrinsic interests and then argue that Railton's full-information analysis is on the horns of a trilemma.

Above all, note that Railton distinguishes intrinsic goodness from other non-moral ones: X is intrinsically non-morally good for a person A if X is in A's objective interests without reference to any other objective interests of A (Railton 1986a, 17; Railton 1986b, 178). In the case of Lonnie, the desire to drink clear liquids is not intrinsically good. Instead, health would

[5] It is perhaps seen as a form of internalism, i.e., if one ought to φ in a circumstance, one must be motivated to φ given the circumstance (Sobel 2001, Schroeder 2007, and Mason 2008).

be intrinsically good for Lonnie.[6] When Railton claims that Lonnie has a normative reason to drink clear liquids, it is only an explanation in terms of instrumental reasons. That is, drinking clear liquids is the means to health, which is efficient and thus desirable as long as Lonnie desires health.

However, it is unclear how Railton can explain the reduction basis of intrinsic interests. To see this, consider a case of listening to music. Intuitively, listening to music is intrinsically good for someone but not good for others. Imagine that a person David also finds himself miserable in a foreign country. He becomes aware of a desire to turn on the radio, although there is only boring news. Unknown to David, there is a music festival near his apartment. Idealized David would want non-idealized David to have a desire to go out to attend the music festival. Nevertheless, are *we* able to tell what idealized David would want non-idealized David to want?

To clarify, listening to the music itself is non-morally good for David. That is, it is *intrinsically* good. So, the idealized David would want non-idealized David to have the desire to go out to attend the music festival.[7] But it is unclear whether *we* can tell this, or how we could *explain* what idealized David wants. I argue that Railton's account is on the horns of a trilemma here: there are three possibilities to deal with this case of intrinsic interests for Railton, but no matter which explanation we give to the question of intrinsic interests, the full-information analysis will lead to an unsatisfactory answer.

3.1. The First Aspect: If We Do Not Know

According to Railton, the condition of full information only includes descriptive information. We want to reduce the goodness to other descriptive facts. As we have different intrinsic objective interests, however, we do not know others' interests. If so, we cannot tell what is desirable or good for David. In this case, only if we know that listening to music is intrinsically non-morally good for David can we know that the idealized David would want the non-idealized David to go out to the music festival.

[6] For more reasons to see health as intrinsically good, see (Raibley 2013).

[7] Some may doubt whether it is possible to compare two possible choices in one's life, which I do not consider in this paper (Baumann 2018).

Nevertheless, we have no chance to know that it is non-morally good for David, given full information and full rationality. Only the idealized David could manage it. If so, it becomes doubtful whether intrinsic non-moral values could be constructed on a reduction basis. After all, it seems the intrinsic interests here depend on some facts that cannot be described objectively or independently.

Railton may reply that the existence of intrinsic goodness is a fact, even though we do not know the content. That is, we could still define intrinsic goodness with the realist account when it is unknown to us. But the point here is that this account includes certain unexplained personal desires. In other words, what is good for person A is fundamentally based on what A *thinks* is good for A. If so, the account may end up in a circular argument and thus fails to provide a reductive account for subjective interests.[8]

If the full-information account of non-moral values is not successful in constructing intrinsic interests on a reduction basis, the realist explanation of the normative force is also implausible. As we can see, the notion of individual rationality is relative to objective interests. If intrinsic interests are fundamental, it implies that the reason to have a certain act that reflects any interests is fundamentally based on intrinsic interests. According to Railton's account, the actual individual A has reason to φ because the idealized individual A wants the actual A to φ. If an intrinsic interest has no satisfying reduction basis, however, such a reason may exist in A's own thoughts. The latter will give A the motivating reason to act, but it does not display the link between the normative and the empirical. And therefore, it is doubtful whether this realist account of non-moral values succeeds.

3.2. The Second Aspect: If We Do Know

To avoid such a direction leading to a circular argument, one may hold that we could know that the idealized David would want the non-idealized David to go out to the music festival. However, it is still problematic

[8] A reductive account is supposed to reduce entities *a* to entities *b*, i.e., provide an analysis of *a* notions in certain other notions such as *b*. In a reductive account of *a*, we do not have the terms of *a* notions anymore.

whether possible naturalist explanations could provide a link between the empirical and the normative.[9]

On the one hand, some may say that the reason why we could know idealized individuals' choices is still based on the present state of the individuals. To produce a normative force, they may appeal to the present desires of the non-idealized David. Consider the case of David again. They may hold that given the whole experience and knowledge, it is reasonable to know that David is a fan of music. If David is a fan of music, then it is natural to think that Davie will desire to go out to the music festival once he knows there is one outside. So, in this case, we know that the idealized David will want the non-idealized David to desire to go out to the music festival. The problem is that, although most individuals who have intrinsic interests in music are fans of music, it is not necessarily so. For example, someone may have never listened to music, even if they would become fans of music after having such an experience. There are plenty of similar descriptions, in which some musicians always describe their first experience in music as amazing, while they have never known anything about music before. Furthermore, this explanation depends on the present desires of an individual. In this sense, what constitutes values is not independent of the individual's mental state. That is, only if an individual has a motivation for X beforehand, can the choice of the idealized individual have a normative force on the non-idealized one.

Another explanation may be based on the idealized states. One may say that with full information it is reasonable to know that an individual would be a fan of music in the idealized condition. Accordingly, we could know that the idealized David as a fan of music would want the non-idealized David to listen to music. Nevertheless, even if we assume that given full information we could know the outcomes of different possible experiences, becoming a fan of music is still not a sufficient indicator of having an intrinsic interest in music.[10] For example, some fans of music may only enjoy

[9] For a general objection to the explanations provided by moral reductionists including Peter Railton, see (Zhong 2012).

[10] Rosati and Sobel discussed how full information is obtained by the idealized self. Rosati thinks there are some difficulties in comparing different possible outcomes and experiences for the idealized self, and Sobel argues that the full-information

the atmosphere of concerts instead of music. Furthermore, although we may know that the idealized David would be a fan of music and thus want the non-idealized David to listen to music, this fact itself still does not provide a reason for non-idealized David to have a desire to go out for the music festival. In this case, David may become a fan of music given more information about music in the idealized condition, but the interests of the idealized David seem not to provide any benefits for the non-idealized one. After all, the non-idealized David may have no feeling about music at all now.[11] If so, it is counter-intuitive to say that it is good for David to go out to the music festival, which he does not enjoy at all at this moment. In conclusion, it is hard to construct a stable link between the behaviors or other states of being a fan of music and having intrinsic interests in music.[12] And therefore, we cannot say it is good for David to go out for the music festival.

3.3. The Trilemma

In conclusion, the full-information analysis of non-moral value is on the horns of a trilemma to answer the question about intrinsic values: if we do not know what the idealized self will want the non-idealized self to desire, it seems this account may lead to a circular argument and fails to be a naturalist account; if we know the choices of the idealized self, however, we may turn to an account based on present states that would not be independent of internal psychology, or turn to an account based on idealized states that lacks a normative force—the main goal of this account. In short, it is difficult for the full-information analysis to keep its basic commitments and achieve its goal together.

---

account omits some important limitations in our psychological and cultural facts, which makes our well-being incommensurable as well. See (Rosati 1995b, 296-325; Sobel 1994, 784-810; Shemmer 2011). I tend to agree with them, but even if their arguments are problematic, it does not hurt my argument here.

[11] As one's desires are changing and unstable, the analysis of welfare based on the desires is temporal (Dorsey 2013).

[12] Bykvist (2010) argues that how you would have felt about a life had you never led it is irrelevant to the question of how good that life is for you. Tiberius (1997) also doubts whether the demand made by an idealized person is intuitive or not.

4. Possible Objections and Responses

In this section, I respond to several possible objections.

First, some may claim that we all share similar intrinsic interests, as Railton admits in his paper (Railton 1986b, 178).[13] That is, all human beings share some basic common interests, which permits us to understand one another's interests. In this way, we could only focus on non-intrinsic interests when providing the full-information analysis of non-moral value, which avoids the trilemma. I doubt the universality of intrinsic interests. Furthermore, it cannot explain why we could know what the idealized self would want the non-idealized self to desire. Consider the case of David again, even though music is intrinsically good for many people, it does not mean that we know it applies to David. In other words, the common basis that leads to similar intrinsic interests shared by human beings could promote our understanding of the choice of idealized individuals but does not help explain how we can know the choice.

Another objection may hold that listening to music is not intrinsically good, as listening to music makes people happy. They may think happiness or pleasure is the only intrinsic interest. As we know the relationship between the means and the end, we could know that the idealized David would want the non-idealized David to go out to the music festival. But this idea takes a risk in holding happiness or pleasure as the only intrinsic non-moral value. We usually all accept the existence of other possible intrinsic values such as liberty, justice, and friendship.[14] It is also intuitive to me that listening to music could be intrinsically good at least for some. A naturalist account in a hedonist version may manage to give a normative force, but it fails to explain our intuitions.

Some may hold that my argument is demanding because health as an intrinsic value could also be questioned in the same way. Consider the original case of Lonnie again, we may ask whether we could know that the idealized Lonnie will want the non-idealized Lonnie to desire to drink clear liquids. But it seems absurd, as it seems obvious that we know that. If so,

[13] For some disagreements, see (Sobel 1999).

[14] I do not want to touch on the theory of well-being or axiology in this paper. Instead, what I focus on are intuitive and ordinary ideas about values.

the arguments will be out of work. However, I do not want to deny the same trilemma applies to the case of Lonnie. That is, the full-information analysis itself does not provide a plausible method for us to know whether the idealized Lonnie would want the non-idealized Lonnie to drink clear liquids to keep healthy. It is the fact of similar interests that makes us have a successful conjecture, and it leads us to think such a full-information analysis is successful, as we already have similar interests with Lonnie. But this is an illusion. The fact of similar interests could not play a significant role in the full-information analysis, as we have seen above. The reason I choose music instead of health as the focus is that music is more obvious to be accepted as unknown intrinsic goodness, as it seems not intrinsically good for everyone.

5. Conclusion

In this paper, I argue that the full-information analysis of non-moral values fails to give a satisfactory realist account of non-moral values. According to Railton, if the idealized self A who has full information and full rationality wants the non-idealized self to desire X, X will be good for A. The problem is whether we know what the idealized self will want the non-idealized self to desire. In the case of intrinsic values, if the theory implies that we do not know what the idealized self will want the non-idealized self to desire, it seems such an account may lead to a circular argument; if we are thought to know the choices of the idealized self, however, we may turn to an account based on present states which would not be independent, or turn to an account based on idealized states which lacks a normative force—the main goal of this naturalist account. In short, either full-information analysis fails to be a serious realist account, or it cannot achieve the goal of producing a normative force on the individual.

There are some alternatives to developing the full-information analysis of non-moral value. First, there may still be other methods for us to know what the idealized self would want the non-idealized self to desire, and they can be described in natural terms. Second, there may be further explanations of how subjective judgments of an individual could be constituted by natural properties. In other words, even if we have no access to the choices of the idealized

self, the normative forces are still provided by independent natural facts about the individual. Third, there may exist other different explanations of intrinsic interests defined in objective terms and thus avoid the trilemma. However, the current version of the full-information analysis is unsuccessful.

In sum, naturalist moral realists need to provide more explanations about how the full-information analysis works, and it means they have more burden of justification. If so, moral realism based on such a realist account of non-moral value will be affected deeply as well, but that is not in the scope of this paper.

References

Baker, Derek Clayton. 2016. "Deliberators Must Be Imperfect." *Philosophy and Phenomenological Research*, 93(2): 321–47. https://doi.org/10.1111/phpr.12199

Baumann, Peter. 2018. "What Will Be Best for Me? Big Decisions and the Problem of Inter-World Comparisons." *Dialectica*, 72(2): 253–73. https://doi.org/10.1111/1746-8361.12219

Bykvist, Krister. 2010. "Can Unstable Preferences Provide a Stable Standard of Well-Being?" *Economics and Philosophy*, 26(1): 1–26. https://doi.org/10.1017/S0266267110000039

Dorsey, Dale. 2013. "Desire-Satisfaction and Welfare as Temporal." *Ethical Theory and Moral Practice*, 6(1): 151–71. https://doi.org/10.1007/s10677-011-9315-6

Finlay, Stephen. 2007. "Four Faces of Moral Realism." *Philosophy Compass*, 2(6): 820–49. https://doi.org/10.1111/j.1747-9991.2007.00100.x

Hume, David. 2007. *A Treatise of Human Nature: A Critical Edition*, edited by David Fate Norton and Mary J. Norton. Oxford: Clarendon. Press. https://doi.org/10.1093/actrade/9780199596331.book.1

Lin, Eden. 2017. "Against Welfare Subjectivism." *Noûs*, 51(2): 354–77. https://doi.org/10.1111/nous.12131

Lin, Eden. 2019. "Why Subjectivists About Welfare Needn't Idealize." *Pacific Philosophical Quarterly*, 100(1): 2–23. https://doi.org/10.1111/papq.12232

Mason, Elinor. 2008. "An Argument Against Motivational Internalism." *Proceedings of the Aristotelian Society*, 108(1pt2): 135–56. https://doi.org/10.1111/j.1467-9264.2008.00240.x

Raibley, Jason. 2013. "Health and Well-Being." *Philosophical Studies*, 165(2): 469–89. https://doi.org/10.1007/s11098-012-9951-2

Railton, Peter. 1986a. "Facts and Values." *Philosophical Topics*, 14(2): 5–31. https://doi.org/10.5840/philtopics19861421

Railton, Peter. 1986b. "Moral Realism." *Philosophical Review*, 95(2): 163–207. https://doi.org/10.2307/2185589

Railton, Peter. 2003. *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511613982

Rosati, Connie S. 1995a. "Naturalism, Normativity, and the Open Question argument." *Noûs*, 29(1): 46-70. https://doi.org/10.2307/2215726

Rosati, Connie S. 1995b. "Persons, Perspectives, and Full Information Accounts of the Good." *Ethics*, 105(2): 296–325. https://doi.org/10.1086/293702

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199299508.001.0001

Shemmer, Yonatan. 2011. "Full Information, Well-Being, and Reasonable Desires." *Utilitas*, 23(2): 206–27. https://doi.org/10.1017/S0953820811000069

Sobel, David. 1994. "Full Information Accounts of Well-Being." *Ethics*, 104(4): 784–810. https://doi.org/10.1086/293655

Sobel, David. 1999. "Do the Desires of Rational Agents Converge?" *Analysis*, 59(3): 137–47. https://doi.org/10.1093/analys/59.3.137

Sobel, David. 2001. "Explanation, Internalism, and Reasons for Action." *Social Philosophy and Policy*, 18(2): 218. https://doi.org/10.1017/S026505250000296X

Suikkanen, Jussi. 2016. "Naturalism in Metaethics." In *Blackwell Companion to Naturalism*, edited by Kelly James Clark, 351–368. Hoboken: Wiley-Blackwell. https://doi.org/10.1002/9781118657775.ch25

Tiberius, Valerie. 1997. "Full Information and Ideal Deliberation." *Journal of Value Inquiry*, 31(3): 329–38. https://doi.org/10.1023/A:1004295329083

Zhong, Lei. 2012. "An Explanatory Challenge to Moral Reductionism." *Theoria*, 78(4): 309–25. https://doi.org/10.1111/j.1755-2567.2012.01147.x

Organon F 32 (1) 2025: 42–49 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32103

RESEARCH ARTICLE

Sadness is not about Loss

Maria Zanella*

Received: 4 July 2024 / Accepted: 4 January 2025

Abstract: I argue that sadness is not about loss. I present two counterexamples to the Loss View: the view that if one is sad about something one takes it to be a loss. I suggest an alternative and defend it.

Keywords: sadness; loss; badness; emotion; formal object.

§1 The Loss View

Philosophers of emotion have often claimed that sadness requires taking something to be a loss. For example:

> Although different episodes of sadness may be related to particular objects as disparate as the disappearance of a loved one, failure at an exam, the melting of a glacier, and so on, the subject who feels sad nevertheless apprehends all these objects as losses. (Deonna and Teroni 2014, 17)

Let us call this view the Loss View:

Loss View: if one is sad about something, one takes it to be a loss.

* University of Warwick
https://orcid.org/0009-0005-3037-7006
Department of Philosophy, Social Sciences Building, University of Warwick, Coventry, CV4 7AL, United Kingdom
mariaz.correspondence@gmail.com

According to the Loss View, even if we can be sad about many different kinds of things, whenever we are sad about something, we take it to be a loss; on this view, all possible episodes of sadness require taking something to be a loss.[1] This view is common: it has been endorsed in (Lazarus 1991, 122), (Solomon 1993, 186), (Roberts 2003, 234), (Prinz 2004, 61), (Müller 2019, 37–38, 43–44) and (Deonna and Teroni 2022, 1–2).

I have brushed over the differences between the views of these philosophers, stating the Loss View using the verb 'take'. One is committed to the Loss View if one endorses more specific views to the effect that, in order to be sad about something, it is necessary to perceive/judge/believe/imagine something to be a loss. Solomon, for instance, is committed to the Loss View because he writes: "My sadness, my sorrow, and my grief are judgments of various severity to the effect that I have suffered a loss" (Solomon 1993, 186).[2]

Before we move to the counterexamples, a word on 'loss'. Defenders of the Loss View write as though all of their examples of things one could be sad about could be taken to be of a single kind of event (or relation) which can be designated with the word 'loss'. Indeed, if they are to specify what all possible episodes of sadness have in common, it is *necessary* that all of their examples be of such a kind. But there may be no such kind. Does something happen when one 'loses' one's mother which is of the same kind as what happens when one 'loses' one's job? What about when one 'loses' hope or when one 'loses' a watch? I doubt whether there is a single kind of 'loss'-event which all of these 'losses' belong to. Importantly, similar things can be said of 'have'. Regardless of whether there is a single kind of 'loss'-event, it is nevertheless true that whenever one takes a 'loss'-event to have occurred one must think that there was something which then there wasn't

[1] The Loss View is entailed by the claim that loss is sadness' 'formal object'. See, e.g., (Teroni 2007, 408) and (Müller 2019, 37). On formal objects see (Kenny 2003, 132–135).

[2] For the purposes of this paper, the differences between the Loss View theorists do not matter. If you believe there is no such genus as *taking* of which *perceiving*, *judging*, *imagining*, etc. are species, then imagine my counterexamples to be counterexamples to the more specific theses, which would, were there to be such a genus as *taking*, entail the Loss View.

or that something was had (in some sense of 'have') which then was no longer had.

I will speak as if there is a single kind of 'loss'-event which all of the 'losses' the Loss View theorists speak about belong to. If there is no such kind, then I should be understood to be arguing for the strongest conclusion—namely, that in order to be sad about something, one needn't take it to be a 'loss'-event of any kind.

§2 The Loss View is false

Here are two counterexamples to the Loss View:

Counterexample 1. Anne (of *Anne of Green Gables*) is sad about having red hair because she thinks it is ugly. Anne does not take her having red hair to be a loss.

Counterexample 2. Bob (of *A Christmas Carol*) is sad about being poor because he cannot afford treatment for his son Tim. Bob does not take his being poor to be a loss.

It is not even clear what it would be for Anne to take her having red hair to be a loss. Nor is it clear what it would be for Bob to take his being poor to be a loss.

Similar cases of sadness which do not seem to involve any thought of loss are cases of sadness on account of friendlessness, congenital illness, or involuntary childlessness. These are things people can be, and are, sad about without taking them to be losses.

In light of these counterexamples, one might think that this weaker alternative to the Loss View ought to be considered:

Weak Loss View: if one is sad about something, one takes something to be a loss.

But the Weak Loss View fares no better. In being sad, Anne doesn't take herself to have had something which she no longer has nor does she take there to have been something which there now isn't. Nonetheless she is sad. And Bob was always poor and always knew that he was poor.

Objection. These cases involve the subject thinking about some loss of opportunity, and therefore they are not counterexamples to the Weak Loss View. Anne is sad about her red hair because she thinks it comes with a loss of opportunity—in having red hair Anne takes herself to have lost the opportunity to be considered beautiful by her contemporaries. Bob takes himself to have lost the opportunity to pay for treatment.

Reply. It is doubtful whether the subjects of these cases take themselves to have lost opportunities: Anne always had red hair, so she wouldn't have taken herself to have ever had, or lost, the opportunities that having red hair precluded in Victorian Canada; Bob never had the money and knew it, so he would never have taken himself to have had, or lost, the opportunities that money would have afforded him.

Objection. But couldn't Anne think she has lost the 'genetic lottery'; and couldn't Bob think that he has lost some sort of lottery that determined his social position?

Reply. Regardless of whether either of these 'lottery' beliefs are true (I doubt it), we need not attribute them to Anne or Bob in order to make sense of their sadness. In order to be sad that she has red hair, Anne doesn't need to think that there was ever a chance of her not having red hair. Bob might think he was doomed to be poor and still be sad that he is.

What might be true is that the characters compare (what they take to be) a(n) (im)possibility and an actuality: their sadness might well be fuelled by counterfactual(/counterpossible) thinking/imagining. But the Weak Loss View cannot account for this: to take something to be a loss requires comparing two things one takes to be actualities—a before and an after.

There are also cases in which people are sad as a result of comparing (what they take to be) two actualities without them taking anything to be a loss: Anne might be sad about having red hair after comparing her hair with that of her friend Diana; Bob might be sad about being poor after comparing his situation with that of Scrooge. There are no 'losses' here. Nor need the characters think that there are.

§3 Sadness and taking something to be bad

So what must one take something to be in order to be sad about it? What was said about comparisons at the end of the last section suggests an alternative to the (Weak) Loss View:

Worse View: if one is sad about something, one takes it to be worse than something else.

The *worseness* is *worseness for*; specifically worseness for something/someone one cares about (which may well be oneself). You are sad about your father being dead because your father's being dead is worse for you (and him?) than his being alive; Anne is sad about having red hair because she cares about how she looks and she takes having red hair to be worse for her looks than having black hair.

It is, I think, unnecessary to compare things in order to be sad. So I think the Worse View is false. However, it is true that if one makes a comparison one might end up taking something to be bad for something or someone one cares about; and in my view it is *this* which is crucial for sadness. What all cases of sadness have in common (including those I have considered here) is taking something to be bad. Hence the Bad View:

Bad View: if one is sad about something, one takes it to be bad.

The *badness* is *badness for*; specifically badness for something/someone one cares about. Anne is sad about having red hair because she takes it to be bad for her and her looks. Bob is sad about being poor because he takes it to be bad for Tim.[3]

I won't offer further reasons here for preferring the Bad View over the Worse View. But I will, in what follows, defend the Bad View from a

[3] A different connection between sadness and badness is suggested by (Mulligan 2016). For me, when Sam says, truly, e.g. that he is sad that his father is dead: what Sam is sad about is that his father is dead; and what he takes his father's being dead to be is *bad*. Mulligan's discussion of happiness (*ibid.,* 138–139) suggests, rather: that what Sam would *really* be sad about is the badness of his father's being dead, a higher-level state of affairs; and what Sam would take that higher-level state of affairs to be is *unlucky*.

methodological objection in order to show that it is indeed a more promising alternative to the (Weak) Loss View.

Objection.[4] Even if the Bad View is correct, it does not shed light on the nature of sadness. In particular, it does not provide us with a suitable formal object for sadness. Badness cannot be the formal object of sadness, for it it were, the *individuation requirement* would be contravened:

> *Individuation requirement*: each emotion type has its distinctive formal object (Teroni 2007, 399).

After all, when we are, e.g., afraid of something, don't we take it to be bad (for something/someone we care about)? Taking something to be bad is not exclusive to sadness, so badness cannot be sadness's formal object.

Reply. Objections have been raised against the individuation requirement (see Teroni 2007, 401–402). But even supposing that the individuation requirement is correct, that is no problem for the view that badness is sadness' formal object. Suppose badness *is* sadness' formal object and assume that one cannot take something to be a danger (to something/someone one cares about) without taking it to be bad (for something/someone one cares about), and hence that one cannot be afraid of something without taking it to be bad. This wouldn't entail a contravention of the individuation requirement, because it would not entail that sadness and fear have the same formal object. For formal objects (are supposed to) determine the correctness conditions of emotions (Teroni 2007, 399), and one could maintain that sadness' formal object is badness while fear's formal object is dangerousness on the grounds that the correctness conditions of fear are determined by dangerousness (as opposed to badness) and the correctness conditions of sadness are determined by badness.[5] Therefore one could uphold the individuation requirement, maintaining that

[4] I thank Giovanna Colombetti for a comment which led me to develop this objection.

[5] I take sadness to be (primarily) about events and states of affairs; if an event or state of affairs is bad (for something/someone one cares about), sadness towards it is correct. If it is not bad for anything or anyone one cares about, sadness towards it is incorrect.

sadness and fear have different formal objects, even if both required one to take something to be bad.

This discussion focussed on fear, but so long as the correctness of the other negative emotions (anger, shame, jealousy, etc.) is not determined by mere badness, the same defence can be given *mutatis mutandis*. The individuation requirement would only be contravened if there were an emotion other than sadness which could claim badness as its formal object; but there is, I think, no such emotion.

Acknowledgements

I would like to thank Eileen John, Christoph Hoerl, Nathan William Davies, the participants of the 2023 Alpine Fellowship Symposium, and the participants of the 2023 Warwick-Geneva-Leipzig Interdepartmental Workshop for helpful discussions. I would also like to thank the Alpine Fellowship and the AHRC for financial support.

References

Deonna, Julien A., and Fabrice Teroni. 2014. "In What Sense Are Emotions Evaluations?" In *Emotion and Value*, edited by Sabine Roeser and Caine Tood, 15–31. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199686094.003.0002

Deonna, Julien A., and Fabrice Teroni. 2022. "Why Are Emotions Epistemically Indispensable?" *Inquiry* 68(1): 91–113. https://doi.org/10.1080/0020174X.2022.2126147

Kenny, Anthony. 2003. *Action, Emotion and Will*. London: Routledge. https://doi.org/10.4324/9780203711460

Lazarus, Richard S. 1991. *Emotion and Adaptation*. Oxford: Oxford University Press.

Müller, Jean M. 2019. *The World-Directedness of Emotional Feeling: On Affect and Intentionality*. London: Palgrave McMillan. https://doi.org/10.1007/978-3-030-23820-9

Mulligan, Kevin. 2016. "Happiness, Luck and Satisfaction." *Argumenta* 1(2): 133–145. https://doi.org/10.14275/2465-2334/20162.MUL

Prinz, Jesse. 2004. *Gut Reactions: A Perceptual Theory of Emotion*. Oxford: Oxford University Press.

Roberts, Robert C. 2003. *Emotion: An Essay in Aid of Moral Psychology*. Cambridge and New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511610202

Solomon, Robert C. 1976. *The Passions*. New York: Anchor Press/Doubleday.

Teroni, Fabrice. 2007. "Emotions and Formal Objects." *dialectica* 61(3): 395–415. https://doi.org/10.1111/j.1746-8361.2007.01108.x

Organon F 32 (1) 2025: 50–91 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32104

RESEARCH ARTICLE

The Synthetic Concept of Truth and Its Descendants

Boris Čulina*

Dedicated to Jozo and Rajka Ištuk,
victims of war and victims of lies on the Internet

Received: 11 April 2024 / Revised: 15 January 2025 / Accepted: 21 January 2025

Abstract: The concept of truth has many aims but only one source. The article describes the primary concept of truth, here called the synthetic concept of truth, according to which truth does not belong exclusively to us nor exclusively to nature: truth is the objective result of the synthesis of us and nature in the process of rational cognition. It is shown how various aspects of the concept of truth – logical, scientific, and mathematical aspect – arise from the synthetic concept of truth. Related to these aspects, (i) the role of the predicate of truth in the concept of truth is analysed, (ii) Tarski's definition of truth and its role in the concept of truth are analysed, and (iii) the position of the paradoxes of truth in the concept of truth is analysed.

Keywords: truth; truth in logic; truth in science; truth in mathematics; the truth predicate; Tarski's definition of truth; paradoxes of truth.

* University of Applied Sciences Velika Gorica
https://orcid.org/0000-0002-4195-6936
University of Applied Sciences Velika Gorica, Zagrebačka Ul. 5, 10410, Velika Gorica, Croatia.
boris.culina@vvg.hr

> "The ideal subject of totalitarian rule is not the convinced Nazi or the convinced Communist, but people for whom the distinction between fact and fiction (i.e., the reality of experience) and the distinction between true and false (i.e., the standards of thought) no longer exist."
> Hannah Arendt: *The Origins of Totalitarianism* (Arendt 1973)

> [Pontius Pilate]: What is truth?
> [Yeshua]: The truth is, first of all, that your head aches…
> Mikhail Bulgakov: *Master and Margarita* (Bulgakov 1997)

1. Introduction

Many of the ambiguities associated with the concept of truth stem from the fact that the concept has various aspects that are not sufficiently differentiated. Tarski's T-scheme (Tarski 1933) is a classic example of this. T-scheme is a set of T-sentences, the sentences (biconditionals) of the form:

$$T(\ulcorner\varphi\urcorner) \leftrightarrow \varphi^*$$

where "T" is the symbol of the truth predicate, φ any sentence of a language L (usually the language we are considering), $\ulcorner\varphi\urcorner$ is the name of that sentence in a language *ML* (usually the metalanguage in which we consider L), while φ^* is a translation of that sentence into *ML*. To get a concrete example of a T-sentence, I will take the English sentence "Svrco is afraid of thunder" (the language L will be part of the English language), and my native language as the language *ML*:

T("Svrco is afraid of thunder") ↔ Švrćo se boji grmljavine

where $\textit{Svrco is afraid of thunder}^* = \textit{Švrćo se boji grmljavine}$ is a translation of the English sentence into my native language (Croatian). Here the concept of truth appears in five places: as the truth values of the left and right sides of the biconditional, as the truth value of the whole biconditional, as the meaning of the truth predicate symbol "T," and as the truth value of the sentence "Svrco is afraid of thunder." Only the last sentence belongs to the language L, while the other sentences and the symbol "T" belong to the language *ML*. However, all of them have a semantic source in the sentence "Svrco is afraid of thunder" of the language L. The left side of

the biconditional through the symbol T allows to speak in *ML* about the truth value of the sentence "Svrco is afraid of thunder" of the language *L*, the right side of the biconditional is related to the truth value of the translation of that sentence into *ML*, while the truth value of the whole biconditional is related to the success of the translation. Thus, the key aspect of the concept of truth is related to the truth value of the sentence "Svrco is afraid of thunder" of the language *L*, while other aspects are connected to this primary aspect for various reasons.

In what follows, I will focus on this primary concept of truth – the truth values of the atomic sentences of the language *L*, leaving aside the truth values of the sentences of the metalanguage in which I will carry the considerations. After analysing the primary concept of truth, I will consider other aspects of the concept of truth. Related to these aspects, in the last part of the article (i) the role of the predicate of truth in the concept of truth is analysed, (ii) Tarski's definition of truth and its role in the concept of truth are analysed, and (iii) the position of the paradoxes of truth in the concept of truth is analysed.

There is a vast philosophical literature on the concept of truth. Although various aspects of the concept of truth have been addressed (Glanzberg 2023), I have not come across a differentiation of the concept of truth as done in this article.

The basic assumption of the analysis of the concept of truth conducted here is that rational cognition and abstract thinking are in their final form the creation and use of language. A common view of the role of language in rational cognition and thinking is that it plays a passive role there: language is a medium for expressing and communicating thoughts, and for describing reality. The first philosophers to fully recognize the essential role of language in rational cognition and thinking were Hamman, Herder and Wilhelm von Humboldt in the second half of 18th century and the first half of 19th century, and Cassirer later, in the first half of 20th century. In the first half of the 20th century, linguists Sapir and Whorf came to the same conclusion. However, they did not systematically analyse the essential role of language in rational cognition and thinking, and the key role of the concept of truth in it. The essential role of language is systematically analysed in (Čulina 2021). In this article, the key role of the concept of

truth is systematically analysed. Some parts of these articles overlap. In addition to the fact that these parts are now improved, I believe that the importance of the concept of truth deserves a separate article dedicated to that concept.

In the analysis that follows, I will stick to two methodical principles. The first is that all the language of rational cognition can be understood as an extension and improvement of everyday language.[1] Experience in using everyday language is more or less unconsciously transferred to the entire language of rational cognition. I will analyse the hidden assumptions of this generalization that are significant for the concept of truth.

Another methodical principle that I will adhere to is that I will deal primarily with the effects of language forms in rational cognition, and not with the biological, psychological, social, empirical and theoretical processes on which these effects rest. For example, I will primarily deal with the question of what it means for my use of language to possess the full meaning of a language form, and not what the full meaning of a language form consists of. I consider that it is just a proper level of abstraction which, on the one side, explicates all precise effects, and on the other side, hides all complexities and obscurities of the use of language in the process of rational cognition. Of course, it does not mean that the meanings of language forms are not important. Moreover, the meanings are necessary. We cannot use language in rational cognition without the meanings of language forms. However, unlike the determinable effects of language use in rational cognition, the meanings of language forms are too fluent to be able to say something definite without limiting their necessary fluency. Although there are essential differences between what Frege calls "sense" and "reference" and what I call in this paper "the possession of the full meaning" and "semantic value" of a linguistic form, my focus on the effects of language forms corresponds to Frege's insistence on reference: "The reference is thus shown at every point to be the essential thing for science." (Frege 1892a).

[1] This is the language form of Einstein's claim that "The whole of science is nothing more than a refinement of everyday thinking." (Einstein 1936, 349)

2. The Synthetic Concept of Truth

Adhering to the principles stated above, I will begin the analysis with the sentence from everyday language: "Svrco is afraid of thunder." To determine the truth value of the sentence "Svrco is afraid of thunder" we must know the meaning of its parts. Knowledge of English grammar tells us which parts they are and what their linguistic meaning is: "Svrco" is the name of an object, and "is afraid of thunder" is a predicate expression. However, to determine the truth value of the above sentence, we must know exactly which object the word "Svrco" names and what the meaning of the predicate expression "is afraid of thunder" is. Svrco is my only dog, and every connoisseur of English possesses the meaning of the word "is afraid of thunder," despite the fact that we do not know clearly enough what the "meaning of a predicate expression" means. The possession of these meanings is necessary but not sufficient to determine the truth value of the sentence "Svrco is afraid of thunder." We still have to do an appropriate experiment, let nature give its contribution, to determine that it is a true sentence.

This example from everyday language illustrates the basic cognitive situation: the use of a predicate expression leads to the creation of a binary experimental framework that we apply to the named object. We generate a binary experiment in which nature chooses one of the two offered values, *yes* or *no*, as the result of the experiment. We apply a predicate symbolized by "P" to an object a and describe the situation with the declarative atomic sentence "$P(a)$." Two possible results of the application are the so-called *truth values* termed *True* and *False*. We take the result chosen by nature as the truth value of the language form "$P(a)$." *True* and *False* are designed by us as a part of the binary experimental framework and selected by nature in the realization of the experiment. These binary experiments are the essence of our rational cognition. We make the question and offer two possible answers, and nature selects an answer. The selected truth value does not belong exclusively to us nor does it belong exclusively to nature. It is the objective result of the synthesis of us and nature in the process of rational cognition: it differentiates what is from what is not. That is why I have termed this primary concept of truth *the synthetic concept of truth*.

An everyday cognitive situation illustrated and described above, simple as it might seem, has a number of underlying characteristics and assumptions that are essential for the process of rational cognition and that need to be clarified. First, it reflects our innate approach to the world which we divide into objects with which something happens and predicates that determine what happens. This division is not absolute – something that is a predicate in one context can become an object to which other predicates are applied in another context. This *object - predicate dualism* is a fundamental characteristic of the cognitive framework described here. It is reflected in language through the structure of the atomic sentence "$P(a)$." Symbols "a" and "P" have different roles in the sentence. We use symbol "a" to name (mention) an object a. We use symbol "P" to say something about the object a. The symbol "P" does not name anything: it leads to a certain binary experiment on the object a.

To my knowledge, Whorf is the first one to recognise that the object-predicate dualism is a prominent feature of Indo-European languages: "Our language thus gives us a bipolar division of nature. But nature herself is not thus polarized." (Whorf 1940, 247). He also recognizes that the dualism and the way we analyse nature is not inherent to nature but to our approach to nature: "We dissect nature along lines laid down by our native language. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds – and this means largely by linguistic systems in our minds. We cut nature up, organize it into concepts and ascribe it significance as we do" (Whorf 1940, 231).

Furthermore, the language form "$P(a)$" is not a passive description of the associated binary experiment: it is a part of the experiment. Although names for objects and symbols for predicates can be arbitrary, their presence in our rational cognitive processes is essential. Through names, we control our connection with objects and through predicates we control our connection with associated experimental frameworks. Moreover, as I will explain below, objects and predicates do not exist by themselves – they also exist as parts of our rational syntheses with nature. Since names and predicates are a means of extracting objects and binary experimental frameworks

in rational cognition, each name is a part of the object it names and each predicate symbol is a part of an associated experimental framework. Thereby, a particular syntactic form is not important. What is important is the very presence of the form.

To my knowledge, von Humboldt is the first to recognize the importance of the above-described connection between language forms and the formation of concepts, and who finds in this relation the key to understanding why language is essential for thinking: "Language is the formative organ of thought. Intellectual activity, entirely mental, entirely internal, and to some extent passing without trace, becomes, through sound, externalized in speech and perceptible to the senses. Thought and language are therefore one and inseparable from each other. But the former is also intrinsically bound to the necessity of entering into a union with the verbal sound; thought cannot otherwise achieve clarity, nor the idea become a concept. The inseparable bonding of thought, vocal apparatus and hearing to language is unalterably rooted in the original constitution of human nature, which cannot be further explained […] without this transformation, occurring constantly with the help of language even in silence, into an objectivity that returns to the subject, the act of concept formation, and with it all true thinking, is impossible." (Humboldt 1836, 50). Umberto Eco says this poetically in the last sentence of the 1980 novel *The Name of the Rose*: "Stat rosa pristina nomine, nomina nuda tenemus."[2]

A fundamental semantic assumption of the use of an everyday atomic sentence "$P(a)$" in rational cognition is that "a" names an object. This rests on the assumption that it is possible to extract from the world something to be named. I will term the named object *the semantic value of the name*. Every name has the same general meaning – to name something. I will say that I *possess the full meaning of a name* if I have means to identify the named with the help of nature. These means can be different, even for the same object. They can be based on the senses. For example, I can identify my dog Svrco by sight, but also by hearing. They can be based on some physical equipment. For example, a star invisible to the naked eye can be identified using a telescope. They can be based on social contact. For example, I cannot directly identify person X but I know person Y who can

[2] "Yesterday's rose stands only in name, we hold only empty names."

identify person *X*. They can be based on some knowledge. Let's take the famous example of the planet Venus. If I know that *Venus = the morning star = the evening star*, then I can identify Venus in various situations in various ways. Sometimes a whole theory can help us identify an object. For example, Newton's theory of gravitation allows us to calculate the coordinates of the planet Venus in the sky at any moment and thus identify it. What are the meanings of the various names for Venus, and whether knowing that all these names refer to the same object changes their meanings, are questions I will not go into.[3] As I stated in the introductory section, I will deal only with the effects of meaning in the use of language. And this is exactly what the concept of possessing the full meaning of a name encompasses. Moreover, for the purpose of analysing the concept of truth, my aim in the next few paragraphs is to present the arguments only for the following two claims about names:

(i) Like the truth value of an atomic sentence, the process of naming is also a kind of synthesis of us and nature.

(ii) When we use language, we assume that every name of the language names an object, no matter how this connection is achieved and whether it is achieved at all. Expressed in terms of meaning: when we use language, we possess general meaning of names (that a name names something), but not necessarily the full meanings of names.

In doing so, I will not deal with defined names, but only with primitive names of the language, because the definition of a name ultimately reduces the possession of its full meaning to the possession of the full meanings of primitive names and primitive predicate symbols.[4]

When looking at my dog, I realize the connection between the word "Svrco" and my dog almost with a pure perception. However, in the moments when I cannot see him, I keep the connection on the basis of some definite knowledge and the theory that my dog exists somewhere as a distinct object. In everyday life, we keep the connection between the name and

[3] An overview of the various approaches to the meanings of names can be found, for example, in (Cumming 2023).

[4] Predicates are analysed below.

the named across time in such a way that, using some commonly established knowledge, we trace the named object, and any changes made upon it until the moment when we decide that it is no longer the same object (because it is destroyed or it is transformed into something else). When this connection terminates depends on an accepted world view. For example, when Svrco dies, whether the name "Svrco" denote his bones or his spirit, or neither, depends on a world view. I like to call this "the problem of Trigger's broom." Trigger is a likeable street sweeper in a British TV Series "Only Fools and Horses." He has got a medal from local authorities because of his thriftiness – he has been using the same broom for the last twenty years. However, we soon learn that in those twenty years he has replaced the broom head 17 times and the broom handle 14 times. Is it the same broom despite the changes? In everyday situations the decision is a matter of an (established) convention, more or less.

Other obscurities emerge when we analyse the connection between names and objects we cannot perceive directly. Here, the connection is more complex and more dependent on a theory. When we investigate in an experiment if a particle x was an electron, how do we know (i): that there is a distinguished object we can investigate, (ii): that the connection between name "x" and the object is preserved during the experiment, and (iii) that another object didn't appear, or the named object of the investigation hasn't changed?

Even if we ignore changes over time, the connection between name and the named is a complex mechanism of our interaction with nature. To begin with, I would use the game of recognizing figures in the clouds. Not only does the recognition of a figure in the clouds depend on the place of observation, but two people in the same place will see different figures. In ordinary situations, we all recognize and name the same beings and objects, so it seems to us that we are only giving names to existing objects. But as soon as we move away from the usual situations, extracting from the situation what will be our object (the named) becomes increasingly dependent on our approach. For example, in fluid dynamics, we distinguish between two approaches to the study of fluids, depending on what we have extracted for study – whether our object is a fluid that occupies a certain space and is constantly changing in time (Euler's approach) or always the same piece

of fluid that is constantly changing space in to which it is located (Lagrange's approach). A step further in the analysis would require us to "dive" into the fluid and turn into, for example, a jellyfish, while retaining the same linguistic abilities. Due to different needs and perception, the world would look completely different to us: the naming abilities would be completely different, and we would extract completely different parts of reality for the named objects.[5]

I believe these considerations are compelling enough to accept the first assertion about names: that, like the truth value of the atomic sentence "$P(a)$," *the process of naming is also a kind of synthesis of us and nature.*

When I use the name "Svrco," I exactly know what is named: my dog Svrco. However, even in everyday situations, we use names for which we don't know the exact object they name, for example, the name of a person we don't know. Even worse, it is possible that such a person does not exist, as it the case today with fake profiles on the internet. In the same unwarranted way, we extend the language used in everyday situations to other situations, when we are involved in science and mathematics, or when we talk fairy tales to children. However, we think "with names" in the same way, whether we know what they name or not and whether they name anything at all. For example, when we are involved in the fairy tale *Snow White and the Seven Dwarfs* we think, discuss and make conclusions as if all the characters in the story exist, because we are "tuned" to think in this way in semantically clear everyday situations. Only, when we step out of the language of the story (and use another language) we acknowledge that there are no such objects. This consideration supports the second assertion about names: concerning names, the moral is that *when we use language, we assume that every name of the language names an object, no matter how this connection is achieved and whether it is achieved at all.* Expressed in terms of meaning: when we use language, we possess general meaning of names (that a name names something), but not necessarily the full meanings of names. In the same way that we use language in everyday situations,

[5] In (Atiyah 1995), the famous mathematician Michael Atiyah described a thought experiment with an intelligent jellyfish, in which he showed that its mathematics would be significantly different from ours, thus arguing that mathematics is human invention, not discovery.

we use it in all situations. We can refine the language, replace it with a precise mathematical model, for example the language of first-order logic, but the assumptions of its use remain the same.

I believe that naming, as a kind of synthesis of us and nature, together with the fundamental assumption of the language use of names, that every name names an object (although we may not even know which object it names and whether it names anything at all), is a key primitive element of language. I think it is wrong to minimize the importance of naming as in Russell's theory of descriptions (Russell 1905), in Quine's reduction to values of variables (Quine 1948) or more radically in Quine's reduction to "ideal nodes at the foci of interesting observation sentences" in his naturalized epistemology (Quine 1990).

The next fundamental semantic assumption of the use of an everyday atomic sentence $P(a)$ in rational cognition is that the predicate "P" applied to the object a gives, with the help of nature, the truth value of the corresponding atomic sentence $P(a)$. The application of "P" consists of finding the associated binary experimental framework which, applied to a, gives an experiment in which nature gives the result: *True* or *False*. Thus, by *predicate* I consider the predicate symbol (predicate expression) together with this interpretation. Each predicate determines, through the intervention of nature, a mathematical function (function in the mathematical extensional sense) from objects to truth values. I will call this extensional function the *semantic value of the predicate*. However, we must not equate the predicate and its semantic value. Otherwise, we would destroy the whole language mechanism of rational cognition. The predicate is a part of the process of rational cognition, while its semantic value on a given object is the final result of this process, in which nature is substantially involved. Each predicate has its own general meaning: to generate a binary experimental framework. But each predicate has its own binary experimental frameworks. I will say that I *possess the full meaning of a predicate* "P" if I have means that for each object a I associate with the predicate a binary experimental framework in which nature will determine the truth value of the sentence $P(a)$. For a given predicate "P" we can have several different means, in the same way as with naming, from perception and experimental apparatus to the theories in which that predicate is incorporated. Using these means, we

can generate multiple experimental frameworks associated with the predicate. I will illustrate it using the example of the predicate "is a dog."

From the moment of birth, we form the possession of the full meaning of the predicate "is a dog," I would say almost by perception, as a part of our ability to differentiate beings. The semantic means of the predicate expression "is a dog" are deeply rooted in our sensory world, and only later do we complete it (make it more precise) with determinations which vary from everyday experience (for example that a dog does not necessarily have fur) to advanced theoretical knowledge (for example about its genetic code). This development does not mean that we did not possess the full meaning of that predicate before, but that the predicate itself changes, although its language form remains the same. The predicate "is a dog" leads to a whole host of binary experimental frameworks, from visual and auditory perception to the examination of the genetic code.

The possession of the full meaning of a predicate is manifested in our ability to apply it to objects by various means. Unlike the insufficiently clarified concept of meaning, the concept of the possession of the full meaning is verifiable to us and that is why I will use it.[6] For the purpose of analysing the concept of truth, my aim in the next few paragraphs is, similar to the analysis carried out for names, to present the arguments only for the following two claims about predicates:

(i) Like the truth value of an atomic sentence and like the naming, predicates are also a kind of synthesis of us and nature.

(ii) When we use language, we assume that we possess the full meaning of every predicate of the language without considering how we possess the full meaning and whether we possess it at all.

In doing so, I will not deal with defined predicates, but only with primitive predicates of a language, because the definition of a predicate ultimately reduces the possession of its full meaning to the possession of the full meanings of primitive names and primitive predicate symbols.

[6] The question of the meaning of predicates is one of the most difficult philosophical questions. An overview can be found, for example, in (Margolis and Laurence 2023) and (Orilia and Paoletti 2022).

From the fact that predicates are essentially connected to binary experimental frameworks, which are our biological or conceptual design integrated with nature, it follows that predicates are also a kind of synthesis of us and nature. The process of seeing leads to such a synthesis: light comes from the world but light processing belongs to our perception and brain. That is why predicates for colours are a typical example of the biological synthesis of us and nature.

For some predicates, it is clear that they were designed according to our intentions. Even in common situations, different people use different predicates. Predicates are the basic means by which we abstract what is important to us from a given situation. Let's imagine a group of hikers who have decided to have lunch. They have found a stone with a flat upper surface which is adequate to put out food and consume it. For them, the stone is a table. It is the same stone on which a ranger stood yesterday because he had a good view from it. For hikers, the stone is a table, for the ranger it is an observation post. Each of them extracted what they needed from the stone using the appropriate predicate. Even when I described that object as a stone, I have abstracted something from it by the predicate expression "is a stone." Even when I described it as an object, I have abstracted something from it by the predicate expression "is an object." All the above abstractions are conditioned by our preferences, but they are abstractions over nature. They also testify to the synthesis of us and nature in the formation of predicates. Further relativization would lead us to thought experiments in which we would analyse what kind of predicates other organisms (elephants or microbes) would develop in the same situation if they had our linguistic abilities. By means of their predicate expressions, they would surely create different abstractions and structure the situation differently. Thus, predicates depend on us as individuals, but also on us as a human community.

Those predicates with which we try to say something objectively about nature are especially important for science. However, in order to possess the full meaning of such a predicate, our presence is necessary. We usually achieve this through complex measuring devices that are a kind of extension of our senses. Thus, objectivity means not that such a predicate belongs to nature itself, but that it is invariant to the individual or group that applies

it. To all of them, nature will give the same answer in the application of such a predicate. However, there is another important problem that I will illustrate with the example of a predicate "is an electron." This predicate is applied to objects out of our direct experience. We must develop adequate experimental tools, built on some theory (world view), to have an indirect experience of such objects. Dealing with the meaning of the term "is an electron" opens up a lot of questions. Does one type of experimental framework determine the meaning of the predicate expression "is an electron"? Or is the essence of "is an electron" something else which only coincides with the concrete meaning in the context of the experiment? We would like that "is an electron" have a deeper meaning than it manifests in particular experimental settings. However, is such a "transcendental" predicate independent of various experimental settings or is it just their "common denominator"? In other words, does the predicate attached to the expression "is an electron" exist independently of us or does it exist only through our cognitive interaction with nature? A simple picture is that all such predicates exist independently of us, and that we only discover them through our interaction with nature. However, we have no rational ground for this claim. On the other hand, if we were to bound ourselves to predicates that strictly correspond to experimental settings we would lose any power of deeper cognition of nature. However, for the predicate "is an electron" to have any cognitive value, it must necessarily be part of our cognitive interaction with nature, otherwise it loses meaning. This problem also occurs in our everyday rational cognition. Moreover, the everyday situation clearly shows us the solution. I will take the already discussed predicate "is a dog" as an example. I can determine that a being is a dog with several types of experiments. One experimental framework is based on seeing that being, another on listening to that being, the third on analysing its genetic code. However, I have the knowledge that all these experiments on the same being will give the same answer. This knowledge allows me to possess the full meaning of the predicate "is a dog" over any of these experimental frameworks and invariant to them, because they all give the same answer. If another experimental framework appears tomorrow that gives the same answers as these, I will include it too in the possession of the full meaning of the predicate "is a dog." The same solution applies to the predicate "is an

electron." The theory of electromagnetism (if we stay within the framework of classical physics) allows us to associate with the predicate "is an electron" many binary experimental frameworks for determining whether a particle is an electron or not. We don't have to prefer any of these experiments because they all give the same answers. As with the predicate "is a dog," we possess the full meaning of the predicate "is an electron" through these experimental frameworks and invariantly to them. It is important to note here that this possession is based on a scientific theory which is also largely our construction. Thus, we participate in the predicate "is an electron" not only through the design of the associated experimental frameworks, but also through the scientific theory to which it belongs.

I believe these considerations are compelling enough to accept the first assertion about predicates: that, like the truth value of an atomic sentence and like the naming, *predicates are also a kind of synthesis of us and nature*.

Already in everyday language we use predicates for which we do not possess the full meaning. On a personal level, this happens constantly while growing up. Let us imagine a situation where a child has heard of kangaroos. At the beginning, all she knows about them is that kangaroos carry young in a pouch on their stomachs. If the child understood this information as a distinguishing characteristic of a kangaroo, then she possesses the full meaning of the predicate "is a kangaroo." She can determine for each animal, by examining whether it has a pouch, whether that animal is a kangaroo or not. Of course, her predicate "is a kangaroo" is different from the predicate "is a kangaroo" established by the human community. Through further learning, the child will have to change the meaning of her predicate and adapt it to the one accepted by the human community. When the child learns additional information, that only female kangaroos have a pouch and that there are other animals that have a pouch, then she knows she doesn't possess the full meaning of the predicate "is a kangaroo," and her further learning of that predicate will consist of completing the predicate. One photo of a kangaroo will allow her to possess the full meaning of the predicate, which is in accordance with the socially established understanding of the predicate. However, the child may see some beings that she is not sure are kangaroos. This means that she still does not possess the full meaning

of the predicate "is a kangaroo," but will have to complete it.[7] If one day she becomes a biologist specializing in kangaroos, only then will she possess the full meaning of the predicate "is a kangaroo." But is it really so? What if a new species is discovered and her knowledge is not enough to determine whether it is a kangaroo or not? Given that she can no longer compare her understanding of the predicate "is a kangaroo" with the understanding of the scientific community, because it has been agreed upon, this situation definitely leads to the conclusion that the scientific community does not possess the full meaning of this predicate but must complete it.

We could carry out a similar analysis for other predicates. In (Waismann 1968), Friedrich Waismann showed that we can almost never be completely sure that we possess the full meaning of a predicate. If I use his terminology, predicates have an "open texture." However, unlike predicates that do not belong to rational cognition (for example, the predicate "is a fairy"), predicates that belong to rational cognition usually develop over time towards greater precision and efficiency. Let us just take the predicate "is an electron" as an example. This predicate not only developed historically but also changed significantly with each more advanced physical theory. It has a different meaning in classical electromagnetism than in quantum mechanics or quantum field theory. Frank Wilczek, winner of the Nobel Prize in Physics writes: "What is an electron? That question was central to the development of quantum theory early in the twentieth century and remains at the frontier of physics today. There are several inconsistent answers, each correct." (Wilczek 2013). This is a good example of a predicate that essentially depends on the entire theory of which it is a part. That within various theories the concepts of electron are mutually inconsistent, yet correct, is not contradictory and can easily be explained by the connection between scientific theories about nature and nature. Scientific theories are only models of nature that approximate it well enough within a certain scope of the phenomenon. Thus, the predicates of the theory are also just approximations that we try to fix within the theory or change them significantly by

[7] Another possibility is to decide to reject everything that she does not recognize as a kangaroo as not being a kangaroo. But that would sooner or later lead to collisions between her understanding of that predicate and the scientific understanding.

changing the theory. That is why, as a rule, we never possess the full meaning of a predicate, but a sufficiently full meaning for the needs of knowing a phenomenon. Possessing a full meaning is an idealization, similar for example to the idealization of point particles (material points) in classical mechanics.

This brings us to the second assertion about predicate symbols: concerning predicates, the moral is that *when we use language, we assume that we possess the full meaning of every predicate of the language without considering how we possess the full meaning and whether we possess it at all.*

Due to the further analysis of different aspects of the concept of truth, it should also be pointed out that there are situations where we do not use predicates as an investigative tool to address questions to nature. Commonly, these are situations which we create and over which we have control, for example, in designing a game, a story or a mathematical world (as I will explain later in Section 6). Then, for some predicates, we directly decide on which objects they give *True,* and on which objects they give *False.* For example, we can decide which character in a fairy tale will be good or which natural numbers less than one hundred will have some (unimportant) property U (we will just enumerate such numbers). This is another use of predicates in which we directly reduce them to their semantic values. The role of these predicates in our rational activities is quite different than the original role of predicates as investigative means in rational cognition.

As I have analysed one-place predicate symbols, I can also analyse multiplace predicate symbols. The analysis of function symbols is similar to the analysis of names. I will say that I *possess the full meaning of a function "f"* if I have means to identify the named $f(a)$ with the help of nature, assuming that I possess the full meaning of the name "a." A nice example of these functions are measurement functions, such as mass or temperature, which associate numbers with parts of nature through an appropriate measurement process.[8] The *semantic value of the function* is the corresponding mathematical (extensional) function between objects. As with predicates, so with functions, it is essential to distinguish the function from its semantic value. A function is a part of the process of rational cognition, while its

[8] These functions are analysed in (Čulina 2022).

semantic value on a given object is the final result of this process, in which nature is substantially involved.

To conclude, the essence of the synthetic concept of truth is the following one. By dis-joining the world and our actions in it into objects and predicates, which we control through language symbols, we put binary questions to nature. By selecting one of the two offered answers, nature brings its contribution to the framework, besides its contribution to the processes of naming and of predicating. In a binary experiment of applying predicate "P" to object a, when nature selects an answer, *True* or *False*, it "says" something about itself. With this valuation of the language form "$P(a)$," the form which describes and controls the experiment, we gain knowledge about nature. This is the starting point for the overall role of the concept of truth in our rational cognition.

It should be noted once again that this is an idealized situation. Often in real situations we do not know exactly what a name names and whether it names something at all, as well as how to apply a predicate to a given object. However, when we use the language to which these names and predicates belong, an integral part of its use is that we assume that these names name objects and that we know how to apply predicates to objects. This is how we use everyday language, and we extend such use to the total language (languages) of rational cognition. Only when we take the names and predicates of that language as objects of our thinking, only then do we deal with the problem of the fulfilment of the assumptions of their use. Then we use another language (the metalanguage of the given language). Then these names and predicates are not means of our (object) language (where we use them) but are objects of another language (where we mention them). In Section 5, dedicated to the scientific aspect of the concept of truth, the use of object language in science will be considered. In Section 7, dedicated to the assertion-valuation distinction, the use of metalanguage in the examination of object language will be considered.

I consider that the synthetic concept of truth is the solution to the philosophical problem of truth – is there any connection between truth and reality and, if so, what is the connection. The synthetic concept of truth shows that there is a connection and precisely shows what the connection is.

3. Comparison of the Synthetic Concept of Truth with Other Concepts of Truth

Clearly, the synthetic concept of truth is not any kind of a deflationary concept of truth that diminishes the importance of the concept of truth.[9] The synthetic concept of truth is of crucial importance for rational cognition. Also, the synthetic concept of truth is not a kind of a correspondence theory of truth where the truth value of the sentence is determined only by whether the sentence corresponds with reality or not. Thereby, reality is considered something independent of us and language: language only serves to describe reality.[10] In the synthetic concept of truth, atomic sentences themselves, with their interpreted parts – names and predicate symbols – and with their truth values, where nature is involved, form reality: reality is the result of the synthesis of us and nature through the creation and use of language.

Although formal parallels can be drawn between Frege's analysis (Frege 1891, Frege 1892a, Frege 1892b, Frege 1892c) and my analysis of the atomic sentence, especially between Frege's insistence on the distinction between the concept and the extension of the concept and my insistence on the distinction between the predicate and the semantic value of the predicate, the results of the analysis are fundamentally different. For Frege, sense and reference have a metaphysical meaning: an atomic sentence expresses a thought, and the thought belongs to a kind of Platonic world (Frege 1918), as well as the truth value of the thought. In Frege, the thought is primary, and it can be decomposed into object and the remaining unsaturated part – concept (Frege 1906a, Frege 1906b). Thus, the object and the concept as part of the thought also belong to the Platonic world. In my analysis, predicates and objects are primary. They belong to the world of our real activities: from predicates and objects we build binary experiments in which, with the help of nature, they are synthesized into the truth values of the

[9] Various formulations of the deflationary concept of truth can be found, for example, in (Armour-Garb, Stoljar and Woodbridge 2023).

[10] Various formulations of the correspondence conception of truth can be found, for example, in (David 2022).

corresponding sentences. Language is not a lifeless description of the Platonic world, but it is a living organism that changes and completes itself in the synthesis with nature.

The ideal situation in which a predicate possesses a full meaning is a kind of formulation of the verification principle, an idea that is at the basis of logical empiricism. In the words of Moritz Schlick: "The meaning of a proposition is the method of its verification" (Schlick 1936). In the words of Rudolph Carnap: "Thus the meaning of a sentence is in a certain sense identical with the way we determine its truth or falsehood; and a sentence has meaning only if such a determination is possible." (Carnap 1936). The only elaborated version of this principle, which at the same time diminishes its value, is Carnap's version of logical empiricism (Carnap 1936, Psillos 2000, Carnap 1966), so I will stick to it. At the level of atomic sentences, Carnap's analysis of the idea of verification leads to the division of predicates into observational predicates (e.g. "is red") and theoretical predicates (e.g. "is an electron"). Although a clear boundary cannot be drawn, we can roughly say that observational predicates have a high degree of verifiability. On the other hand, theoretical predicates are not directly verifiable, and that is why Carnap requires the introduction of correspondence rules that will connect them with observational predicates. These rules will not make them verifiable but will give them a certain indirect empirical meaning. Carnap develops the entire structure of such a language in which he tries to give each sentence, not only atomic sentences, some degree of empirical meaning.[11] If we stay at the level of atomic sentences, unlike Carnap's division into observational and theoretical predicates, which is quite questionable and heavily criticized by Quine (Quine 1951), my approach is uniform. All primitive predicates have the same status in the idealized situation of possessing their full meaning, and in the assumption of this possession we adhere to when we use language. How much we really possess the meaning of a predicate, i.e. to what extent the assumption of language use is fulfilled on it, are questions that belong to metalanguage and which, in my opinion, due to the fluency of the concept of meaning (its sensitivity to a multitude

[11] I will comment on that language in Section 5, dedicated to the scientific aspect of the concept of truth.

of factors that include even the accepted theory on which we rely on) cannot be adequately formulated in the language itself, as Carnap tried.

On the level of ideas, in addition to a different understanding of language, my move away from logical empiricism is a move towards the ideas of pragmatism: placing the human being at the centre of rational cognition, as an active biological and sociological being whose needs and motives significantly shape their rational cognition. In William James memorable words: "In our cognitive as well as in our active life we are creative. We add, both to the subject and to the predicate part of reality. The world stands really malleable, waiting to receive its final touches at our hands. Like the kingdom of heaven, it suffers human violence willingly. Man engenders truths upon it." (James 1907, 254). Not various dualisms but a synthesis: "Does the river make its banks, or do the banks make the river? Does a man walk with his right leg or with his left leg more essentially? Just as impossible may it be to separate the real from the human factors in the growth of our cognitive experience." (James 1907, 250). However, due to the wide variety of pragmatism and its remaining at the programmatic level, it is difficult for me to draw some more specific connections with my approach.

The founders of pragmatism based their insistence on the integrity and uniqueness of the human being in the cognition of the world on the theory of evolution. Today, in their work, M. R. Bennett and P. M. S. Hacker base it on cognitive neuroscience: "A human being is a psychophysical unity, an animal that can perceive, act intentionally, reason and feel emotions, a language-using animal that is not merely conscious, but also self-conscious – not a brain embedded in the skull of a body....it is human beings who think and reason, not their brains." (Bennett and Hacker 2022, 3). In their view of language and concepts, I find similarities with my approach. These authors also give a key importance to language in human thinking and cognition: "...it is the capacity to speak and the mastery of a language that is a condition of all that is distinctively human, and hence too a condition for the sciences and the arts of humanity" (Bennett and Hacker 2022, 13). The connection between form and meaning is very similar to mine. For me, a predicate is a predicate symbol together with its meaning, where the concrete form is not important, but its very presence is important. They have a similar relation between words and concepts: "A concept is an abstraction

from the use of a word" (Bennett and Hacker 2022, 404). I am talking about the possession of the predicate; they are talking about the possession of the concept. However, for them, possessing a concept means knowing how to use the concept-word in the "language game": "To have a concept is to have mastered the use of a word" (Bennett and Hacker 2022, 404). For them, to master a concept means to master "the rule-governed use of the word that expresses the concept" (Bennett and Hacker 2022, 404). And this is achieved through "emulation: learning by doing; parental and sibling training and teaching; learning by engaging in language games; learning from informal instruction" (Bennett and Hacker 2022, 403). In my opinion, for a predicate that is part of rational cognition, the acquisition of its application to the world is of primary importance and not the acquisition of the use of its expression in language, although these processes are connected. Here my move away from these authors is a move towards the ideas of logical empiricism.

Davidson (2005, 77) points out the key problem of "unity of proposition" that the theory of truth and predication must solve:

> …if we do not understand predication, we do not understand how any sentence works, nor can we account for the structure of the simplest thought that is expressible in language. At one time there was much discussion of what was called the "unity of proposition"; it is just this unity that a theory of predication must explain. The philosophy of language lacks its most important chapter without such a theory, the philosophy of mind is missing its crucial first step if it cannot describe the nature of judgement; and it is woeful if metaphysics cannot say how a substance is related to its attributes.

For Davidson, the concept of truth is a primitive concept, as it is for me. But to me it is more than that. The truth value of an atomic sentence, as the result of our synthesis with nature in the process of rational cognition, gives unity to the atomic sentence that Davidson seeks: it makes the atomic sentence to be something more than just the concatenation of its parts, the predicate symbol and the name involved in the sentence.

4. The Logical Aspect of Truth

We can build various language structures over atomic sentences. The object-predicate dualism naturally leads to first order languages, which not only have a simpler and clearer semantics than other languages but also prove to be the most important type of logical language. In what follows, I will assume this type of language.

The basic building blocks of a first-order language are atomic sentences which are analysed above, and which are the primary bearers of the synthetic concept of truth. Consequently, all the assumptions of the use of atomic sentences are now the assumptions of the use of an interpreted first-order language. These are all the assumptions mentioned above that we accept when we use (not when we mention) these atomic sentences: that each name names an object, that for each predicate we possess its full meaning, and consequently and with the help of nature, that each atomic sentence is true or false.

Each complex sentence of an interpreted first order language is a description of a binary experiment which is a combination of binary experiments associated with atomic sentences. For example, the sentence $P(a)$ *and* $Q(b)$ describes a binary experiment composed of the binary experiments described by the sentences $P(a)$ and $Q(b)$. The associated binary experiment applied to a and b yields *True* when both atomic experiments yield *True*, otherwise it yields *False*. Likewise, the sentence *for all* x $P(x)$ describes an experiment that gives the value *True* when for each valuation of the variable x the experiment described by $P(x)$ gives the value *True*, while otherwise it gives the value *False*. Why do we need these combinations at all, given that there is nothing new in them concerning rational cognition which is not present in atomic sentences? There are several reasons but by far the most important reason to combine binary experiments is to recognize and determine a regularity that is repeated in certain types of combinations. For example, every time when we assert that an object is a dog, we, or somebody else, sooner or later, will also assert that the object is mortal. We combine the experiments "x is a dog" and "x is mortal" into the experiment "if x is a dog then x is mortal," which gives the value *True* for each evaluation of x. We capture in a simple way the observed regularity

by claiming that the sentence "For all x, if x is a dog then x is mortal" is true. However, quantification poses the so-called problem of induction (Hume 1738 – 1740). We can determine the truth value of "if x is a dog then x is mortal" for any value of x (in principle) but we cannot do it for all (potentially infinite) values. This is a situation in which we can possibly get the answer "no" but never the answer "yes." We could conclude that this sentence does not describe a binary experiment at all, and we could exclude this type of sentences from language. However, then we could not express regularities which we observe and which are the main sources of knowledge, as the history of science confirms.[12] *As with naming and predicating, we extend the use of language in ordinary situations to all situations and assume that every sentence of an interpreted first-order language is true or false, regardless of the way we find its truth value, and even regardless of whether we can find it at all.* We accept such universal and existential sentences (and corresponding experiments) despite all uncertainty they bring. This assumption is of foremost importance for the scientific concept of truth, which will be described in the next section, but also for the logical concept of truth to which this section is dedicated. This assumption and all the assumptions of the use of atomic sentences I will term *the external assumptions of an interpreted first-order language*. Their fulfilment is crucial for the application of the language but not for the logic of the language. The only important thing for the logic of the language is that these assumptions are part of the specification of the language, not whether they are fulfilled. By *the logic of a language*, I mean the internal organization of the language – the connection of semantic values of language forms, which is independent of the reality that the language speaks about – together with the external assumptions of the language use.

For a first order language, a mathematical (extensional) function is connected with each language construction of a sentence from simpler sentences. The function determines the truth value of the constructed sentence on the basis of the truth values of the sentences from which it is constructed. For example, the construction of the conjunction *A and B* is connected with the two-place Boolean function that outputs *True* only when both inputs

[12] As C. D. Broad said: "induction is the glory of science and the scandal of philosophy" (Broad 1952, 143).

A and *B* are *True*. The important property of any such function is that it is an internal semantic function, a function that connects semantic values independently of the reality the language speaks of. So, it belongs to the logic of the language. I will term such a function *the semantic function of the construction*. These semantic functions give recursive conditions for truth values which, together with the truth values of atomic sentences, determine the unique mathematical function that assigns, in a given evaluation of variables, a truth value to each sentence. This means that in an interpreted first order language, under the external assumptions of its use, the truth value of each sentence is entirely determined by the truth values of atomic sentences. According to the synthetic concept of truth, the truth values of atomic sentences are primitive semantic elements of language determined by the process of rational cognition. In this way, with the assumptions of language use, the truth value of each sentence is connected with reality in a completely determined way.

Because the semantic functions of the sentence constructions in a first order language belong to the logic of the language, they determine the logical connection of truth values of the sentences. This aspect of truth, the internal interconnectedness of the truth values of sentences of a language, I will term *the logical aspect of the concept of truth*. Important concepts of logical truth and logical consequence belong to this aspect. *Logical truth* is the sentence whose truth is determined, under the external assumptions of language use, by the internal semantic structure of the language regardless of its particular connection with reality. E.g. the sentence *not A or A* is a logical truth, because its truth is determined by the internal semantics of the connectives *not* and *or*, regardless of the truth value of sentence *A*. Also, that from a set of sentences $\{A_1, A_2, \dots\}$ *logically follows* a sentence *B*, means that starting from the truth of the sentences $A_1, A_2, \dots$, the internal semantic structure of the language, not the reality the language speaks of, determines the truth of *B*. Thus, for example, the internal semantics of the connective *and* determines that a sentence *B* logically follows from the sentence *A and B*. The relationship of logical consequence between sentences is one of the crucial language mechanisms in the development of rational cognition.

The logical elements of first order languages are analysed in detail in (Čulina 2024).

5. The Scientific Aspect of Truth

As analysed above, the first order language built upon interpreted atomic sentences has the external assumptions of its use. These are: (i) the fundamental assumption of the language use of names: every name names an object, (ii) the fundamental assumption of the language use of functions: for each function symbol we possess its full meaning, (iii) the fundamental assumption of the language use of predicates: for each predicate symbol we possess its full meaning, and (iv) the fundamental assumption of the language use of sentences: every sentence is true or false. In a real process of rational cognition, already in everyday situations and especially in scientific theories, we use names for which we do not know completely what they name, predicate and function symbols for which we do not possess the full meaning, and quantified sentences for which we do not know if they are true or not. However, it is important to emphasize that regardless of whether the exterior assumptions are fulfilled or not, the logic of the language demands that when we use the language, we assume that they are fulfilled. In thinking itself there is no difference whether we think of objects that really exist, or we think of objects that do not really exist and whether the predicate symbols we use can be applied to such objects at all or not. That difference can be registered only in a "meeting" with reality.

Furthermore, although semantic values of the complex language forms are determined by semantic values of the simpler forms from which they are built, in the process of rational cognition we invert this original priority. An assertion about a particular object is more confident and more determined rational cognition then an assertion about all objects. However, we cannot apply all primitive (undefined) predicates to all objects because there are too many objects, potentially infinitely many. Furthermore, some objects disappear, some come into existence. So, we cannot know the truth values of all atomic sentences. We rely more and more on the regularities which we notice. These regularities are formed by universal and existential sentences (laws). These sentences gradually become the main basis for rational cognition, although we cannot perform completely the complex binary investigations they determine. Moreover, these sentences speak often about idealized situations and idealized objects using idealized predicates.

For example, in classical mechanics, we analyse a motion of the so-called material particles which at each moment of time occupy exactly one point in space. Hence, we assert something about objects which even do not exist in the strict sense of this word. We make assertions about such objects without any corresponding atomic sentence we could verify experimentally. Despite this, such assertions are the result of a deeper analysis of real situations and, through a kind of synthesis, give us powerful knowledge of real situations.

All this means that our real knowledge, regardless of the degree of its accuracy, is almost always only a fragment of some assumed ideal semantically complete language. The whole dynamics of a scientific theory can be understood as the dynamics of completing and changing an appropriate language. In the process of rational cognition, we decrease unspecified parts of the language, even change the meanings and the semantic values that had been already formed. However, this process is not chaotic, but it is, looking over longer periods, a constant advance in rational cognition of nature.[13] That is because it has powerful regulatory mechanisms which control and drive it – the exterior interaction with nature through experiments and the logic of language. Namely, for a theory to be a scientific one, at least some names and some function and predicate symbols must have an exterior interpretation, an interpretation in the exterior world, not necessarily a complete one. This partial external interpretation enables us to perform at least part of the binary experiments described by atomic sentences. This allows nature to put its answers into our framework, so that we can test our conceptions experimentally. Without this part the theory is unusable. On the other hand, the language disciplines us in a way that we shape our cognition and understanding into a set of sentences which we consider to be true. In an ideal case, we choose a not too big set of sentences we are pretty sure to be true, the axioms of the theory. Then, we are obligated, by the logic of the language, to consider true all sentences which logically follow from the axioms. So, another rationalized part of our conceptions consists of a set of sentences we consider to be true and to which we try to give an axiomatic organization.

[13] Even Kuhn's scientific revolutions (Kuhn 1962) can be interpreted as radical changes of established language frameworks.

Therefore, a scientific theory about nature is in its most explicit form a junction of a set of sentences (the sentence part of the theory) and partial external interpretation of the language (the interpreted part of the theory). From the axioms of the theory, we logically deduce the truth values of sentences. Particularly, we deduce the truth values of atomic sentences which belong to the external interpretation, and which are, therefore, experimentally verifiable. If the truth values do not coincide with the truth values which nature gives, then the theory is wrong. If they are identical, it makes the theory trustworthy but, as we know, it is not a proof that it is right. As Popper emphasizes, theories must be experimentally verifiable so that they can be falsifiable. In this interaction of the sentence part and the externally interpreted part of a theory, the real dynamics of the theory takes place: the axioms, as well as the interpreted parts, evolve, even change, and the same happens with the whole language framework. Science is the construction of the language which is not semantically complete in any phase of the construction.

I will term this aspect of the concept of truth *the scientific aspect of the concept of truth*. At the core of this scientific dynamics is the synthetic concept of truth. It gives legitimacy and perspective to scientific research described above as a development of truth valuations of sentences and external interpretation of a language.

This approach is fundamentally different from Carnap's approach. This difference is not only in the approach to primitive predicates, as commented at the end of Section 3, but also at the level of complex sentences. By dividing the language into empirical and theoretical sentences and connecting them using the correspondence rules, Carnap strives to obtain a language that is a semantically complete language for empirical sentences and, through the correspondence rules, complete in a way for theoretical sentences at every stage of development (Psillos 2000, Carnap 1966). In the approach developed here, science is the construction of the language which is not semantically complete in any phase of the construction. On the level of ideas, my move away from logical empiricism is again, as with the interpretation of atomic sentences, a move towards the ideas of pragmatism and the active role of the human being in rational cognition. A scientific theory is Neurath's ship, which we repair in parts, but which sinks as a whole.

6. The Mathematical Aspect of Truth

The concept of truth in mathematics essentially depends on the accepted philosophy of mathematics (Horsten 2023). Thus, the mathematical concept of truth presented here also depends on a certain philosophy of mathematics, which is elaborated in (Čulina 2020).

I consider mathematics primarily the internal organization of rational cognition, a thoughtful modelling of that part of the process of rational cognition that belongs to us. Building a logical language is one such modelling. So, I consider that logic is part of mathematics. A first order language is a mathematical model constructed for the use in rational cognition just like natural numbers are constructed for counting. It is the result of thoughtful modelling of intuition about our natural language. Thoughtful modelling of other intuitions about our internal world of activities, for example, intuitions about quantity, symmetry, flatness, nearness, etc., lead to other mathematical models. By "our internal world of activities" I mean the world that consists of activities over which we have strong control and which organize and design by our human measure (e.g., movements in space, grouping and arranging small objects, writing on paper, painting, playing music, …).

It is from these concrete activities that the idea of an idealized mathematical world emerges, the world that expands and supplements the internal world of activities. Let's take real numbers, for example. Although we can approximate irrational numbers by rational numbers with arbitrary precision (if we had enough space, time and materials – again an idealization), their existence is outside our means of construction – we have just imagined irrational numbers.[14] By choosing names, function symbols and predicate symbols, we shape the initial intuition into one structured conception. However, here the role of functional and predicate symbols, as well as the truth values of sentences, is different than in rational cognition. Predicates are not investigative tool to address questions to nature, there is no intervention of nature, and thus no synthesizing role of truth values. Truths are truths "by fiat." Because we create a mathematical world we

[14] In his book (Mac Lane 1986), Sounders Mac Lane describes this process of idealization on a multitude of examples.

have a complete control in its design. We determine on which objects the predicate will give truth, in the same way as we decide which character in a fairy tale will be good. It's the same with functions. We cannot experimentally verify that $|| + || = ||||$ $(2 + 2 = 4)$ because it is not the truth about nature – it is the way we add tallies.

However, since the conception usually goes beyond our constructive capabilities, the constructed language has only partial interpretation in our internal world of activities. Since the interpretation is only partial, and because the imagined domain of interpretation is usually infinite, we cannot determine the truth values of all sentences of the language. Therefore, we must further specify the conception by appropriate choice of axioms. When we describe a mathematical world by some set of axioms, inferring logical consequences from the axioms, we establish what is true in that world. This can be very creative and exciting work, and it seems that we discover truths about some existing exotic world, but we only unfold the specification. The inferred sentences are not true because the world they describe is such, but that world is so conceived that those sentences are true in it. They are the conditions that the world must satisfy. I will term this aspect of the concept of truth, as a specification of an imagined mathematical world that emerged from our internal activities, *the mathematical aspect of the concept of truth*. Since I consider logic to be part of mathematics, the logical aspect of the concept of truth is also part of the mathematical aspect of the concept of truth. I would note that we have already encountered this mathematical aspect in logic on the example of a linguistic construction using the connective *and*. This connective is directly associated with its semantic value, the corresponding Boolean function, without an intensional intermediate step.

Although, in contrast to synthetic truths, mathematical truths are completely determined by us, the very fact that mathematics is our tool of cognition results in the subordination of the mathematical concept of truth to the synthetic concept of truth. We can understand mathematical truths as extreme examples of synthetic truths, in which nature does not participate at all, but everything is subordinated to our actions. An important consequence of this is that mathematical and cognitive language have the same semantic and logical structure, a structure that has its source in the synthetic concept of truth. In both languages, we think in the same way,

and that thinking is based on the same assumptions of language use. This uniformity enables the double connection of mathematics as a tool of cognition with cognition itself. I will illustrate the first way using the example of Euclidean geometry. In my view, Euclidean geometry is an idealized conception derived from intuition about our internal spatial activities (Čulina, 2018). However, we can preserve the sentence part of the theory but change the interpretation. If we ask ourselves whether the physical space obeys the axioms of Euclidean geometry, we must extract from space what we consider as points (maybe enough localized parts of space), as directions (maybe directions of light rays), and the distance between two points (maybe the time needed for light to pass from one point to another). If in such an interpretation the physical space satisfies the axioms of Euclidean geometry then we have an experimentally verifiable theory. Its sentence part is the same as in our mathematical theory of the space of our human activities, so we can transfer all results to the structure of physical space. Only the interpreted part is different. It does not belong to mathematics anymore, but it is a base for an experimental verification of the theory about the external world. Thus, thanks to the uniformity described above, the mathematical concept of truth can be understood as a matrix for the synthetic concept of truth: by changing the interpretation, we directly turn a mathematical statement into a synthetic statement. Considering that through various interpretations the same mathematical statement can generate various synthetic statements, with this mechanism we achieve great efficiency in thinking.

The uniformity of mathematical and cognitive language described above is the basis for an even more essential connection between these two languages in the process of cognition: when mathematical language is literally part of cognitive language. Perhaps the best example for this is quantum mechanics, where we associate a Hilbert space with a physical system: we associate the cognitive language of physical systems with the mathematical language of Hilbert spaces. In that common language, mathematical terms are organically combined with physical terms in the formulation of claims. However, we can already illustrate this organic connection with the well-known simple use of natural numbers. Through the process of counting, we connect nature with the world of natural numbers. For example, the true

statement about the world that there are now 3 objects on my table includes an imaginary mathematical object, the number 3. When I add a coffee cup to the table, the true statement about the world, that there will then be 4 objects on my table, in addition to some assumptions about the world (for example, that there will be no explosion), follows from the mathematical statement that $3 + 1 = 4$.

Despite its uniqueness, both in its structure and in its purpose, the mathematical concept of truth arises from the synthetic concept of truth.

7. The Role of the Truth Predicate in the Concept of Truth

All considerations about the concept of truth in this article have been done in the appropriate metalanguage whose sentences also have their truth values. The basic connection between the truth of the sentences of the language L we are considering and the language ML in which we are considering L is achieved through the truth predicate "T." The truth value of the sentence φ in the language L corresponds to the truth value of the sentence $T(\ulcorner\varphi\urcorner)$ in the language ML. It is precisely Tarski's T-schema that expresses this role of the truth predicate:

$$T(\ulcorner\varphi\urcorner) \leftrightarrow \varphi$$

This aspect of the T-scheme underlies the basic idea of deflationism that the truth predicate is unnecessary: asserting $T(\ulcorner\varphi\urcorner)$ is the same as asserting φ. However, in the context of language and thinking, the left and right sides of the biconditional differ significantly. Without loss of generality, I will show this difference on the example of the sentences "Svrco is afraid of thunder" and "'Svrco is afraid of thunder' is a true sentence."

The main difference in the use of these sentences is that when I say "Svrco is afraid of thunder," the subject of my expression and thought is my dog Svrco, and when I say "'Svrco is afraid of thunder' is a true sentence," the subject of my expression and thought is the sentence "Svrco is afraid of thunder." This is a typical use-mention distinction. In the first case I use the sentence "Svrco is afraid of thunder" to say something about Svrco and in the second I mention the sentence to say something about it. What is specific here is that one sentence speaks about the truth of another

sentence, where each of the sentences has its own truth value. If, for example, we were talking about the number of letters in the sentence "Svrco is afraid of thunder," nothing would be disputable. The very fact that one sentence speaks of the truth of the other leads to an important difference between these sentences in the process of thinking. I will term it *assertion-valuation distinction*. Namely, the very way we use a (declarative) sentence conveys the information that we consider it true. So, when I assert "Svrco is afraid of thunder," in addition to the information about Svrco, I convey the information that it is a true sentence. So, there is no need to assert it in a roundabout way with the sentence "'Svrco is afraid of thunder' is a true sentence" (by which I again convey the information that this sentence is true). However, if someone considers the truth of the sentence "Svrco is afraid of thunder," they will not use this sentence but will mention it and evaluate its truth. If they conclude that it is true, they will end their analysis with the assertion "'Svrco is afraid of thunder' is a true sentence."

This assertion-valuation distinction distinguishes the left and right sides of Tarski's T-scheme. It is also a mechanism for stopping or prolonging truth regression. For example, using sentences of the metalanguage *ML* I discussed the truth values of sentences of a language *L*. The insights I gained that way are composed of sentences which also have truth values, which may be the subject of other sentences. And so on indefinitely. The assertion aspect stops the regression, and the valuation aspect continues the regression. So, if we agree on something, that's where the regression ends. Usually, the regression stops in the metalanguage. If someone disputes what I have said about the truths of sentences of the language *L*, they dispute the truth of the corresponding *ML* metalanguage sentence. But the subject of their analysis will again be the language *L* and the conclusion they draw will be the assertion of the metalanguage *ML* and not its metalanguage *MML*.

As far as I know, the importance of the linguistic mechanism of assertion was first pointed out by Frege (1897). How subtle and important the concept of assertion is in Frege can be read in (Pedriali 2017). Contemporary considerations on the concept of assertion can be found in (Brown and Cappelan eds. 2011).

8. Tarski's Definition of Truth

As analysed in the introductory part of the article, Tarski's T-scheme is a classic example in which various aspects of the concept of truth are mixed. This extends to Tarski's definition of truth (Tarski 1933), too: some see the definition as an argument for the correspondence theory of truth, others for the deflationary theory of truth. A comprehensive analysis of Tarski's work and various critiques of the work can be found in (Patterson 2012). In this section, Tarski's T-scheme and Tarski's definition of truth are analysed in relation to the aspects of truth differentiated in this article, especially in relation to the synthetic concept of truth.

Regarding the analysis of the concept of truth, the assertion-valuation distinction shows that truth value occurs in two ways, implicitly as part of an assertion or explicitly through the truth predicate symbol, i.e. through mentioning the truth value of a sentence. To assert the sentence $T(\ulcorner \phi \urcorner)$ which explicitly says that the sentence ϕ of a language L is true is to assert the sentence ϕ, and vice versa. If we ignore the translation problems and consider that the metalanguage *ML* is an extension of the language L, this means that all T-sentences are true. We can assert that for every sentence ϕ of the language L:

$$T(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$$

The nature of the truth of these T-sentences can be viewed in various ways, depending on how we view the truth predicate symbol through which the truths of the left and right sides of the biconditional are equated, as I will show below. However, regardless of these differences, the truth of T-sentences belongs to the mathematical aspect of the concept of truth because their truth belongs to the internal organisation of rational cognition. If we were to use the more general T-scheme $T(\ulcorner \varphi \urcorner) \leftrightarrow \varphi^*$ related to a metalanguage that is not an extension of the language L, due to the question of correctness of translation, the scientific aspect of the concept of truth could be present, too.

It is common to consider T-sentences $T(\ulcorner \varphi \urcorner) \leftrightarrow \varphi^*$ as partial definitions of the truth predicate of a language L. In this case, T-sentences are analytical truths of the metalanguage *ML*. So, this is a logical aspect of the concept of truth. This view is directly related to Tarski's analysis of the concept

of truth. Tarski's definition of the truth predicate for the language L in the language *ML* (Tarski 1933) is a formally correct definition because it enables the elimination of the defined predicate symbol T in every sentence of the language *ML*. The definition is also a materially adequate definition in the sense that all T-sentences logically follow from it. However, Tarski's definition of truth has the role of a definition in the proper sense of that term only when we want to set the truth of the sentences of one yet uninterpreted language L by using the truth of the sentences of another language *ML*. This definition transfers the meaning, and thus the truth value of the sentence φ^* of *ML*, to the truth of the sentence φ of L via the appropriate T-sentence. That is why Tarski's definition is so important in mathematical logic.

However, for the interpreted language, Tarski's definition is not a definition in the proper sense of that term because it "defines" something that has already been determined. In such a context, Tarski's definition simply gives a *translation* from the language L to the language *ML* via the T-scheme: each sentence φ of the language L is translated into the sentence φ^* of the language *ML*. If the translation is correct, it preserves the meanings and thus the truth values of the sentences. In this situation, Tarski's definition is simply a mathematical construction of the translation function. It makes possible to connect the truths of sentences of two different languages.

But whether Tarski's definition is a substantive definition or just a mechanism of translation from one language to another, it only transfers the problem of the truth of a sentence of one language to the same problem of the truth of the corresponding sentence of another language. For example, using the T sentence T("*Svrco is afraid of thunder*") ↔ *Švrćo se boji grmljavine* from the introductory part of the article, instead of examining the truth of the statement "Svrco is afraid of thunder," we can now examine the truth of the statement "Švrćo se boji grmljavine," and vice versa. If the translation is correct, it is the same problem. This is best seen when the metalanguage *ML* is an extension of the language L, i.e. when we have a T-scheme $T(\ulcorner\varphi\urcorner) \leftrightarrow \varphi$. Then Tarski's definition translates the problem of the truth of the sentence "Svrco is afraid of thunder" to the problem of the truth of the sentence "Svrco is afraid of thunder" ($\varphi^* = \varphi$).

The problem with Tarski's definition of the concept of truth and the interpretation of his contribution to the analysis of the concept of truth is as follows. Tarski says: "We should like our definition to do justice to the intuitions which adhere to the classical Aristotelian conception of truth – intuitions which find their expression in the well-known words of Aristotle's metaphysics: 'To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true.'" (Tarski 1944, 342). However, Frege showed (Frege 1897) that it is not possible to give an absolute definition of truth, because the application of such a definition depends on the truth of definiens, so it is a circular definition. As a special case, Frege shows that a correspondence theory of truth is impossible because it reduces the problem "is a sentence true" to the problem "is it true that the sentence corresponds with reality," which again leads to circularity. Tarski's definition of the truth of a sentence is not an absolute definition of truth neither does it refine an intuition about truth as correspondence with reality. It is a relative definition of the truth of sentences in one language (object language) by the truth of sentences in another language (usually metalanguage). The definition enables a translation of the truth for sentences in one language into truth of sentences in another language, as Tarski explicitly states in his T-convention (Tarski 1933). Hence, in Tarski, the intuition about a correspondence theory of truth is realized as a correspondence of truth between two languages and not between language and reality. Tarski's recursive definition of truth reduces the truth values of compound sentences to atomic sentences. Tarski's and the synthetic conception of truth differ in the way they treat atomic sentences. Tarski finishes his definition by giving a translation of atomic sentences to metalanguage, and by this transferring the concept of truth from language to metalanguage. Contrary to this, in the synthetic conception of truth, the truth values of atomic sentences are undefined primitive elements determined by the process of rational cognition. In this way, the truth value of every sentence is connected with reality in a completely determined way. Tarski's definition of the concept of truth correctly formulates recursive conditions that connect the truth of a constructed sentence with the truth of the sentences from which it is constructed, while by translating the truth of atomic sentences of language L into the truth of

sentences of metalanguage, or vice versa, it ceases to be a content-wise theory of truth.

9. The Truth Predicate and the Paradoxes of Truth

In this section, my goal is to show that the paradoxes of truth are not paradoxes of the synthetic concept of truth but are limited to a certain part of the logical aspect of the concept of truth.

The basic purpose of the truth predicate "T" is that we can use it, in the corresponding metalanguage *ML*, to describe the truth values of the sentences of the language *L*. According to the meaning of the truth predicate "T," the sentence $T(\ulcorner\varphi\urcorner)$ is a true (false) sentence of *ML* when φ is a true (false) sentence of *L*. When the language *L* is not part of the language *ML*, the role of this predicate is the same as, for example, the predicate expression "is a diesel engine." Just as in the language of mechanical engineering we speak about engines using the predicate expression "is a diesel engine," so in *ML* we speak about the truth values of the sentences of *L* using the "T." "T" is a non-logical symbol of the language *ML*, just as "is a diesel engine" is a non-logical expression of the language of mechanical engineering. As "is a diesel engine" connects engine types with the truth values of the corresponding sentences of the language of mechanical engineering, so the truth predicate "T" connects the truth values of the sentences of the language *L* with the truth values of the corresponding sentences of the language *ML*.

However, when *L* is part of the language *ML*, then the truth predicate "T" connects the truth values of sentences of the same language. Truth conditions on the truth predicate "T," that $T(\ulcorner\varphi\urcorner)$ is a true (false) when φ is a true (false), where φ belongs to *L*, now belong to the internal semantics of the language in the same way as, for example, truth conditions on connectives. In this case, the truth predicate "T" is a logical symbol of the language *ML*, like connectives and quantifiers. The only difference in relation to connectives and quantifiers is in universality. Only a language that has its own sentences in the domain of its interpretation (possibly through coding) can have a logical symbol of its own truth predicate.

However, this situation, when *ML* is an extension of *L*, and so the truth predicate symbol is a logical symbol of *ML*, opens the possibility of the paradoxes of truth. In a standard situation in science, atomic sentences of the language *L* do not contain the truth predicate "T," and they have a certain truth value as the result of rational cognition. Such a situation does not lead to paradoxes. Namely, according to the above-described truth condition on the logical symbol "T," in order to examine whether the atomic sentence $T(\ulcorner\varphi\urcorner)$ of the language *ML* is true, we need to examine whether the sentence φ of the language *L* is true, and its truth is completely determined by the truth of the atomic sentences of the language *L*. Thus, the truth value of the sentence $T(\ulcorner\varphi\urcorner)$ is unambiguously determined. However, in a natural language the truth predicate symbol is applicable to all its sentences ($L = ML$): *L* contains "T." Now, too, by the truth condition on the logical truth predicate "T," the examination of the truth of the atomic sentence $T(\ulcorner\varphi\urcorner)$ is reduced to the examination of the truth of the sentence φ, and the examination of its truth is reduced to the examination of the truth of atomic sentences. But now some of these atomic sentences can again be of the form $T(\ulcorner\psi\urcorner)$, so that the process does not stop but continues again. While for the standard language *L* which speaks of some natural phenomenon and does not contain its own truth predicate symbol, this procedure gives a unique answer, now we have no guarantee that the reduction procedure will stop at some step or that we will get unique truth values of sentences covered by such procedure. Let us consider the two simplest examples where the truth determination procedure is not successful:

the sentence L: not $\mathrm{T}(\overline{\mathrm{L}})$ (The Liar)

the sentence I: $\mathrm{T}(\overline{\mathrm{I}})$ (The Truth-teller)

For the sentence L we have the following chain of reduction:

$$\mathrm{L} \mapsto \text{not } \mathrm{T}(\overline{\mathrm{L}}) \mapsto \mathrm{T}(\overline{\mathrm{L}}) \mapsto \mathrm{L} \mapsto \ldots$$

It is easy to see that no evaluation along this chain satisfies the truth conditions: the assumption that L is true gives that L is false, and the assumption that L is false gives that L is true. Thus, we cannot assign any truth value to the sentence L. On the other hand, for the sentence I we get the following chain of reduction:

$$I \mapsto T(\overline{I}) \mapsto I \mapsto \ldots$$

Now both evaluations, the evaluation according to which I is true and the evaluation according to which I is a false sentence, satisfy the truth conditions along the chain. So, this sentence can be both true and false in an equally (un)convincing way.

The paradoxes of truth stem precisely from the fact that the classical procedure of determining truth values, which grew out of everyday language use, does not always have to give a classically assumed (and expected) unique answer. Such an assumption is an unjustified generalization from common situations to all situations. We can preserve the classical procedure, but we must reject universality of the assumption of its success. The awareness of that transforms paradoxes of truth to normal situations inherent to the classical procedure.

The conclusion is that the paradoxes of truth arise from the internal organization of language, so they belong to the logical aspect of the concept of truth and do not concern the synthetic concept of truth. Thus, the solution should be sought in the internal organization of the language.[15]

10. Epilogue

In 1991, Milošević and Tuđman, presidents of Serbia and Croatia, met in Karađorđevo, in the former Yugoslavia. They talked behind closed doors, with no witnesses, and no record was left of the conversation. Did they then make an agreement on the partitioning of Bosnia and Herzegovina along so-called ethnic lines, and thus destroy so many human lives and cause so much human suffering? The synthetic concept of truth gives us the legitimacy to ask that question, and all of the above aspects of the concept of truth can help us get the answer.

[15] A good overview of various solutions to the paradoxes of truth can be found in (Beall, Glanzberg and Ripley 2023). The author's solution can be found in (Čulina 2001, Čulina 2023).

References

Arendt, Hannah. 1973. *The Origins of Totalitarianism*. NewYork: Harcourt Brace Jovanovich.

Armour-Garb, Bradley, Daniel Stoljar, and James Woodbridge. 2023. "Deflationism About Truth." *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition), edited by Edward N.Zalta and Uri Nodelman. https://plato.stanford.edu/archives/sum2023/entries/truth-deflationary/

Atiyah, Michael F. 1995. Review of *Conversations on Mind, Matter and Mathematics*, by Jean-Pierre Changeux and Alain Connes. *Times Higher Educational Supplement*, September.

Brown, Jessica, and Herman Cappelen, eds. 2011. *Assertion: New Philosophical Essays*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573004.001.0001

Beall, Jc, Michael Glanzberg, and David Ripley. 2023. "Liar Paradox." *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2023/entries/liar-paradox/

Bennett, Max R., and Peter M. S. Hacker. 2022. *Philosophical Foundations of Neuroscience*. Hoboken, New Jersey: Wiley.

Broad, Charlie D. 1952. *Ethics and The History of Philosophy*. London: Routledge.

Bulgakov, Mikhail. 1997. *Master and Margarita*. London: Penguin Books.

Carnap, Rudolf. 1936. "Testability and Meaning." *Philosophy of Science*, 3(4): 419–71. https://doi.org/10.1086/286432

Carnap, Rudolf. 1966. *Philosophical Foundations of Physics*. New York: Basic Books.

Čulina, Boris. 2001. "The Concept of Truth." *Synthese*, 126: 339–60. https://doi.org/10.1023/A:1005237331006

Čulina, Boris. 2018. "An Elementary System of Axioms for Euclidean Geometry Based on Symmetry Principles." *Axiomathes*, 28(2): 155–80. https://doi.org/10.1007/s10516-017-9358-y

Čulina, Boris. 2021. "The Language Essence of Rational Cognition with Some Philosophical consequences." *Tesis* (*Lima*), 14(19): 631–56. https://philpapers.org/archive/CULTLE.pdf

Čulina, Boris. 2022. "A Simple Interpretation of Quantity Calculus." *Axiomathes* (online first) https://doi.org/10.1007/s10516-021-09609-9

Čulina, Boris. 2023. "How to Conquer the Liar and Enthrone the Logical Concept of Truth." *Croatian Journal of Philosophy*, 23(67): 1–31. https://doi.org/10.52685/cjp.23.67.1

Čulina, Boris. 2023. "Mathematics: An Imagined Tool for Rational Cognition." https://philpapers.org/archive/CULMA.pdf. The first seven sections have been accepted for publication in the *Annals of Mathematics and Philosophy*.

Čulina, Boris. 2024. "Logic as an Internal Organisation of Language." *Science and Philosophy*, 12(1):62–71. https://doi.org/10.23756/sp.v12i1.1582

Cumming, Sam. 2023. "Names." *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2023/entries/names/

Davidson, Donald. 2005. *Truth and Predication*. Cambridge, Massachusetts: Harvard University Press. https://doi.org/10.4159/9780674030220

David, Marian. 2022. "The Correspondence Theory of Truth." *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/sum2022/entries/truth-correspondence/

Einstein, Albert. 1936. "Physics and reality." *Journal of the Franklin Institute*, 221(36): 349–82. https://doi.org/10.1016/S0016-0032(36)91047-5

Frege, Gottlob. 1891. "Funktion und Begriff." *Vortrag, gehalten in der Sitzung vom 9. Januar 1891 der Jenaischen Gesellschaft für Medizin und Naturwissenschaft.*

Frege, Gottlob. 1892a. "Ausfuhrungen über Sinn und Bedeutung." In *Gottlob Frege: Nachgelassene Schriften*, edited by Hermes at al., 1983, 128–36.

Frege, Gottlob. 1892b. "Über Sinn und Bedeutung." *Zeitschrift für Philosophie und philosophische Kritik*, 100: 25–50.

Frege, Gottlob. 1892c. "Über begriff und gegenstand." *Vierteljahresschrift für wissenschaftliche Philosophie*, 16: 192–205.

Frege, Gottlob. 1897. "Logik." In *Gottlob Frege: Nachgelassene Schriften*, edited by Hermes at al., 1983, 1–8.

Frege, Gottlob. 1906a. "Kurze Übersicht meiner logischen Lehren." In *Gottlob Frege: Nachgelassene Schriften*, edited by Hermes at al., 1983, 213–18.

Frege, Gottlob. 1906b. "Einleitung in die Logik." In *Gottlob Frege: Nachgelassene Schriften*, edited by Hermes at al., 1983, 201–12.

Frege, Gottlob. 1918. "Der Gedanke." *Beiträge zur Philosophie des deutschen Idealismus*, 2: 58–77.

Glanzberg, Michael. 2023. "Truth." *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/fall2023/entries/truth/

Hermes, Hans, Friedrich Kambartel, and Friedrich Kaulbach, eds. 1983. *Gottlob Frege: Nachgelassene Schriften*. Hamburg: Felix Meiner Verlag.

Horsten, Leo. 2023. "Philosophy of Mathematics. *The Stanford Encyclopedia of Philosophy* (Winter 2023 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/win2023/entries/philosophy-mathematics/

Hume, David. 1738–1740. *A Treatise of Human Nature*. London: John Noon.

Humboldt, Wilhelm von. 1836. *Über die Verschiedenheit des menschlichen Sprachbaus und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. Bonn: F. Dümmler.

James, William. 1907. *Pragmatism: A New Name for Some Old Ways of Thinking: Popular Lectures on Philosophy*. New York: Longmans, Green. https://doi.org/10.1037/10851-000

Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Mac Lane, Saunders. 1986. *Mathematics, Form and Function*. New York: Springer-Verlag. https://doi.org/10.1007/978-1-4612-4872-9

Margolis, Eric, and Stephen Laurence. 2023. "Concepts." *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/fall2023/entries/concepts/

Orilia, Francesco, and Michele Paolini Paoletti. 2022. "Properties." *The Stanford Encyclopedia of Philosophy* (Spring 2022 Edition), edited by Edward N. Zalta. https://plato.stanford.edu/archives/spr2022/entries/properties/

Patterson, Douglas. 2012. *Alfred Tarski: Philosophy of Language and Logic*. London: Palgrave Macmillan. https://doi.org/10.1057/9780230367227

Pedriali, Walter B. 2017. "The Logical Significance of Assertion: Frege on the Essence of Logic." *Journal for the History of Analytical Philosophy*, 5(8). https://doi.org/10.15173/jhap.v5i8.3073

Psillos, Stathis. 2000. "Rudolf Carnap's 'Theoretical Concepts in Science.'" *Studies in History and Philosophy of Science Part A*, 31(4): 151–72. https://doi.org/10.1016/s0039-3681(99)00031-x

Quine, Willard Van Orman. 1948. "On What There Is." *Review of Metaphysics* 2(5): 21–38.

Quine, Willard Van Orman. 1951. "Two Dogmas of Empiricism." *The Philosophical Review*, 60: 20–43.

Quine, Willard Van Orman. 1990. *Pursuit of Truth*. Cambridge, Massachusetts: Harvard University Press.

Russell, Bertrand. 1905. "On Denoting." *Mind*, 14(56): 479–93.

Schlick, Moritz. 1936. "Meaning and Verification." *Philosophical Review*, 45(4): 339–69. https://doi.org/10.2307/2180487

Tarski, Alfred. 1933. *Pojęcie prawdy w językach nauk dedukcyjnych*. Warszawa: Towarzystwo Naukowe Warszawskie.

Tarski, Alfred. 1944. "The Semantic Conception of Truth." *Philosophy and Phenomenological Research*, 4: 341–76,

Waismann, Friedrich. 1968. "Verifiability." In *The Theory of Meaning*, edited by George Henry Radcliffe Parkinson, 35–60. Oxford: Oxford University Press.

Whorf, Benjamin Lee. 1940. "Science and Linguistics." *Technology Review*, 42: 229–231, 247–48.

Wilczek, Frank. 2013. "The Enigmatic Electron." *Nature*, (498): 31–2. https://doi.org/10.1038/498031a

Organon F 32 (1) 2025: 92–106 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32105

DISCUSSION NOTE

Rectification Note to "Riemann's Philosophy of Geometry and Kant's Pure Intuition"

(Organon F, 31(2), 114–140)

Dinçer Çevik*

I would like to extend my gratitude to those who have given me the opportunity to present my perspective on this matter.

The philosophy of geometry is a highly specialized field of study. More specifically, research focusing on Riemann's philosophy of geometry is exceptionally limited. Of course, this cannot justify inappropriate use of others' work, it will be observed in the details provided, many of the analyses (if not all) rely on a shared set of interpretations regarding Riemann's *Habilitationsvortrag*. As a result, it is exceedingly challenging to introduce entirely novel claims or terminology. In this regard, based on my observations in the literature, similarities are often unavoidable within the scope of what is considered "common knowledge." What is common knowledge? According to *The Harvard Guide to Using Sources*, it is defined as follows: "The only source material that you can use in an essay without attribution is material that is considered common knowledge and is therefore not attributable to one source. Common knowledge is information generally known to

* Mugla Sitki Kocman University
https://orcid.org/0000-0001-5897-7381
Mugla Sitki Kocman University, Faculty of Letters, Department of Philosophy, Mugla, 48000, Turkey
dincercevik@mu.edu.tr

an educated reader, such as widely known facts and dates, and, more rarely, ideas or language."[1]

One should be careful to distinguish between specific details, claims, and views, and what is considered common knowledge—characterized by its cumulative, general, and widely shared nature. This pertains to the content. Additionally, one should be cautious about similarities in form, such as structural and grammatical patterns. As the simple yet powerful motto states, "Correlation does not imply causation."

Therefore, apart from content-related issues—which largely rely on common knowledge, as I will demonstrate—the other examples provided by the anonymous scholar pertain to grammatical structures.

Particulars

Arguably, Riemann and Gauss are two of the most important figures in the field of geometry; their ideas were groundbreaking. It is not uncommon for discussions about revolutionary concepts, claims, and approaches within a specific area to lack significant expansion or diversification.

I have been in ongoing communication with prominent researchers in the field, including those with the scholars I am accused of not properly citing. I have been in communication with them since the period of my master's degree, including during the preparation process of this paper, and I am grateful to them. On the other hand, I only heard of names like Windham—whom the anonymous researcher claims I cited improperly—for the first time. I was not even aware that this scholar worked in this field.

Different versions of the paper were reviewed by the prominent researchers. Before being submitted to *Organon-F*, the paper was also sent to two other well-known international journals in the fields of philosophy and history of science.

In fact, during the revision process of the paper, I again contacted one of the researchers in question regarding a reference for a piece of information

[1] https://usingsources.fas.harvard.edu/exception-commonknowledge#:~:text=Common%20knowledge%20is%20information%20generally,and%20must%20always%20be%20cited. (Accessed 16.10.2024).

I used in the paper. All these details, along with evidence, have already been presented in the letter previously sent to the editorial board of *Organon-F*.

My detailed explanations, along with examples from the literature, are provided in the continuation of the letter.

DÇ (PAGE 125): In the mathematical part of his *Habilatitionsvortrag*, Riemann follows Gauss' most fundamental steps, by extending Gaussian concepts and results for surfaces to n -dimensional manifolds, such as the measure of curvature and some properties of geodesic lines. Like Gauss, Riemann's approach is metric; the concept of distance plays a fundamental role both in the theory of curved surfaces and in Riemannian manifolds; in addition, the essential properties of manifolds are expressed by means of the linear element.

Original Rossana Tazzioli: "In his memoir Riemann follows Gauss's fundamental steps, by extending Gaussian concepts and results for surfaces to n-dimensional manifolds, such as the measure of curvature and some properties of geodesic lines. Riemann's approach, as that of Gauss, is metric; in fact, the concept of distance plays the fundamental role whether in the theory of curved surfaces or in Riemannian theory of manifolds; moreover, the essential properties of manifolds are expressed by means of the linear element."

Explanation: It is well-known that Gauss developed a method to describe the geometry of surfaces using what we now call the metric tensor. This tensor encodes information about distances and angles on a surface. Riemann later introduced the concept of a Riemannian manifold, a generalization of surfaces to spaces of any dimension. He extended the notion of a metric from surfaces to arbitrary spaces, introducing what is now known as the Riemannian metric. This innovation allows for the measurement of distances and angles in spaces of any dimension, whether they are curved or flat. To the best of my knowledge, there is no alternative approach to the metric other than the one I have described, which directly traces Riemann's ideas back to Gauss.

It is common to analyze the relationship between Riemann and Gauss using the conceptual resources mentioned above: the role of distance, the line element, curved surfaces, geodesic lines, and similar concepts. These

ideas, among others, were inherited by Riemann from Gauss. There are numerous sources that illustrate these connections. For instance, Olivier Darrigol's 2015 paper frequently addresses the underlying relationships between the methodologies and conceptual frameworks of Gauss and Riemann. I will quote a few key points from the paper to highlight these connections.

In the second section, I use this Gaussian background to analyze Riemann's results regarding the curvature of a manifold in his habilitation lecture on the one hand, and his results regarding the transformation properties of quadratic differential forms in the *Commentatio* on the other hand (a *quadratic differential form* is an expression of the type…) (2015, 48-49).[2] See another example where V. F. Kagan presents and explains the line element (mentioned seven times in the article) and other conceptual resources used to define how Riemann generalizes Gauss's metric approach[3]. Kagan also elaborates on the generalization process:

> All their originality notwithstanding, Riemann's ideas are an extension of the methods of investigation of surfaces presented by Gauss in his *Disquisitiones generales* [4] of 1827. Gauss's key idea is that a point on a surface (to be sure, in ordinary Euclidean space) is determined by two coordinates x^1 and x^2 (this is modernized symbolism), and a line element is expressed in terms of a given positive definite quadratic form in the differentials of these coordinates. Specifically (again using modem symbolism) (2005, p.80).

You will notice similar points. It is well-known that both Riemann's and Gauss's approaches are metric-based. It is also widely recognized that both make use of curvature, geodesic lines, and similar concepts. Additionally, the concept of the line element is well-established. Therefore, the conclusion is that Riemann recognized the path Gauss had opened and followed the "fundamental steps" Gauss had already taken. It is generally agreed that

[2] Darrigol, O. (2015). The mystery of Riemann's curvature. *Historia Mathematica*, 42 (1): 47-83.

[3] V. F. Kagan, (2005). Riemann's Geometrical Ideas. *The American Mathematical Monthly*, Vol. 112, No. 1, pp. 79-86).

Riemann followed the path laid out by Gauss, relied on the conceptual framework he inherited from him, and built his own methodology upon it.

DÇ: Developing this approach enabled Riemann to investigate the links between different laws of nature-knowledge of which is based on the exactness of our description of phenomena in infinitesimal regions. Gaining knowledge of the external world from the behaviour of infinitesimal parts constitutes the backbone of Riemann's research program.

Rossana Tazzioli: This approach is typical of all Riemann's work and leads him to investigate links between different laws of nature, knowledge of which is based on the exactness of our description of phenomena in infinitesimal regions. As Hermann Weyl wrote in *Space, time, matter*, "the principle of gaining knowledge of the external world from the behaviour of its infinitesimal parts is the mainspring of the theory of knowledge in infinitesimal physics as in Riemann's geometry."

Explanation: Riemann's philosophy of geometry is also closely related to physics. One of the key starting points for analyzing this relationship is the use of infinitesimals and the laws of nature. For example, W. Ehm emphasizes points similar to those I focus on in my paper. Riemann too, wants to circumvent mysterious actions at a distance and postulates the existence of a space-filling stuff of which he assumes that it behaves like anincompressible homogeneous fluid without inertia. He largely avoids using the term aether and rather speaks of 'stuff' (Stoff), but he clearly adopts the aether hypothesis. Often he speaks of Stofftheilchen if he wishes to indicate that *infinitesimally small portions* of the aether are being considered (emphasis added, 2010, p. 148).

See another example from Papadopoulos, who discusses the importance of the infinitely small and large in Riemann's research program: "Likewise, Riemann's speculations on the infinitely small and the infinitely large go beyond the mathematical and physical setting, and they had a non-negligible impact on philosophy"[4] (Papadopoulos 2017, p.1).

See Papadopoulos quoting Klein:

[4] Ji, Lishen et. al. (2017). *From Riemann to Differential Geometry and Relativity.* Springer.

> I must mention, first of all, that Riemann devoted much time and thought to physical considerations. Grown up under the tradition which is represented by the combinations of the names of Gauss and Wilhelm Weber, influenced on the other hand by Herbart's philosophy, he endeavored again and again to find a general mathematical formulation for *the laws underlying all natural phenomena*. [...] The point to which I wish to call your attention is that these physical views are the mainspring of Riemann's purely mathematical investigations (emphasis added, 2017, p.8).[5]

Elsewhere, Papadopoulos directly quotes Riemann himself in the context of the laws of nature:

> Right after the submission of my HabilitationsschriftindexRiemann! habilitation text I resumed my investigations into the coherence of the *laws of Nature* and got so involved in it that I could not tear myself loose. The continuing preoccupation with it has become bad for my health, in fact, right after New Year's my usual affliction set in which such persistence, that I could only obtain relief through the strongest remedies. As a result I felt very ill, felt unable to work, and sought to again put my health in order through long walks (emphasis added, 2017, p.37).[6]

In fact, it is easy to provide numerous examples showing that one of the fundamental issues in Riemann's research program was to delve into the details of nature and the laws governing it. In this regard, W. Ehm quotes Riemann: "[The] purpose [of the paper] is to penetrate the inner of nature, beyond the foundations of astronomy and physics layed by Galilei and Newton." (Riemann quoted in Ehm, 2010, p.147).[7]

---

[5] Ji, Lishen et. al. (2017). *From Riemann to Differential Geometry and Relativity.* Springer.

[6] Ji, Lishen et. al. (2017). *From Riemann to Differential Geometry and Relativity.* Springer.

[7] Ehm, W. (2010). Broad views of the philosophy of nature: Riemann, Herbart, and the "matter of the mind." *Philosophical Pschhology*, Volume 23, Issue 2, pp. 141-162.

In another quotation that illustrates the significance of discovering the laws of nature within the context of Riemann's research program, L. Ji quotes Klein:

> He [Riemann] endeavored again and again to find a general mathematical formulation for the laws underlying all natural phenomena… these physical views are the mainspring of Riemann's purely mathematical investigations.[8](emphasis added, Ji, 2017, p. 167).[9]

DÇ: Accordingly, Riemann set himself two tasks: The first (a philosophical task) was to define a manifold extension. The second (an empirical task) was to give definitions of intrinsic curvature and measure determined from within extension.

Original Banks: The first task is to define a manifold extension. The second task is to give definitions of intrinsic curvature and measure determined from within this extension, say by introducing rigid meter sticks or light beams.

Explanation: Indeed, in the abstract of the paper, I already provide hints about the philosophical and empirical tasks: "The aim of this paper is twofold: first to explicate how Riemann's philosophy of geometry is organized around the concept of manifold. Second, to argue that Riemann's philosophy of geometry does not dismiss Kant's spatial intuition." What I want to emphasize by saying this is that in the article, I rely on Riemann's methodology, which is built upon philosophy, mathematics, and physics through the concept of the manifold. Papadopoulos puts this idea by saying that "Roughly speaking, the first part is philosophical, the second one is mathematical, and the third one deals with applications to physics. But to some extent philosophy and physics are present in the three parts" (2017, p. 27).[10]

---

[8] Ji, Lishen et. al. (2017). *From Riemann to Differential Geometry and Relativity.* Springer.

[9] Ji, Lishen et. al. (2017). *From Riemann to Differential Geometry and Relativity.* Springer.

[10] Ji, Lishen et. al. (2017). *From Riemann to Differential Geometry and Relativity.* Springer.

Anyone familiar with Riemann's interest in philosophy would explain the methodology of his 1854 paper as relating the concept of a manifold to its philosophical and mathematical context, before transitioning to its physical applications. For instance, see the following: "For B. Riemann it was *an intricate task to formulate his general concept of manifolds (discrete or continuous) in 1854*. At that time, it was technically impossible to give a formal definition of topological spaces and to specify manifolds among them. Not even a general concept of set was around; Riemann's step rather contributed to bring it about (Ferreiros [Ferreiros 1999]. Around the turn from the 19th to the 20th century the situation started to change [Scholz 1999]"[11](emphasis added). E. Scholz also explains that "On the other hand Herbart's epistemology and his ideas on the relationship between philosophy and sciences do seem have influenced Riemann and thus Riemann's perception of *the task of mathematics*" (p.427).[12] He goes on to say that "In fact, it is precisely for this reason that Riemann's approach to mathematics is sometimes referred to as 'conceptual mathematics'" (Ibid.). [13]

In another paper Scholz underlines the following:

> As is well known, Riemann organized his approach to geometry around the new concept of manifold (Mannigfaltigkeit) which for obvious reasons he could not define in a mathematical technical sense. He therefore did it in a semi-philosophical way, drawing consciously and cautiously upon hints by C. F. Gauss who had spoken geometrically about complex numbers (Gauss 1831) and J. F. Herbart who had argued for the use of geometrical imagery in all kind of concept formation, his so-called serial forms (Reihenformen) (emphasis added, Scholz, 2005, p.22).[14]

[11] http://www.map.mpim-bonn.mpg.de/Axiomatization_of_the_manifold_concept (Accessed: 16.10.2024).

[12] Scholz, E. (1982). Herbart's Influence on Bernhard Riemann. *Historia Mathematica*, Volume 9, Issue 4, pp. 413-440.

[13] Scholz, E. (1982). Herbart's Influence on Bernhard Riemann. *Historia Mathematica*, Volume 9, Issue 4, pp. 413-440.

[14] Scholz, E. (2005). Riemann's Vision of a New Approach to Geometry. In *1830-1930: A Century of Geometry,* pp. 22-34.

The second task (i.e., the empirical task) is directly related to the first. This is already evident in Riemann's own article:

> … Now it seems that the empirical concepts on which *the metrical determinations of space* are founded, namely, the concept of a rigid body and that of a light ray, are not applicable in the infinitely small; it is therefore quite conceivable that *the metrical relations of space* in the infinitely small do not agree with the assumptions of geometry; and indeed we ought to hold that this is so if phenomena can thereby be explained in a simpler fashion (emphasis added, Riemann, quoted in Boi, p.199).

Hence, it is evident that Riemann's program is related to a) the concept of the manifold (philosophical task), and b) the determination of metrical relations within extension.

DÇ: Riemann discusses the problem of what he calls 'multiply extended magnitude' in his famous lecture "On the Hypotheses Which Lie at the Foundation of Geometry." Riemann's introduction clearly shows that he saw himself involved in a philosophical as well as mathematical enterprise.

Original Banks: "Bernhard Riemann discusses the problem of what he calls multiply extended magnitude in his famous lecture "On the Hypotheses that Lie at the Foundation of Geometry.'"

Nowak: "Riemann's introduction made it clear that he saw himself involved in a philosophical as well as a mathematical enterprise."

Explanation: The concept of a manifold is also referred to as a multiply extended magnitude. Riemann articulates this fundamental concept in his "On the Hypotheses Which Lie at the Foundation of Geometry." Erhard Scholz's 1982 article has already put the idea that the Riemann's concept of manifold is a "semi-philosophical" in character. Also, Laugwitz (1999)[15], Ferreirós and several other scholars have pointed out that Riemann's approach in his lecture is philosophical, mathematical, and physical. That is

[15] Laugwitz, D. (2008). *Bernhard Riemann, 1826-1866: Turning Points in the Conception of Mathematics*. Birkhäuser.

why Ferreirós refers to Riemann's *Habilitationsvortrag* as the "Magic Triangle" (2006, p.67).[16]

See Freudenthal regarding the philosophical dimension of Riemann's geometry:

> one of the most profound and imaginative mathematicians of all time, he had a strong inclination to philosophy, indeed, was a great philosopher. Had he lived and worked longer, philosophers would acknowledge him as one of them[17] (Freudenthal, 1975, p. 448).

DÇ: Riemann's main concern was construction of space, rather than construction in space.

Nowak: "First, Herbart's constructive approach to space, already cited, mirrored the content of Riemann's reference to Gauss in that both discussed construction of spaces rather than construction in space."

This section relates to a conclusion I reached as a result of discussions conducted via email with another prominent philosopher in the field (the details of which, along with supporting evidence, have been submitted to the journal's editorial board).

I chose to mention Nowak in a controlled manner because, in discussions with a prominent philosopher of geometry in the field, it was agreed that his views are no longer considered relevant, and his paper was rather inconclusive. On the other hand, anyone who has worked to some extent on Riemann's 1846 paper will see that his primary concern was the construction of space. The concept of a manifold is primarily related to the construction of space, not to specific geometric and topological objects within a given space.

DÇ: In Herbart's view, experience shows us properties and bundles of properties, while the underlying reality must be searched for within the things to which properties are ascribed. This distinction between the phenomena and a more stable underlying reality, and an investigation of the

[16] Ferreiros, J and Gray, J. (2006). *The Architecture of Modern Mathematics: Essays in History and Philosophy*. Oxford University Press.

[17] Freudenthal, H. (1975). Riemann, Georg Friedrich Bernhard. In: *Dictionary of Scientific Biography* vol. 11. New York, 447-456.

relationship between them, is essential in Riemann's own reflections about the epistemology of science.

Original Scholz: In the first place, according to Herbart, experience shows us properties and bundles [Complexionen] of properties, theunderlying reality of which must first be sought in things to which the properties are ascribed..... The distinction however between the phenomena and a more stable underlying reality, with an intense relationship between both became an essential point in Riemann's own reflections about the epistemology of science.

Explanation: Herbart himself explains that:

> Are the forms of experience given? Yes indeed they are given, although only as determinations of the manner in which sensations are bound up together. Were they not given, we could not only sunder them from sensation in such a way that the sensed could occur completely isolated, without any connection; rather we could also, at pleasure, see different shapes, hear other time intervals; similarly we could put things together arbitrarily out of properties and change them (Herbart quoted in Banks, 2005, p. 209).[18]

Someone who recognizes the empiricist aspects of Herbart's philosophy would make similar observations. Herbart's initial adherence to Kant also enables analyses concerning the relationship between empiricism and the unchanging reality found here. With this understanding, Riemann's effort to grasp the laws of nature and nature itself from within, through infinitesimals, can also be linked. Similar interpretations have also been made by other researchers in the field:

> Another important element that Riemann inherited from Herbart was a developmental, genetic understanding of science. Far from the usual idea that there exists (in some Platonic realm) a ready-made theory of everything, in his view all concepts of natural science, and of mathematics in particular, have evolved gradually

[18] Banks, E. (2005). Kant, Herbart and Riemann. *Kant-Studien*, 96(2), pp. 208-234.

> from older explanatory systems. Scientific theories are for Riemann the outcomes of a process of gradual transformation of concepts, starting from the basic ideas of object, causality, and continuity. Development takes place under the pressure of contradictions or else implausibilities [*Unwahrscheinlichkeiten*] revealed by unexpected observations – unexpected in light of the hypotheses proposed by reflection at some particular stage (Ferreiros, 2006, p. 77).[19]

> Riemann and Herbart tended to prefer a Leibnizian view on space, seeing space not as a basic entity, but as an expression of the relations between physical entities (see end of Section 3.1). It is particularly interesting that this viewpoint led Herb- art to think geometrically about all kinds of subjects, which again was noticed and further developed by Riemann. This is a feature of Herbart's work that makes it vaguely reminiscent of modern mathematics, and perhaps it has not been sufficiently appreciated (Ferreiros, 2006, p.75).[20]

DÇ: In *Habilitationsvortrag*, Riemann generalizes the Gaussian theory of curved spaces to n -dimensions. Such manifolds are characterized by the fact that each point within them can be uniquely specified by n real numbers. The introduction of the concept of distance into a manifold follows the Gaussian model. Analogously to the two-dimensional case, infinitesimal distances are expressed by processing differentials given in terms of some internal coordinate system, u, with the help of the metric tensor g ij. Thus, Riemann arrives at a formula that is identical to the Gaussian expression for the surfaces.

Original Carrier: In 1854, Bernhard Riemann generalized the Gaussian theory of curved surfaces to n-dimensional manifolds. Such manifolds are characterized by the fact that each point in them can be uniquely specified by n real numbers. The introduction of the concept of length or dis-

[19] Ferreiros, J and Gray, J. (2006). The Architecture of Modern Mathematics: Essays in History and Philosophy. Oxford University Press.

[20] Ferreiros, J and Gray, J. (2006). The Architecture of Modern Mathematics: Essays in History and Philosophy. Oxford University Press.

tance into such a manifold closely follows the Gaussian model. Quite analogous to the two-dimensional case we express infinitesimal lengths by processing coordinate differences (or differentials) as given in terms of some internal coordinate system u with the help of the metric tensor gik. One arrives at a formula that is identical, mutatis mutandis, with the corresponding Gaussian expression for surfaces.

Explanation: This part is technical and common. One of the most important results of the Riemann's 1846 lecture is his generalization of Gaussian view of curved spaces to n-dimensions. This is a *technical fact.*

> […] he [Riemann] expresses the square of a line element by means of a positive definite quadratic form in the differentials dx^i of the coordinates whose co- efficients are functions of the coordinates x^i .This *relation is not just an extension of Gauss's formula to an n-dimensional manifold*. Rather, it introduces the completely new idea of determining the metric on a manifold by specifying it in an infinitely small portion of that manifold… (Kagan, 2005, p. 81).[21]

Indeed, it can be said that one of the most important developments Riemann brought about was this generalization. What I mean is that this is a paradigmatic example of common knowledge that illustrates how Riemann reaches n-dimensions by following Gauss's footsteps. This technical aspect can be seen in any discussion of Riemannian spaces:

> The theory of Riemannian spaces. A Riemannian space is an n-dimensional connected differentiable manifold M^n on which a differentiable tensor field g of rank 2 is given which is covariant, symmetric and positive definite. The tensor g is called metric-tensor. Riemannian geometry is a multi-dimensional generalization of the intrinsic geometry (cf. İnterior geometry) of two-dimensional surfaces in the Euclidean space E^3. The metric of a Riemannian space coincides with the Euclidean metric of the domain under consideration up to the first order of smallness. The

[21] V. F. Kagan, (2005). Riemann's Geometrical Ideas. *The American Mathematical Monthly*, Vol. 112, No. 1, pp. 79-86).

> difference between these metrics is (locally) estimated by Riemannian curvature- a multi-dimensional generalization of the concept of the Gaussian curvature of a surface in E^3.[22]

See also the following which explains the generalization procedure in a less technical sense:

> Gauss's theorem suggests that one could see the surface as an independent curved manifold and then to generalize this concept to higher dimensions via "the concept of a multiply extended magnitude," … which is what Riemann did. These concepts also helped him to generalize to higher dimensions Gauss's concept of curvature. To do so yer another new concept, another great invention of Riemann, the tensor of curvature, and a new form of differential calculus, tensor calculus on manifolds, a generalization of differential calculus[23] (Plotnitsky, p.344-345, 2017).

Conclusion

Of course, points of differentiation and approaches should be conveyed according to the relevant references. On the other hand, in studies on the philosophy of geometry and Riemann, some themes, fundamental concepts, Riemann's innovations, and observations, claims, and concepts about how he achieved them have now become mostly common knowledge.

One could establish connections between various relations and believe that the relations they themselves connect are factual. However, there is a significant difference between believing something and justifying it. As I have shown above, the allegations put forward by the anonymous scholar are unjustifiable. One may notice some similarities between words, names, claims, or views. However, the fact that there are *similarities* in how the

[22] https://encyclopediaofmath.org/wiki/Riemannian_geometry (Accessed: 16. 10. 2024).

[23] Ji, Lishen et. al. (2017). *From Riemann to Differential Geometry and Relativity.* Springer.

issues are presented does not mean they are the *same*; fundamentally, *similarity* does not equal *equality*, especially given that there is limited literature regarding Riemann that is common to all scholars.

Sincerely,
Dinçer Çevik

Organon F 32 (1) 2025: 107–110 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32106

REPORT

Pavel Cmorej (1937 – 2025)

Pavel Cmorej, philosopher, logician, researcher and university professor, died on January 16, 2025. As one of the most prominent representatives and defenders of analytic philosophy in Slovakia, he contributed to its promotion and expansion in our geographical context. Cmorej excelled in his extraordinary scientific, pedagogical, and editorial activities. With his passing, our philosophical scene loses not only an original and inspiring thinker who stood out for his intellectual honesty and passion for thorough conceptual analysis, but also an excellent teacher who was able to explain complex philosophical and logical issues in an accessible and witty manner.

Pavel Cmorej was born on January 1, 1937 in Spišské Podhradie. He completed his studies of philosophy and psychology at the Faculty of Arts, Comenius University Bratislava in 1961, and in 1967 he defended his Ph.D. thesis. Since 1963, he became a member of the faculty's Department of Logic and Methodology of Sciences. In the 1970s, during the period of so-called normalization, he was forced to end his teaching activities due to his civic stances. Similarly, he could publish the results of his scientific work only rarely, but mostly he could not do so under his own name. Despite this, he did not interrupt his intellectual work. In the late 1980s, he resumed

his teaching career. He received his habilitation in 1992 and became a professor of philosophy in 2001. Although he started working at the Institute of Philosophy of the Slovak Academy of Sciences in 1990, he lectured at several Slovak universities until 2003.

The expansion of analytic philosophy in our field was significantly helped by the founding of the journal *Organon F*. It was Pavel Cmorej who, in the early 1990s, was behind the idea of a Slovak and Czech-Slovak journal focusing on analytic philosophy. This project was quickly implemented, and the first issue of the journal was published as early as in 1994. Cmorej led *Organon F* as its editor-in-chief for more than a decade and made a significant contribution to establishing it as an internationally respected journal, which publishes research articles by prominent analytic philosopher from all over the world.

Cmorej's most significant research achievements have undoubtedly been in the fields of logical semantics, philosophy of language and philosophy of logic. He has authored four monographs, edited several books, and written more than a hundred research papers. His logical and philosophical views are most comprehensively presented in the books *On the Border of Logic and Philosophy* (2001) and *Analytical Philosophical Investigations* (2009). His book *Introduction to Logical Syntax and Semantics* (2001) is valuable from a pedagogical perspective.

One of the most important contributions is Cmorej's discovery of the existence of so-called empirical essential properties. This result was published in a series of research articles, namely "Essentialism versus Anti-Essentialism" (1988), "Empirical Essential Properties" (1996), and "Bare Individuals and Predication" (2006). The existence of empirical essential properties challenges radical anti-essentialist positions.

Cmorej's work in the field of the philosophy of language covers a wide range of topics. Most notably, he examined in detail several ways in which linguistic expressions relate to extralinguistic objects and some of the ways in which speakers identify objects by using expressions (most notably, in "Denotation and Reference" (1998), "Reference, Identification, and Intentional Acts" (2000), "What Do Linguistic Signs Denote?" (2000), "Usage, Mode of Use and Meaning of Linguistic Expressions" (2004), "On the Nature of Intentional Reference" (2005), etc.). He also developed intriguing

views regarding linguistic signs (especially in "On the Explication of the Concept of a Linguistic Sign" (1985)) and the distinction between expressions and semi-expressions (in "Semi-Expressions and Expressions" (2005)). Cmorej also connected problems from the philosophy of language with difficult questions concerning the nature of objects and their existence (for example, "Towards the Concepts of Identification" (2008), "From Descriptions to Their Referents" (2013), "Possible Individuals and Possibilities" (2015)).

Cmorej was a keen champion of intensional logic and semantics. In a series of articles, he contributed to the refinement of the conceptual apparatus of the Transparent Intensional Logic (TIL), which currently represents one of the most complex semantic approaches to natural language (for example, "'Paradoxes' of Intensional Logic" (1989), "Temporal Semantics and Speech Acts" (1990), and "On the Explication of Some Meanings in TIL" (2000)). Cmorej also widely contributed to other intensional and non-classical logics (for example, "Synonymity and Extensional Isomorphism" (1969), "Some Problems of the Synonymy of Language Expressions" (1970), "Three-Valued Epistemic Logic" (1986), "Unverifiable and Unfalsifiable Empirical Propositions" (1988)).

One of the most provocative papers in the whole Cmorej's output – namely "The Development of Concepts" (1990) and "Unhappy Consequences of the Developmental Understanding of Concepts" (1990) – deal with the notion of concept. By developing a series of ingenious arguments, he demonstrated that concepts are static abstract entities that do not undertake any changes. Debates over this issue belong to the most interesting in Slovak logic and analytic philosophy.

An interesting part of Cmorej's research concerns some of the key figures in logic and analytic philosophy. He addressed the philosophy of early Ludwig Wittgenstein in the studies "Objects and States of Affairs in the Metaphysics of Wittgenstein's *Tractatus*" (1989) and "On the First Sentences of Wittgenstein's *Tractatus*" (1999). He discussed some aspects of Frege's work in the article "On Frege's Theory of Concepts" (2000), and elaborated on Gödel's philosophy in "Gödel's Proof of the Existence of God and the Set-Theoretical Understanding of Properties" (1996). The results of his extensive research into the history of logic, scientific methodology, and

analytic philosophy in Slovakia were mainly published in two edited books in 2002 and 2003.

Through his research and teaching activities, which Pavel Cmorej successfully pursued for more than half a century, he shaped several philosophical fields, including philosophical logic, philosophy of language, semantic analysis of natural language, analytic metaphysics, and formal epistemology. On this basis, he managed to build his own school of analytic philosophy. He influenced the academic careers of several of his students and his work inspired many of his colleagues. As a brilliant debater, he saw discussion as a valuable tool for developing philosophical thought. He enjoyed engaging in polemics on diverse philosophical and logical topics, many of which drew on his original philosophical insights. Without his work, Slovak philosophy would be greatly impoverished.

Not only as a remarkable thinker, but also as a person characterized by valuable human qualities, Pavel Cmorej deserves lasting respect and admiration.

Marián Zouhar

Organon F 32 (1) 2025: 111–113 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32107

REPORT

David Miller
(1942 – 2024)

David Miller, a preeminent philosopher of science, logician and one of the foremost proponents of critical rationalism, passed away on 20 November 2024, at the age of 82. His academic journey began at the University of Cambridge, after which he pursued studies in Logic and Scientific Method at the London School of Economics. There, he became one of Karl Popper's research assistants – a pivotal relationship that deeply influenced his future philosophical work. Unlike some of Popper's other students and colleagues, such as Imre Lakatos, Alan Musgrave, and Joseph Agassi, who sought to moderate the radical anti-inductivism and falsificationism inherent in critical rationalism, Miller fully embraced and even extended its radical tenets. Miller's critique of justificationism significantly advanced the Popperian agenda by emphasizing that knowledge cannot and need not be justified and must remain constantly open to criticism. In this way, Miller further developed Popper's robust philosophical stance against traditional epistemology and reinforced the non-foundationalist nature of critical rationalism.

Miller is best known for his seminal work, *Critical Rationalism: A Restatement and Defence* (1994). In this book, he addressed and dismantled most common objections against critical rationalism, and formulated the key principles of his version of critical rationalism with exceptional clarity – let us mention for instance the famous thesis asserting the divorce of reason from good reasons, stating that rationality not only does not depend on good reasons but is contrary to them; the delineation of the domain within which reason operates as testing the consequences of our conjectures not in order to consolidate our knowledge, but to liquidate it; or the description of justificationism as an addiction we need to get rid of by undergoing detoxification. By defining reason exclusively in negative terms, he

brought renewed vigour to Popperian thought. While his work resolutely upheld the critical and negative definition of reason, he also confronted challenges within Popper's framework. Notably, he identified flaws in Popper's theory of verisimilitude, which aimed to demonstrate scientific progress as ever-closer approximations to the truth. And although he endeavored to develop an alternative theory, he candidly acknowledged that it left unresolved the challenge of empirically testing verisimilitude appraisals.

David Miller engaged with numerous fundamental issues in logic and the philosophy of science that he further elaborated in his second book *Out of Error. Further Essays on Critical Rationalism* (2006). During his career his research addressed, among others, the problem of the aim of science and its rational control, the problem of the empirical basis in science, the challenge of rational decision-making (including the so-called pragmatic problem of induction), the role of arguments in cognition, and various interpretations of probability (including the propensity interpretation). His work also dealt with the problem of inductive probability that culminated in the renowned Popper-Miller Theorem. In less technical terms and under the evocative title "Being an Absolute Sceptic" Miller lists the following ways in which critical rationalism could beneficially influence science: scientists should stop exaggerating the power of scientific rationality and likewise, the public should moderate their expectations of what can be accomplished in science. Further, scientists should be less harsh on their colleagues who make interesting but false conjectures and be more ready to admit ignorance. If scientists (as well as philosophers of science) cease to attribute to well-tested hypotheses a security or reliability that they do not possess yet stop short of a slide into relativism – then it is possible to pursue science as a rational enterprise par excellence. An inspiring vision, indeed.

Both authors of this obituary remember David Miller as a colleague and a friend. I, Zuzana, met Popper in the 90ies. I first came to the University of Warwick to discuss Popper's philosophy with David, holding the naive belief that some "good reasons" were necessary for theory assessment and theory choice. His curt rejection of this assumption was intimidating – more so, even, than Popper's own criticism when I visited him in Kenley. However, after further reading and reflection, and through ongoing communication with David, who displayed remarkable patience and a genuine willingness to help

me grasp critical rationalism, I found myself drawn to the "radical side" of this philosophy. Over the years, we became good friends, meeting often in Prague or the UK. David profoundly enriched my philosophical development, and I will always cherish his friendship.

I, Miloš, knew David, and will always remember him that way, as an incredibly kind and patient listener who was at the same time a tough critic with a specific sense of humour. This manifested itself not only in expert debates, but also in ordinary conversational situations. So, for example, on one of my early musings about how corroboration might be an indicator of verisimilitude, he laconically remarked that it "is just wishful thinking", which made me realize the fallacy of my reasoning. And to my confession that I am so used to the mountains that I cannot imagine living on the plain, he replied that I have a poor imagination. David was, in short, always substantive, witty and critical, even of his own opinions or attitudes. I am extremely grateful to have had the good fortune to know him and learn from him.

Thank you, David.

Zuzana Parusniková
Miloš Taliga

Organon F 32 (1) 2025: 114 ISSN 2585-7150 (online)
https://doi.org/10.31577/orgf.2025.32108

ADDENDUM

Addendum to "Is Extensible Markup Language Perspectivist?"

In Tambassi, T. (2024). Is Extensible Markup Language Perspectivist? *Organon F*, 31(4): 399–410. https://doi.org/10.31577/orgf.2024.31404 the following information is missing:

Funding. This paper has received funding from the European Research Council under the European Union Horizon Europe Research and Innovation Programme (GA no. 101041596 ERC—PolyphonicPhilosophy).

Disclaimer. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.