European Journal of Mental Health 14 (2019) 109–142
https://doi.org/10.5708/EJMH.14.2019.1.6

Ferenc Mújdricza*

THE ROOTS OF TRUST**

(Received: 21 January 2019; accepted: 25 March 2019)

The present article addresses the question of the 'roots' of trust: a debate between cognitive and non-cognitive trust theories, ongoing since the dawn of modern theorising on trust. On the one side, there is the assumption of conceiving trust as a learnt capacity, based on Erikson's concept of basic trust. On the other side is the hypothesis of innate, built-in trust. After a critical overview of the cognitive and non-cognitive approaches, given that neither side was able to build up a decisive argument, the paper proceeds to some relevant discoveries of the life sciences that serve as proofs of the concept. Michael Polányi's principles of marginal control and boundary conditions help avoiding the pitfall of any reductionist determinism. The analysis results in a rejection of the early learning concept of the cognitive approaches. Trusting is proven to be an *a priori* given human faculty inscribed in our neurobiological system, but neither biologically, nor in any other way, entirely determined. The possibility to trust is always present in the human: the trusting being.

Keywords: cognitive trust; non-cognitive trust; basic trust, innate trust, Michael Polányi

1. Introduction

Trust is without doubt one of the core determinants of individual as well as of social existence, integrity, and (mental) health. The fact that the human is a *trusting being* is among the predominant and, at the same time, most peculiar constitutive factors of human sociality. Theorising trust has become a true 'hot topic' of the social sciences

* Ferenc Mújdricza, Semmelweis University, Károly Rácz School of PhD Studies: Mental Health Sciences, H-1428 Budapest, Pf. 2., Hungary/Ungarn; Hungarian Central Statistical Office, Methodology Department, Survey and Data Preparation Methodology Section, H-1525 Budapest, Pf. 51., Hungary/Ungarn; Ferenc.Mujdricza@ksh.hu.

** I wish to express my gratitude to my PhD supervisor, Prof. Dr. J. Endre Nagy for his overall guidance and vital insights not only with regards to this paper, but to the entire doctoral research as well, whose subtopic is the present article (e.g. incorporating Tillich and Polányi, pointing out the ontological nature of Weil's claim, etc.). I wish to thank to my fellow PhD candidate, Mónika Gubányi for drawing my attention to the relevance of oxytocin regarding my idea of trust and to Zoltán Kacsuk, my colleague at the Hungarian Central Statistical Office, for his suggestions on how to address some of the requests of the reviewers. The anonymous reviewers of the final and earlier versions of the manuscript have also contributed greatly to the paper as well as to my research in general with their comments and requests, therefore they also deserve acknowledgements.

since Niklas Luhmann's groundbreaking work *Trust and Power* (1979). Despite its popularity, there are still many serious contradictions and unresolved issues. The present paper addresses one of the fundamental issues with the theories of (social) trust.

The vast literature on trust can be categorised by various aspects (Delhey & Newton 2003; Örkény & Székelyi 2009). The categorisation in two major 'schools' is not uncommon, though; and it serves the purposes of the present article at best. Namely, categorising into cognitive and non-cognitive groups of theory (Becker 1996, Balázs 2008) that are more or less analogous to the strategic and moralistic types of trust as described by Eric M. Uslaner (2002). The 'conceptual confusion' (Lewis & Weigert 1985, 975; Misztal 1996, 13.) of trust cannot only be derived from the highly increased interest since Luhmann's work (1979) was published. Most probably it stems from the unusual 'rootlessness' of the topic. This rootlessness amounts to a relative absence of the classical philosophical background. Surprising it may be but it is still a fact: the great thinkers and philosophers of the past have rarely discussed the phenomenon of trust in a systematic and thorough fashion and mostly sporadic, statement-like remarks can be found. To provide but an incomplete list, Hesiod, Plato, Aristotle, Cicero, Seneca, or Confucius from the ancient times, the Christian scholar Thomas Aquinas, Machiavelli of the Italian Renaissance, anglophone thinkers like Hobbes, Locke, Hume, Burke, and John Stuart Mill as well as German philosophers such as Hegel, Nietzsche, and Simmel, all paid more or less attention to the importance of trust in various contexts[1]. Although their remarks on trust are not proper conceptualisations of the phenomenon, the recurrence of the topic can be traced throughout the ages.

The aim of the present article is to provide the reader with a possible solution to a problem responsible for one of the most fundamental differences between the cognitive and the non-cognitive theories of trust. This problem is a debate that is yet unsolved between the assumption of trust as a learnt capacity versus the concept of trust that is an in-built, innate component of human personality (in its individual as well as in its social sense). This work is only a small but crucial part of a larger-scale endeavour unto a novel understanding of the notion and concept of trust the appropriate elaboration of which would greatly exceed the frames of a journal article in length as well as in format. Therefore, further details will have to be published in future papers.

1.1. Methodological remarks

Due to the theoretical focus, the present study applies pure systematic theoretical argumentation. As for the collection of the literature discussed in the paper, a mixed approach was used following the suggestions of Greenhalgh and Peacock (2005)

[1] The discussion of their views is not the aim of the present paper. Grant Duncan (2019) gives a thorough account on many of them in his recent book.

and RANDOLPH (2009) on the priority of 'traditional' offline methods, treating online indexing services and searches (Web of Science, Google Scholar) as auxiliary, complementary means in the efforts unto 'saturation'[2] (RANDOLPH 2009). The analysis of the literature was done simultaneously with the collection process. The analytic method is close to Grounded Theory approach:[3] going back and forth between data collection and the levels of analysis, constantly comparing the results, categories, inferences, etc. to the contents of old and newly found texts.

2. Theories of trust

This chapter gives a brief overview of the two major branches of trust theories in order to show the background to the reader, i.e. the context in which the subject of the present article can be situated. The aim of this paper is not to give a thorough literature review, because an adequately detailed description and assessment would far exceed the frames of a journal article. Hence, the paper provides the most important characteristics, assumptions, and critiques of each school within the text. For the more interested readers, two tables were created (one for each branch) where they can find further important details.

2.1. The cognitive theories

Despite the severe critiques on its shortcomings, the cognitive / strategic approach still holds a firm influence to date. It is perhaps due to a fallacy that it would be easier to operationalise and empirically analyse in contrast with the non-cognitive / moralistic approaches. Trust as described by these concepts generally assumes uncertainty and risk, knowledge- and experience-based, (object-)person-bound rational decision or expectation, which is extremely fragile (USLANER 2002). Due to the radical differences between the major cognitive schools that tend to sharply criticise each other, it is better to discuss the different schools one by one, showing how they build and react on each other instead of attempting to group them together by the analytic aspects in *Table 1*. For a summarising view, see *Table 1* below. However, the major shortcomings of the 'softcore' cognitivist schools are presented separately, because that is the terminal, that is to say, the final destination where the

[2] The term 'saturation' comes from the (qualitative) methodology of Grounded Theory. It refers in general to a point in the data collection after which no new evidence comes to light (GLASER & STRAUSS 1967). Although it sounds rather tempting, actual saturation can hardly ever be reached for reasons beyond the scope of the present article, still, one has to strive for getting close to it (CHARMAZ 2006).

[3] As it can be seen, the analytic approach differs from the approach of Grounded Theory due to the fact that predefined as well as newly discovered categories were used because of the generally unconvincing positivist assumption of Grounded Theory. Namely, that any theoretical construction could be built up based solely on and thus 'emerging from' the data by ruling out (or bracketing) previous knowledge (tacit or not) on the subject. The present paper does not delve further into this – otherwise highly interesting – methodological debate, of which MÚJDRICZA and FÖLDVÁRI (2018) have recently given an overview.

rationalising cognitivist approach was able to get to by its failed attempt on incorporating the irrationality of trust in its argument.

2.1.1. The early rationalists

In the less refined concepts adopting the simple economic approach of Rational Choice Theory (Dasgupta 1988; Good 1988; Coleman 1990; Offe 1999), trust appears as reasonable calculation, rational choice, assessment of costs and gains based on pure cognition (Marková 2004) and past experiences – a variety of uncertainty-reducing, risk-handling means: a substitute for certainty (Balázs 2002). Sztompka (1999), too, can be listed here – even though he also appears among the theorists of a more advanced version of the cognitive approaches – due to his woefully uncritical acceptance of Coleman's (1990) definition of trust as gambling, i.e. placing a bet. The most succinct critique of this approach, as paradoxical as it may be, actually predates them. Luhmann (1979; 1988) had already pointed out that such definition describes nothing but sheer calculation that leaves trust no room. Trust, according to him (Luhmann 1979), can only be perceived as rational from the level of the social system as a phenomenon increasing the complexity on the systemic level but reducing it for the individuals.

2.1.2. Trust as 'encapsulated interest'

From the critique of the simplifying applications of Rational Choice Theory on trust arose a later version of an interest-driven, calculative conception of trust. Russell Hardin (1993; 2006) – and his followers, most notably Cook and Gerbasi (2009) and Kohn (2008) – shifted the focus from rational choice to rational expectation. It entails that trusting would amount to having a reason to believe that being trustworthy in the relevant way at the relevant time will be in the trusted person's interest, and so the interests of the trustor are 'encapsulated' in the interests of the trusted. Although Hardin (1993; 2006) opposes the assumption that trust was a risky choice or a choice at all, his concept still carries 'shady manipulative' (Lagerspetz 2015, 39.) overtones. In this mutually interest-driven approach to trust, the trustor takes advantage of the desire of the trusted for good reputation. Following Philip Pettit's (1995) and Olli Lagerspetz' (2015) arguments, such cunning can be discovered in Hardin's concept that not only the trustor, but the trusted can also become a victim of a shrewd calculation. Thus, this concept distorts the phenomenon of trust into a cunning, manipulative, selfish relationship of interest (i.e. relationship based on interest only), not refraining from beguiling, either. Hardin's huge contribution to trust research notwithstanding, his core conception seems to demonstrate a misunderstanding of its subject. In the words of Olli Lagerspetz (2015, 39.), it can be called a 'mutually distrustful cooperation' at best.

2.1.3. The 'terminal' of the cognitive accounts

Alas, the early theorists seem to have been unaware of two short, but important remarks of a few authors of Social Exchange Theory (Homans 1961; Blau 1964)[4] on trust. Trust has a crucial role in Blau's Social Exchange Theory, for trust is essential for reciprocity, social exchange, and stable social relations, and vice versa. He binds trust to risk-taking in social exchange relationships. In his description, trust emerges gradually out of pure self-interest in the course of recurring cost-benefit calculations in (social) exchange situations. So far, he might be taken as a forerunner of the early cognitivists (their apparent ignorance of Blau notwithstanding). But this exchange is somewhat different from rational, economic exchange, for prices and rewards cannot be exactly measured in social exchange relations. Social exchange, this basis and derivate of trust is calculative, but not a pure calculation of advantages – somewhere in the middle between love and calculation. This intermediacy is one of the important remarks to be found in Social Exchange Theory, for it opens up the possibility of irrational, or at any rate, not entirely rational trusting. Homans (1961, 386.) went significantly further in assuming trust as a 'true belief' that (at least sometimes) prevents the betrayal of others for short-term gains. Furthermore, this trust is supported by a specific form of social capital: a moral code – which idea transcends the cognitive concepts and foreshadows Fukuyama's (1995) and Uslaner's (2002) moralistic trust concept[5].

The 'softcore congitivists', Putnam (1993; 2000), Giddens (1990), Sztompka (1999), and Möllering (2006) diverge from the mainstream of cognitive trust theories. Even Luhmann (1979; 1988) himself refrained from perceiving trust as a simple relationship of interest, this is why he is listed here, in spite of the many self-contradictions[6] in the works of this pioneer of trust conceptualisation. Self-contradictory or not, it is beyond doubt that he was against the simplistic rational choice approach, and took a much softer stance by assuming that trust is only rational on the system-level, on the individual, it is irrational. He labelled trust as a 'risky undertaking' (Luhmann 1979, 26.), an expectation sans calculation and self-interest.

Putnam's (1993; 2000) notion of trust – and Hollis' (1998) concept that further developed it – built upon the idea of 'generalised reciprocity'[7] is a solution far more sensitive to the 'sociality' of trust compared to the 'hardcore' cognitive concepts. It seems to be very much in line with Blau's Social Exchange Theory as well (aside from the lack of reference).

Its main shortcoming is also similar to the concept of Blau's. As Putnam (1993) admits it, being a combination of short-term altruism and long-term self-interest, his

[4] Other notable works on Social Exchange Theory are Thibaut and Kelley's (1959) book and Emerson's (1976) paper, but they do not emphasise trust as Blau (1964) or Homans (1961).

[5] Alas, neither refer to him.

[6] This interesting topic cannot be pursued here.

[7] A term he 'borrowed' most likely from Gouldner (1960) – failing to indicate the reference, unfortunately.

highly popular trust concept still remains experience-based and self-interest-driven. The robust nature and the self-affirming, cumulative character of trust and distrust (HIRSCHMAN 1984; GOOD 1988; DASGUPTA 1988; GAMBETTA 1988) is impossible to explain persuasively without leaving the domain of the cognitive theories. They remain unexplained in Putnam's concept as well. Moreover, it does not provide any explanation on trusting *complete* strangers, as neither of the cognitive approaches do.

Trust, so to say, 'brackets', suspends risks in GIDDENS' (1990) and SZTOMPKA's (1999) approaches. Their reading suggests that not only we *do not count* with the risks: we positively act *as if* they did not exist at all. Although HARDIN (1993) also built the supposed *as if* -characteristic of trusting in his theory, GIDDENS (1990) and SZTOMPKA (1999) put it in the focus – although SZTOMPKA (1999), as noted above, took a major step back by utilising COLEMAN's (1990) 'gambling' approach, supposing that trust was like placing a bet, an assumption that HARDIN (1993) had utterly rejected. It is important to note that neither of these authors mention Milton Friedman. It is hard to assume that FRIEDMAN's (1953) widely known *as if* doctrine attempting to 'fix' some issues of the Rational Choice Theory-based neoclassical economics had no influence on the *as if* assumption in trust theories that was also (how surprising!) an attempt to fix the issues caused by the simplistic Rational Choice Theory-based 'early' cognitive accounts.

2.1.4. The major shortcomings of the most developed cognitive theories

Apart from the fact that this *as if* concept only tells about a *mimesis* (imitation) of trust rather than actual trusting, the failure of the cognitive concepts in grasping the essence of the phenomenon in a rationalising way become ultimately clear at this point. They are bound to let in an extremely irrational element, the so-called 'leap into faith' (GIDDENS 1991, 3, 244.). Following Guido MÖLLERING (2006) – who does a fairly worthy attempt at unifying Giddens', Hardin's and Sztompka's *as if* approach[8] – the nature of this suspension, the 'leap of faith' – Giddens' suggestion

[8] to be precise, the idea can be traced back to LEWIS and WEIGERT (1985) and Bernard WILLIAMS (1988) if we remain within the boundaries of trust literature. Hardin just picked it up and used it – in a slightly incoherent manner, to be true, with his encapsulated interest theory (and he mistakenly referred to GAMBETTA 1988 instead of Williams, too. . .). GIDDENS (1990; 1991), SZTOMPKA (1999), and MÖLLERING (2006) only demonstrated the assumption in a fairly systematic fashion (which statement is the truest with regard to Möllering). Not one of them – Williams included – refers to FRIEDMAN (1953), by the way, a mistake that is intolerable on other fields of science. But such mistakes are not uncommon among the theorists of trust. Williams failed to provide reference on Lewis and Weigert, Giddens on LUHMANN (1979; 1988) and LEWIS and WEIGERT (1985), PUTNAM (1993; 2000) on GOULDNER (1960) and BLAU (1964), YAMAGISHI and YAMAGISHI (1994) on BAIER (1986) and DUNN (1988), as well as MAYER and colleagues (1995), ROUSSEAU and colleagues (1998), and others following them since on Baier, and the list could go on and on. It is not up to the author of this paper to judge whether this pattern is due to sheer negligence, ignorance, or less excusable reasons, but at least a considerable share of the aforementioned 'conceptual confusion' of trust might indeed have been sorted out, were such mistakes avoided. We have to give credit, though, to MÖLLERING (2006) in that he does discuss BLAU's (1964) concept, unlike so many other cognitive trust theorists.

was adjusted slightly to fit to KIERKEGAARD's (1983) well-known term[9] – became the most important question for (cognitive) trust research. The problem with this approach is, on the one hand, that although simple calculation disappears by the suspension and bracketing of risks and faith would fill its place, faith, according to TILLICH (2009), cannot be the complementary element of incomplete knowledge. On the other hand, suspension and bracketing of risks, perils, and vulnerability amounts to sheer ignorance or negligence – naïveté rather than trust.[10]

Besides the faith and the naïveté issue, the leap of faith – or hope (LI 2015) – can be taken as a so-called 'black box' (again, it is suspiciously similar to where behavioural economics based on Friedman's *as if* doctrine ended up; BERG & GIGERENZER 2010). It is rather important to emphasise this shortcoming: as this is the furthest the conceptualisation of trust with a cognitive approach got to (already suspiciously irrational), it sheds light on the ultimate mistake of the cognitive school. Certain inputs (information, experiences, perils, risks, vulnerabilities, etc.) are assumed to be processed by this black box and – no one really knows how – trust would come out on its other side. Such black boxes as this are usually unwelcome in sound conceptualisation efforts (BUNGE 1979). Proper theoretical deductions should be free from any 'leaps' or 'holes' in the argumentation, at least if a simpler solution is at hand for the issue in question. And it is just the case with the 'inputs' for the supposed leap of faith of trust. The notion of reliance, which the related critiques (BAIER 1986; BALÁZS 2002; HERTZBERG 1988; 2010, BERNSTEIN 2011; LAGERSPETZ 2015) thoroughly discuss, does not require a black box (BAIER 1986) to form out of these inputs. Their definition of reliance can be more or less identified with the calculative, rationalising cognitive / strategic notion of trust, consequently, any – not only the 'leap of faith-ist' – cognitive account of trust mistakes reliance for trust.

Reliance (therefore the cognitive trust definitions also, or in USLANER's (2002) reading, strategic trust) assumes a three-place, purposive relationship, formalised as 'A relies on B to φ' (LAGERSPETZ 2015, 16.; see also USLANER 2002, 21). In contrast, BAIER (1986) got rid of purposiveness in trust, even though she still saw trust as a three-place relationship.[11] However, perhaps not explicitly following this move of Baier's, but certainly in line with it, the true non-cognitive trust theorists – take for example HERTZBERG (1988), BECKER (1996), or LAGERSPETZ – see trust only two-

[9] The first to introduce the assumption of a cognitive 'leap' in trusting was, quite unsurprisingly, LUHMANN himself (1979). He did not use the term faith yet – his famous phrase is 'leap into uncertainty' (p. 33.), wherefor information about the trusted 'simply serve as a springboard' (p. 33). While LEWIS and WEIGERT (1985) refer to him, Giddens, again, fails to provide the reader with reference on either Luhmann or Lewis and Weigert, even though the similarity between 'leap into uncertainty' and 'leap into faith' is rather unmistakable.

[10] Moreover, basing the concept of trust on faith runs the risk of circularity if we remain on the level of the everyday usage of these notions, given the overlapping meanings and use of the two words in the English language (DUNCAN 2019).

[11] Where purposiveness is shaken off by shifting the focus on entrusting a valued thing to someone – a motion later to be called in question, too (JONES 1996).

place: 'A trusts B' (2015, 16).[12] Although Lagerspetz' (2015) critique is quite sharp, his argumentation seems flawless: any reason for trusting must fall short of a justification of trust, for were trust a kind of risk-taking agency in a state of uncertainty, it would either be redundant (if we already have assurance pertaining to the given situation) or ill-advised (if uncertainty reigns). The inquiry attempting to find reason for trusting itself 'assures that my attitude can at most amount to reliance' (Lagerspetz 2015, 134). According to these, the rational, interest-based, cognitive / strategic concepts of trust, as a matter of fact, might not even speak *about* trust (Becker 1996; Hertzberg 2010; Lagerspetz 2015). Rather, in certain cases, they assume a manipulative and / or suspicious frame of mind that is, actually, the direct opposite of trust (Lagerspetz 2015). The bottom line of Lagerspetz' (2015) critique (based on Jones' (1996) account) is that since trust is with us from the start as a background of the interpretation of the experienced facts and information, it cannot be *based on* them.

To sum up: any 'reasonable', rational, calculative, or interest-based 'trust' definition speaks instead about distrust, manipulation, or even in the best case, reliance. *As if* trusting leads to a sheer mimesis, imitation of trust. Following the path set by the 'leap of faith' paradigm and leaving the 'black box' of faith, bracketing, and suspension untouched, naïveté emerges; if the black box is cut out with Occam's razor, only a refined description of reliance is left – not so much more than any other theory of the cognitive approaches could provide us with.

By introducing these critiques, we have already arrived at the domain of the non-cognitive concepts. Before proceeding, though, I need to note: should we accept the 'leap of faith' in trust, faith would become a key element of the 'genesis' of trust. An element, which, being irrational, is unexplainable in a rationalising way. Remember: the original goal was the *rational* explanation of trust on cognitive grounds, yet they admit, at least, that trusting can be viewed as irrational (Möllering 2006) – arriving to the problem that Blau's (1964) remark had foreshadowed decades earlier. As failed as the cognitive approaches are, the assumption of the 'leap of faith' indicates a significant move towards the non-cognitive theories, whose starting point is similar to this view.

2.2. The non-cognitive theories

The non-cognitive concepts (Baier 1986; Govier 1993; Becker 1996; Jones 1996; Bernstein 2011, Lagerspetz 1998; 2015) rooted deeply in Moral Philosophy challenge the less phenomenologically focused cognitive concepts from the start. While the cognitive theories seem to strive for getting a grip on 'everyday trust', not delving deeper into phenomenological / philosophical arguments, the masterminds of the

[12] Uslaner moves another step further in his summarising work – his 'moralistic' trust (which is more or less similar to the core trust concept of the non-cognitive accounts) is only one-place: '*A* trusts' (Uslaner 2002, 21. Italics in the original).

Table 1

	Cognitive theories of trust				
	Early (hardcore) rationalists	*Encapsulated interest*	*Softcore cognitivists*		
Short description	Simple, crude concepts based mainly on the economic approach of Rational Choice Theory	a more refined economic approach	transitions or intermediaries between the cognitive and non-cognitive accounts		
Key authors	P. Dasgupta, D. Good, J.S. Coleman, P. Sztompka	Hardin and his followers (K. Cook, A. Gerbasi, M. Kohn)	N. Luhmann	R. D. Putnam, M. Hollis	A. Giddens, P. Sztompka, G. Möllering
Origins	Rational Choice Theory, Game Theory, Axelrod's experiments	Criticism of the simplistic Rational Choice Theory-based early cognitivists	His own System Theory, & Simmel, Parsons, etc.	Blau's Social Exchange Theory, Gouldner's generalised moral norm of reciprocity (based on Malinowski's observations)	Luhmann, Friedman (*as if*) Lewis & Weigert (*as if & leap*), Williams (*as if*), Hardin (*as if*), Kierkegaard (*leap of faith*)
Conceptualisation	**risk-taking, rational, strategic choice/decision**, a calculative substitute for certainty = a means of handling relational uncertainty by an assessment of costs and gains.	**rational expectation**: having a reason to believe that being trustworthy in the relevant way at the relevant time will be in the trusted person's interest - trustors interests are 'encapsulated' in the trusted's interests	**risky undertaking'**, expectation, but not a calculative and/or self-interest-based relationship: **a reduction of complexity**. Irrational on the level of the individual, rational on the level of the social system.	(Thin) trust is based on the **generalised reciprocity**, a continuous relationship of exchange: combination of short-term altruism/solidarity and long-term interests. Irrational on the individual level, rational on the social level.	Free of calculation, bracketing/suspension of risks/threats/uncertainty (acting *as if* they not existed). Trust implies a **'leap into/of faith'**: a leap into uncertainty/unknown. It 'transmutes' experiences and information into trust.
Roots of trust: key assumptions	Trust is rooted in past experiences collected through iterative interactions with individuals or groups	Experience, evidence- and interest-based, maintained by the desire to continue the relationship (financial interests, emotional ties, reputational effects, etc.). Parents' early year investments form trust.	Based on familiarity and evidence from the past. Self-contradictory: based on early childhood experiences vs. innate.	Based on early childhood experiences. Altruism/solidarity, 'rightly understood' self-interest, shared social networks and expectations of reciprocity.	Trust is not always rational. (Past) experiences, information, and evidence on the trusted individual or group are the bases of the leap
Key elements	Descriptions of 'everyday trust' Cognitive, conscious action Unstable, fragile Hard to create, easy to destroy Three-place relationship: 'A trusts B to ϕ' Strategic agency in a sense Based on experiences/information/knowledge on the trusted				
	Rational		Can be irrational		
Shortcomings	Describing 'everyday' trust only No explanation on the core/essence of trusting No adequate explanation on unconscious trust Mistaking trust for calculation/manipulation/reliance/naïveté No clear distinction among reliance, trust, and naïveté No explanation on trust towards complete strangers No satisfactory explanation on the robust, self-affirming nature of trust				
	Overly simplifying, outdated, dehumanising *homo oeconomicus* approach Sheer calculation, leaves trust no room If trust is a risk-taking agency, it is either redundant or ill-advised Describes *reliance* rather than trust No explanation on irrational trust	Calculative, interest-driven *homo oeconomicus* approach Mistaking trust for a (cunning, manipulative, beguile, selfish) mutually distrustful 'cooperation'. No explanation on irrational trust	burdened with self-contradictions, i.e. trust is innate vs. learned in early childhood	Less harshly, but still interest-driven (with all of its disadvantages) Incomplete conceptualisation of the phenomenon	Introducing a 'black box' - a token of failure in conceptualisation (crucial hole in the theory) Faith cannot complement incomplete knowledge Mistaking trust to a mimesis of trust (*as if*) Describes naïveté rather than trust Without the black box/mimesis, it is only a refined description of reliance

non-cognitive concepts take a completely different approach, supported by firm argumentative powers of their philosophical background. They address the issue of the definition of trust by approaching it on a more general as well as 'deeper' philosophical level. Reaching to the core, the very essence of trusting on the grounds of Moral Philosophy is their ultimate goal (except for Lagerspetz and the Wittgensteinians). In their endeavour to define trust, they even sacrifice the rationality of it. Although this different stance might be taken as the reason for the irreconcilable discrepancies between the two aspects, it is against intuition that any general, essential, core notion of a phenomenon would be so much different from its everyday form. To quote Philippe Rochat (2010, 32): 'the basic requirements or mechanisms underlying trust are basically the same: the same invariant and universal properties underlying the phenomenon, however filtered they might be in their expression at various levels of content and complexities.' They should not be so much in contrast with each other – rather, definition(s) of everyday trust should at least be *possible* to derive from a general non-cognitive idea of it. Conversely, the non-cognitive theories culminate in the rejection of cognitive or rational trusting.

Although different 'schools' of non-cognitive trust theories can be discerned, it is better to discuss them together, given the significantly less deviations and the more pronounced common elements in contrast with the diverse cognitive theories.

2.2.1. Key authors and schools – and the less significant

Systematic non-cognitive trust theory is rooted in Annette Baier's *Trust and Antitrust* (1986) – an article inspired greatly by Sissela Bok's (1978) considerations on trust. Still, a long conceptual 'journey' leads from this groundbreaking work to the 'fully-fledged' non-cognitivism of Jay M. Bernstein and, even more importantly, Olli Lagerspetz. Trudi Govier and Karen Jones, who treated trust as an (affective) attitude, were still more of forerunners than true non-cognitivists, incorporating cognitive elements in their theoretic approach alongside the non-cognitive ones, but the paramount importance of their contribution is beyond doubt. Lawrence C. Becker is the first among the key authors to be considered as openly non-cognitivist. There are important common elements as well as noteworthy differences in the non-cognitive trust literature, the following paragraphs and *Table 2* summarise the most significant.

Some prominent names that might come to mind (Dunn, Hertzberg, Yamagishi & Yamagishi, Fukuyama, Uslaner, etc.) are not listed among the 'key' authors. This is due to the fact that these authors cannot be taken as theory-builders on their own account, rather, they provided inspiration to others, or summarised the work of others with few original additions. This is not to say that their works were not of exceptional quality or importance. For instance, Dunn's (1988) view of trust as a passion introduces the assumption of trust as an unconscious, not strategic / calculative, confident expectation of benign intentions in a free agent. But his account is slightly controversial, for he views trust as a two-sided phenomenon, which can also be a strategic,

consciously chosen policy for handling the freedom of other agents. Still, the assumption of trust as a passion inspired many to come after him. HERTZBERG (1988; 2010) laid the foundations for Lagerspetz (both Wittgensteinian scholars), the latter elaborated on the bases of the former's work[13]. DUNCAN (2019) builds on the work of Hertzberg and Lagerspetz in his excellent book, utilising, complementing, and furthering their ethical understanding of trust by an especially thorough 'observation' of the uses of trust in the works of some of the greatest theorists over time. Toshio and Midori Yamagishi made good use of the insights of some of the non-cognitive forerunners (e.g. Baier and Dunn, alas, without references) in their definition of general trust as 'an expectation of goodwill and benign intent' (YAMAGISHI & YAMAGISHI 1994, 131–32), and they also mention the cognitive bias trust generates in the evaluation of others. However, Toshio YAMAGISHI (2011) later opted for information-based trust in his more recent book on trust, thus turning towards the cognitive theories. It goes without saying that trust viewed as '*complex information* processing' (YAMAGHISHI 2011, 23. Italics in the original, emphasis removed) is rather inconsistent with the non-cognitive theses. FUKUYAMA (1995) took an apparently non-cognitivist stance arguing against the assumption that the moral community that gives rise to trust would be possible to acquire through rational investment decisions. It is quite important, though, that he – under the unreferenced influence of HOMANS (1961), perhaps – shifted the emphasis on the sociality of trust, and on the common moral norms and ethical habit trust is based on. Others, like USLANER (2002) made good use of Fukuyama's motion. But neither Yamagishi and Yamagishi nor Fukuyama delve deeper into the phenomenology of trust, hence their exclusion from the key authors. USLANER (2002) has provided us with one of the most comprehensive summaries on the two major schools, giving us the notions of strategic and moralistic as well as generalised and particularised trust. His exceptional summary has but only a few original assumptions – perhaps the most significant of them is the formalisation of (moralistic) trust only as a one-place 'relationship', without any inherent need for subject or intention: '*A* trusts' (USLANER 2002, 21. Italics in the original). This assumption is of paramount importance that I will address in a future paper.

2.2.2. Key elements of the non-cognitive theories

The non-cognitive branch of theories initially treated trust as a (mental) state saturated with a certain *affective warmth* (BAIER 1986; BECKER 1996; JONES 1996; BERNSTEIN

[13] Given that philosophers drawing on Wittgenstein's philosophy generally dislike the idea of developing philosophical theories, the lack of a 'proper' theory is not to be counted as a shortcoming in the case of Hertzberg. The aim of Hertzberg and Lagerspetz was to give an account of how the notion of trust is used in ethical life and how it may solve or create problems in that life. In this specific case, however, Lagerspetz had such essential insights inspired by Simone Weil that seems to have taken him beyond the fixed boundaries of (Wittgensteinian) ethics. As implicit as this crossing of borders may be, it is nevertheless palpable, as it will be shown below.

2011), in opposition with the bleak rationalism of Game Theorists. With the Wittgensteinians, Hertzberg and Lagerspetz, however, the non-cognitivist account moved away from emotional characterisation, not by turning back to the (f)rigid rationalism of the cognitivists but by extending the phenomenon beyond, above and below any specific feeling or state of mind. As DUNCAN (2019) summarises, instead of striving for a conceptualisation, Hertzberg and Lagerspetz emphasised and studied the role of trust in human interaction. Paraphrasing Wittgenstein, LAGERSPETZ (2015, 95) argued for the omnipresence of trust: 'a pattern in the weave of life'. Consequently, two major non-cognitive approaches can be outlined based on their focus: one focusing on emotions and the other putting the emphasis on the mainly irreflective, unconscious character of trusting, treating it as a 'business-as-usual' attitude in many cases. It should be noted, though, that this is the result of a shift in emphasis, not an absolute opposition or rejection of the other aspect. The authors viewing trust as an affective attitude also recognise that trustors are usually unaware of their trusting, and Lagerspetz does not deny the possibility of affective trusting either, he just does not allow for *limiting* it to feelings, emotions, or any specific psychological state. However, LAGERSPETZ (2015, 157), too, admits in the concluding chapter of his book that the notion of trust is 'closely associated with friendship and love' – rather warm affections, indeed.

Although the non-cognitive theories are thus not less heterogeneous than the cognitive ones, there are some common elements in the concepts of the forerunners (i.e. Baier, Govier, Jones) as well as of the actual non-cognitivists (i.e. Becker, Bernstein, Lagerspetz). Given that most of them have a strong background in (moral) philosophy, they are less concerned with the 'everyday', superficial perception of trust, thus avoiding the trap the cognitive accounts seem to have fallen into. In general, their efforts enable fathoming the very essence of the phenomenon. Even though the Wittgensteinians should not be 'accused of' having essentialist intents, Lagerspetz' reading carries crucially important general insights – at least tacitly – even though he treats trust as a situative, empirically unobservable notion. A specific common conceptual element is that all of them see the trustor as someone who assumes the benevolence / goodwill / benignity of the trusted. They agree that trust is something that cannot be willed (although JONES (1996) seems a bit hesitant on the matter, see her footnote 13), and they all argue against the cognitivist conflation of trust and reliance, arguing for their clear distinction.

The problem of the robust nature of trust is also a recurring topic, to which the non-cognitive theorists found an elegant solution. JONES (1996) introduced the assumption that trust acts as an interpretive filter, making it highly resistant to information and counterevidence. Their interpretation is fundamentally affected by the general trusting / distrusting disposition. BECKER (1996, 60) made the same assumption implicitly asserting that trust is not damaged by lies or mistakes 'insofar as I construe them as well meant or innocent'. Bernstein and Lagerspetz further developed Jones' assumption. BERNSTEIN (2011) suggested that the interpretive filter of trust could block or distort any information or evidence that might pose a threat to trust.

LAGERSPETZ (2015) solved Jones' issue with the 'start-up' problem of trust[14] by shifting the view from the observers' third-person perspective to the trustor's first-person perspective – the only real view that shows trust in its actuality. Lagerspetz asserts that trust is the very *background* of interpretation, and this is the reason why it drives and even changes the perception and apprehension of facts and situations, which is a characteristic that makes us prone to bias in favour of the trusted. Thus, '[t]rust (. . .) is never *forced* on us by the facts alone. We must trust *at first*, in order then to give facts the interpretations that we do' (LAGERSPETZ 2015, 86. Italics added).

Of the non-cognitive accounts, only Annette BAIER (1986, 260) referred to trust as 'a fragile plant', but she connected this fragility to conscious trust, asserting that trust rarely endures awareness. As for the awareness of trust, BAIER (1986) thought that trust can be conscious as well as unconscious: she compared it to air – we (mostly) notice it 'when it becomes scarce or polluted' (p. 234). According to Karen JONES (1996), trust can appear even if one has all the reasons not to trust. One only becomes aware of her trusting attitude when she attempts to fight against it on the ground of the available information. Until then, trust remains unnoticed. BERNSTEIN (2011) and LAGERSPETZ (2015) took Baier's assumption of awareness destroying trust even further, developing – as Lagerspetz put it – the 'dys-appearance' doctrine of trust: trust is invisible, unperceivable, and appears only posthumously, when it is already disappeared, broken, destructed – like health. That is to say, only the *absence* of trust is consciously realised.

However, the recognition of Annette Baier's concept in general is ambiguous (LAGERSPETZ 2015). HERTZBERG (2010), for instance, sorted it right among the cognitive trust concepts. At any rate, her work signifies a transition from the cognitive towards the non-cognitive approaches – she was a true a 'forerunner' of the latter. Baier was the first to explicitly suggest that trusting attitude is driven by goodwill, and thus she transposed the emphasis on the intimate, 'warm', emotive aspects of trust (BAIER 1986; LAGERSPETZ 2015). Her definition of trust (1986, 235) as 'accepted vulnerability to another's possible but not expected ill will (or lack of good will)' seems to be mistaken though, despite the fact that the element of vulnerability has become remarkably popular since then, mainly in cognitive trust definitions (MAYER et al. 1995; ROUSSEAU et al. 1998).[15] True, though, that accepted vulnerability introduces the factor of risk and danger in the phenomenon of trust, which casts a shadow of game theoretical logic and rational choice perspective on Baier's concept. Although her account is still far from the one of the utter rationalists, LAGERSPETZ (2015) argues that such attitude is to be considered as a suspicious, wary reliance. JONES (1996, 19) also points to the flaw in Baier's concept: it is not

[14] Namely, if trust is based on an interpretation of facts which favours the trusted and trust makes us interpret facts in a certain way, it follows that we must trust in order to trust (LAGERSPETZ 2015). It should be noted that BLAU (1964) also struggled with a somewhat similar start-up problem but from a different aspect, unable to sufficiently demonstrate the gradual, simultaneous emergence of trust and social exchange by assuming their essential interdependence. MÖLLERING (2006), however, accepted Blau's incomplete argument.

[15] Yet these authors 'forgot' listing Baier among their bibliography entries. . .

'sufficiently distinguished' from relying on. LAGERSPETZ' (2015) reasoning is quite right that in their trustful relationships people are not likely to feel more, but rather less vulnerable than in their relationships without trust. Lagerspetz solves the problem by showing a completely different, yet comprehensive aspect: vulnerability is not necessarily something to *avoid*. Rather, in trustful relationships and true communities, we actively seek it through 'freely offering up ourselves to each other' (LAGERSPETZ 2015, 64). Instead of Baier's (reluctantly) *accepted* vulnerability, LAGERSPETZ (2015, 67. Italics in the original) recognises the exposition in trusting as an openness of *shared* vulnerability, which 'in itself *is* a valued piece of goods, perhaps the most valuable ever available'.

In line with this conclusion, trust is characterised by the *absence* of the perception of risks, social uncertainty and vulnerability in BECKER's (1996) 'noncognitive security' theory and the 'full-fledged non-cognitivism' of BERNSTEIN (2011) and LAGERSPETZ (1998; 2015). Trust thus makes one *less* vulnerable than not trusting (LAGERSPETZ 2015). The perception of risks is replaced by a feeling of certainty and security. Despite their indisputable similarity, this approach cannot be identified with the cognitive solution of the 'leap of faith'. Whereas this non-cognitive assumption is characterised by an *absence* of the perception of vulnerability, uncertainty, and risk, the theories linking trust to a 'leap of faith' are characterised by their mere suspension. By the aforementioned shift to the first-person perspective, the Lagerspetzean non-cognitive account of trust – unlike the negligence of the cognitive accounts about this aspect – take into account the subjective dimension of the personal nature of trust. I call it the 'positionality' of trust. According to LAGERSPETZ (1998), the element of risk taking in the act of trust can only be conceived by an external observer. But for the trustor – it is simply not present. (Lagerspetz' idea of trust is discussed in the next chapter, being inseparable from his account on the formation of trust.)

Besides the common ground of the assumption of goodwill / benevolence, some authors include several other elements among the constitutive factors of trust. The core of trust in Karen JONES' (1996) reading is a certain attitude of optimism characterised by the trustor's assumption of goodwill and the competence of the trusted. Lawrence C. BECKER (1996) mentions conscientiousness and reciprocity as an element of the sense of security accompanying trust.

2.2.3. Major shortcomings of the non-cognitive theories

Although the non-cognitive accounts seem to be conceptually better than the cognitive ones, they are not flawless. Some of the shortcomings of the earlier representatives were already addressed by the later non-cognitivist scholars such as Baier's aforementioned insufficient distinction of trust and reliance and Jones' 'start-up problem' of trust. The three-place relationship trust models of Baier, Govier, and Jones are also suspiciously reliance-like. The accounts of Govier and Jones can also be criticised for the inclusion of competence in their model: as BECKER (1996, 57) writes, '[c]ompetence connects with the issue of reliance'.

The common shortcomings of all non-cognitive theories seem to be the following. They fail to provide a satisfactory analysis on 'everyday trust', and cannot sufficiently explain conscious trust either, perhaps because they are too preoccupied with the demonstration of its essentially unconscious nature. Another common shortcoming is that although they took a giant leap forward by properly distinguishing trust and reliance, they did not pay attention to the similarly crucial demand for distinguishing trust from naïveté. From the ethical point of view of Hertzberg and Lagerspetz, the distinction might only depend on the ethical attitude of the person who makes the distinction, that is, who s/he blames for the trustor's failure (alas, they did not delve deeper into this topic). Judging the (failed) trustor's attitude incorrect is to label them naïve, whereas condemning the trustee for their betrayal is to prevent the immediate transformation of trust into naïveté upon its failure. But such a solution for the problem would require a third-person position[16] that seems to be inconsistent with the first-person view that Hertzberg and Lagerspetz – quite correctly – argued for in their descriptions of trust. As for an appropriate first-person perspective understanding of how to distinguish trust from naïveté and thus avoid conflation or relativisation of the two, the non-cognitive accounts have nothing to offer – not unlike the cognitive approaches. Still, by introducing the first-person explanation of trust, they point the direction, at any rate, in which to proceed with the issue of naïveté, as well.

3. Learnt capacity vs. built-in faculty

The issue of the formation of trust arises from the fundamental differences between the cognitive and non-cognitive theories of trust, hence the choice of this categorisation (see the *Introduction* chapter) and their somewhat detailed overview above. The debate revolves around the question of their irreconcilable assumptions: whether trust is a learnt capacity or a 'built-in' human faculty.

3.1. The cognitive approach – trust as a learnt capacity

The majority of the cognitivist authors who paid attention to the formation of trust consider it a learnt capacity based on early infantile and childhood experiences (Luhmann 1979; Hardin 1993; Putnam 2000). The origins of this view can be traced back to the developmental psychologist Erik H. Erikson's (1977) hypothesis of *basic trust*, which fits rather well to the common assumption of the cognitive accounts that trust were a result of past experiences and information on the trusted. In short, Erikson's idea proposed a 'blank slate' account assuming that the most elementary form of trusting would entirely be a result of early infantile experiences, dependent on the quality of the relationship with the primary caregiver. This basic trust would affect the sense of trust and trusting / mistrusting attitude throughout life. Although some

[16] In this regard it is irrelevant if the judging person is the trustor or the trusted him/herself – s/he has to take a third-person perspective in order to make such a judgment.

Table 2

	Non-cognitive theories of trust			
	Forerunner of non-cognitive trust - Accepted vulnerability	*(Affective) attitude*	*Noncognitive security*	*Full-fledged non-cognitivism*
Short description	Transition from the cognitive towards the non-cognitive concepts	Does not rule out cognitive elements, but opposes the cognitive approaches	Clearly oppositing cognitive approaches	Refined, critical non-cognitive theories of trust
Key authors	A. Baier	T. Govier, K. Jones	L.C. Becker	J. M. Bernstein, O. Lagerspetz
Origins	Moral philosophy (Hume, Nietzsche, Bok)	Moral Psychology, Baier, Baker	Luhmann, Dunn, Baier, Baker, Govier	Moral philosophy, Wittgenstein, Hertzberg, Løgstrup, Baier, Jones, Simone Weil
Conceptualisation	Reliance on another's good will, and accepted vulnerability to another's ill will/lack of good will toward one	Belief/optimism about the kindness/benevolence/goodwill (not harming us) and (moral) competence (loyalty, kindness, generosity) of another	A sense of security about others' benevolence, conscientiousness, and reciprocity	Shared vulnerability Tacit expectation of goodness in human encounters Mostly unconscious, 'business-as-usual'
Roots of trust: key assumptions	Innate ('ur-confidence'), but selective	Avoiding to answer the question	Avoiding to answer the question	Inescapable, given ('basic trust')
Key elements	Aimed at grasping the essence of trust Trustor assumes the benevolence/goodwill/benignity of the trusted Distinguishing trust and reliance Trust cannot be willed			
	Trust is like air - we (mostly) notice it when becomes scarce or polluted Three-place relationship: 'A' (en)trusts 'B' with valued thing 'C' It can be conscious as well as unconscious 'Fragile plant' Affective warmth	Affective warmth *Govier:* Two-place relationship, open-ended expectation: 'A' trusts 'B' *Jones:* Three-place relationship - 'A' trusts 'B' within a domain of interaction 'D' Trust acts as an interpretive filter - highly resistant to information (robust nature of trust)	Two-place relationship ('A' trusts 'B') Absence of the perception of social uncertainty, risks, and vulnerability Replaced by a feeling of certainty and security Trust is a disposition Balance of cognitive control and non-cognitive stability Implicitly assumes that trust acts as an interpretive filter Affective warmth	Trust acts as an interpretive filter - highly resistant to information (robust nature of trust) Background for interpretation Trust has an ontological priority compared to interpretations 'Dys-appearance' of trust - only the absence/breakdown of trust is consciously realised *Bernstein:* Affective warmth
Shortcomings	Three-place relationship is problematic: we should trust first to entrust The assumption of trust cannot be willed is problematic because of the central role of entrusting Insufficient distinction of trust and reliance Accepted vulnerability is a 'suspicious, wary reliance' Does not discuss the robust nature of trust	Including competence pushes the definition towards reliance Jones' trust as a three-place relationship amounts to reliance We have to trust in order to trust		
	No true account on 'everyday trust' No (satisfactory) explanation on conscious trust No proper differentiation of trust and naïveté			

of the trust theorists, such as Luhmann and Putnam, ignore ERIKSON's (1977) concept on basic trust (at any rate, they do not refer to it), their views bear suspiciously great resemblance to it. HARDIN (1993; 2006), who emphasises that trust is a learnt capacity based on past experiences and it is maintained by the desire to continue the relationship (financial interests, emotional ties, reputational factors, etc.), presents a rather crude, 'economising' reading of ERIKSON (1977). HARDIN (1993) suggests that the formation of the capacity to trust depends greatly on the parents' early age 'investments' (hence the 'economisation'). He still deserves the credit for referring to Erikson, at any rate, unlike many others.

Putnam and Hardin's stance is quite understandable: they are strong representatives of the cognitive concepts for which the experience-based concept of basic trust comes really useful (as for Luhmann, the next paragraph will show that he was indefinite on the matter). It is rather surprising, though, that USLANER (2002), too, recognises the 'roots' of moralistic trust (which is supposed to summarise the non-cognitive accounts of trust) in early age experiences, whereby an optimistic world outlook and a 'trusting instinct' develops. Reading Uslaner, it is clear that he sympathises with the moralistic (non-cognitive) 'version' of trust, yet he builds his argumentation about its origins upon the grounds of the Eriksonian concept of basic trust. ERIKSON's (1977) theory matches much better with the experience-based strategic (cognitive) approaches. Still, Uslaner suggests that trust would be a result of an early learning process. Should then the goodwill- / benevolence-expecting notion of trust of the non-cognitive concepts be regarded as an (early) experience-based, 'tacit cognitive' phenomenon after all? Uslaner's motion can only be understood by noticing his twofold intention. He tried to argue with the assumption that (generalised) trust would be possible to earn by any 'trustworthy' behaviour and he attempted to provide explanation for the robust nature of trust at the same time. '*Trust must be learned, not earned*' (USLANER 2002, 77. Italics in the original, emphasis removed) – so his catchy conclusion goes. He recognised, though, that by solving one problem he created another: he admitted that his hypothesis on the formation of trust is not a strong claim. The issue that he did not recognise, however, is that by using Erikson's concept to explain the robust nature of generalised, moralistic (i.e. non-cognitive) trust, he set its 'roots' into experiences, as it were. Consequently, in essence, he made trust similar to particularised, strategic (i.e. cognitive) trust. The difference between them (and thus between cognitive and non-cognitive trust) would only be in the *time* of the experiences earned. It follows that the earlier the basis experiences of actual trust were, the more generalised / moralistic / non-cognitive it would be. The two, ostensibly irreconcilable types of trust might not be that different after all, if his hypothesis is correct.

3.2. The non-cognitivist account – trust as a built-in faculty

But the main problem is not yet solved, for the fact that the cognitivists and Uslaner seem to have ignored should not be forgotten: that is, the concept of basic trust has received serious critiques since Erikson put forward his hypothesis. Even LUHMANN

(1979) stated at the beginning of his work that if the exclusive alternatives to trust are chaos and paralysing fear, the capacity to trust is a natural faculty of the human being. With this he suggested the *a priori* given nature of trust, though later, contradicting himself, he also assumed early childhood learning. A partial critique of Erikson's concept is Philippe ROCHAT's[17] more recent (2010) psychological study in which the author opposes the assumption of learnt basic trust. ROCHAT (2010, 37) suggests that trust belongs to 'pre-adapted action systems' of neonates that already exist at birth. Neonates are 'endowed' with it, although this innate trust goes through rapid changes depending on external effects and acquired experiences between ages 0 and 5 (he divides the early development of trust in 6 different phases or levels). Rochat considers trust as a biologically determined (but by no means reflex-like!), inborn, 'built-in' phenomenon but not unchanging and unalterable, of course.

The non-cognitive concepts argue similarly. Annette BAIER declared that a certain innate but fragile trust seems necessary in any creature who is initially nourished by another in contrast with the 'ultra-Hobbist child' (1986, 241) who is so fearful that she rejects even maternal nursing. Zoltán BALÁZS (2002) also argued that trust is present in us from the first moment of our existence. The conclusion of Olli LAGERSPETZ (2015, 86) at the end of his argument on the background-characteristic of trust is that trust 'is with us from the start'. Lagerspetz' assumption is very much in unison with ROCHAT's (2010) implicit trust inscribed in neonates. Alas, other significant figures of the non-cognitivist theories – such as Karen JONES (1996) and Lawrence C. BECKER (1996) – did not discuss the issue of the formation of trust, although BECKER (1996, 51–52) emphasised the importance of this 'line of inquiry', implying that non-cognitive trust might be 'partly hardwired', even assuming the possibility of genetically encoded trust.[18]

However, LAGERSPETZ pays special attention to the issue of basic trust, devoting the entire final chapter of his book *Trust, Ethics and Human Reason* (2015) to this topic. Due to its paramount importance, his concept is reviewed in the subsequent paragraphs in more detail than the ones of others, which are outlined previously in this paper. He starts from the generally accepted premise that 'some form of generalized trust is necessary for human life and even sanity to prevail' (LAGERSPETZ 2015, 131). He goes further, asserting that '[w]e need to believe in the essential continuity of the social life around us, and we must act on the assumption that people generally are not hostile or out to deceive us' (LAGERSPETZ 2015, 131). In general, he provides an alternative argument for the claims that view basic trust essentially groundless: innate, primordial – a self-evident sense. This self-evident sense of trust is a 'natural substratum for the subsequent growth of more reflexive approaches to the world'

[17] Neither Erikson nor Rochat are counted among the key theorists of trust in the present paper for they discuss trust only from the viewpoint of developmental psychology.

[18] The latter assumption has become quite popular since then, one of the most recent studies based on this approach is NISHINA and colleagues' (2018) research conducted under none other than the late trust theorist Toshio Yamagishi's supervision (he passed away on 8. May 2018 – let us remember him by mentioning his last efforts in the service of trust research).

(Lagerspetz 2015, 131). Instead of focusing on the ontological / empirical question of the existence of basic trust, Lagerspetz (2015, 153. Italics in the original) conceptualises 'essential features' of trust, using Simone Weil's and K. E. Løgstrup's works in the demonstration of the concept of basic trust as a possible descriptor of '*the* human ethical condition'.

Lagerspetz questions the conceptions considering trust as a response for some general-level scepticism (Bernstein 2011; Giddens 1990; 1991) based on Wittgenstein's *On Certainty* (1969). As Wittgenstein (1969, §160. Italics in the original) noted with respect to childhood learning: '[d]oubt comes *after* belief', so concludes Lagerspetz (2015) – summarising Wittgenstein's view on the matter – that the faculty to trust should also come '*after* belief'. 'The child learns by believing the adult' – Wittgenstein (1969, §160.) thus turned the table, and showed that early learning is dependent on believing the adult: in this sense, trusting the adult. It follows that it is learning that would be impossible without trust, not the other way around. Furthermore, according to Lagerspetz (2015), critical faculties would not be possible to develop if the child was not characterised at the start by *not-doubting*. However, Lagerspetz (2015) opposes Wittgenstein somewhat by arguing that in a realistic situation babies do not trust adults or their parents instantly, from the very start of their existence. Children's (trust-like) attitude towards their parents should only be conceived as 'fully fledged' believing if the children are already capable of disbelieving, ignoring, and opposing them. Hence, as true as the assumption of trust being present from the earliest infancy may be, it is just as possible that 'we should perhaps not say that babies trust their parents from the very start but only that they *do not distrust* them' (Lagerspetz 2015, 146. Italics in the original.)

At this point, it seems that Olli Lagerspetz is not intent on deciding the debate between the approaches assuming the innate, *a priori* built-in nature of basic trust and the *tabula rasa* ('blank slate') concepts. But shortly afterwards, he states that '[a]n infant who did not start life with a basically trustful attitude towards those around him (. . .) would be considered abnormal' (Lagerspetz 2015, 146). In addition, he thinks that in early childhood, the opposite of trust is not doubt, distrust, or suspicion but fear. Consequently, the lack of basic trusting attitude in an infant would reveal an overwhelming fearfulness. This, essentially, is analogous to Annette Baier's (1986) above-mentioned 'ultra-Hobbist child' image, who, albeit, is not distrustful, s/he 'just' happens to be in the state of lack of trust and fearfulness.

Consequently, even though Lagerspetz himself did not intend to provide an empirical solution, the ontological priority of trust seems to be evident in contrast with distrust[19] – Lagerspetz refers to this particular ontological 'imbalance' in writing about the asymmetry in human growth between trust and distrust. The essentials and the final conclusion of Lagerspetz' argument is included in his earlier book as well (1998). He built his reasoning upon the well-known words of Simone Weil:

[19] Further particulars on the ontological priority of trust over distrust can be found in Robert Stern's (2017) essay, who gives an exhaustive account on Løgstrup's concept of basic trust.

> At the bottom of the heart of every human being, from earliest infancy until the tomb, there is something that goes on invincibly expecting, in spite of all experience of crimes committed, suffered, and witnessed, that good and not evil will be done to him. It is this above all that is sacred in every human being.[20]
>
> (1957, 13; 2005, 71)

This quote serves as a foundation for the trust concept of Lagerspetz: the core of trust is Weil's good-expecting disposition ever-present in the human heart. By combining this good-expecting attitude with LØGSTRUP's (1997) radical account on the *a priori* given trust, LAGERSPETZ (1998; 2015) identifies it to be behind Løgstrup's 'tacit demand' of trust. The combination of Weil and Løgstrup's ideas results in Lagerspetz' own reading of trust as tacit demand. There are two main reasons for the tacitness of this demand. Firstly, its hidden, empirically unobservable characteristic. Secondly, the trusting individual typically does not utter or express it in any way, she herself may not even be aware of it (it is also consistent with the assumption of the unperceivable character of trust). Trust – as LAGERSPETZ (2015) argues citing Løgstrup's related lines in his own translation (for the 'official' English translation, see LØGSTRUP 1997, 18) – is inescapable and *given*, not a matter of any arbitrary or capricious choice.

Simone Weil and Lagerspetz consider the capacity to assume good – the optimistic attitude that USLANER (2002) considers the basis of trust and views as a capacity learnt in early infancy – a pre-inherent phenomenon 'at the bottom of the human heart' that no experience can extinguish. Weil's insightful statement, if we take the liberty of relating it to trust in the light of the non-cognitive interpretations of trust is consonant with ROCHAT's (2010) above-mentioned concept of built-in trust. According to these claims, trust cannot be a learnt, experience-based capacity – it is rather a natural, *a priori* given, unending component of human personality.

It must be noted here that Lagerspetz made perfectly clear that he does not imply (at least not intentionally) any essentialist understanding of trust or any empirical or psychological theory of human states of mind, and reads Weil accordingly. His idea of basic trust is an attempt to establish an ethical perspective through which human life can be viewed – trust included. LAGERSPETZ used Weil's description as a 'point of comparison' whose value 'depends on the extent to which it can help us organize our [ethical] understanding of existing situations' (2015, 150).

[20] I made an alternative translation using the French original (WEIL 1957) and the English translation (WEIL 2005), for the latter is imprecise at key places. The word *indomitably* in the English translation does not carry the same meaning as *invinciblement* in the original French, so I changed it to *invincibly*. The translation of the French word *malgré* to *in the teeth of* is also misleading. This option might be more poetic, but it blurs and significantly changes the meaning of the original by emphasising the seriousness of the experiences mentioned in the passage instead of the original meaning *in spite of* or *despite*. Therefore, I replaced *in the teeth of* with in *spite of* that serves the purposes of an accurate translation and the researcher's interest much better.

Weil's claim on the expectation of goodness in the depths of every human heart, however, has another possible interpretation. Indeed, the passage has no *empirical* value – she did not *ask* each and every person about this, as Lagerspetz argues. By all means, the expectation of *goodness* has the ethical value that Lagerspetz emphasises. Still, WEIL's (2005) major point cannot be dismissed: she clearly stresses the good-expecting characteristic of *every* human being, 'from earliest infancy until the tomb'. It is not an empirical observation but a declaration of human reality, that is, an *anthropo-ontological* statement. Now, it may seem that she talks about *Sein* rather than *Sollen* (WEBER 2012) – ontology and (human) existence instead of ethics and normativity. Ontology, in Wittgenstein's account, precedes the ethical,[21] but given that Weil talks about *goodness*, an inherently ethical notion, her statement seems to represent a pre- or sub-analytic reality of human life wherein *Sein* and *Sollen* are not (yet) divided: in which anthropo-ontological existence and ethical value coincide. The coincidence of ethics and ontology can also be grasped in her view of the Good 'steering' the world (FINCH 1999). Furthermore, this coincidence of existence and value enables the ethical value to be an 'unchoice'.[22]

Wittgenstein claimed the lack of relation of good (values) and the world (facts) (FINCH 1971). WITTGENSTEIN (1961, 79e-80e) proposed that '[g]ood and evil only enter through the subject', and '[w]hat is good and evil is essentially the I', the 'bearer of ethics'. As FINCH (1999, 15) wrote on Weil's ineradicable expectation of goodness, it is one of the two 'tiny points' that links the utterly fragile human beings to other things – and both points are *impersonal*. In this regard, Wittgenstein's account is irreconcilable with Weil's.[23]

The inherently onto-ethical nature of Weil's claim is 'inescapable' (just as trust is). By utilising it as a core element, this onto-ethical characteristic pervades and shines through Lagerspetz' strictly ethically aimed idea of trust as a tacit demand of goodness, despite the paradigmatically ethical framework of his dealing with the issue. Even though Lagerspetz never intended to step over the boundaries of ethics, his description of basic trust carries tacit essentialist overtones, just as 'Weil's insistence on the (. . .) primordial character of our reactions to human beings' (LAGERSPETZ 2015, 151). He also admits the ontological nature of the reactions Weil described:

[21] The late expert of Wittgenstein as well as Weil's philosophy, Henry LeRoy FINCH (1971) gave an exhaustive analysis of Wittgenstein's separation of facts and values.

[22] And on the one hand, one might also suspect it to be the very factor that is responsible for (basic) trust's inescapable and rationally unchosen character as well. On the other hand, this coincidence of what is (*Sein*) and what ought to be (*Sollen*) enables Weil to safely assume the sacredness of this essential human expectation of goodness.

[23] FINCH also shows in his remark on Weil's view on rights that a demand (of fairness) arising from personality would necessarily be 'connected with calculation, possessiveness, and envy' (1999, 15). It follows that Weil's expectation of goodness cannot be personal / subjective, thus cannot *only* be ethical in the Wittgensteinian sense. For if it could be, it would be impossible to build Lagerspetz' notion of (basic) trust on it, because calculation only results in reliance or distrust, as the non-cognitivists (Lagerspetz included) argue.

'These are *natural* responses, seeds from which ethical concepts (. . .) subsequently develop (2015, 151. Italics added)'.[24]

There is but one glitch: however elegant the argument of Lagerspetz and Rochat might be, however catching Weil's insight is, however rational and reasonable Uslaner's, Putnam's, and Erikson's reasoning sounds, none of them provided any true *proofs* of their concept. To be fair, the strictly ethical analysis of Lagerspetz would not require any such proofs, still, it makes sense to call for empirically observable indicators of his tacitly onto-ethical claim. Lagerspetz demonstrated basic trust on the level of human ethics – the attempt to find its equivalent on the sub-ethical and subhuman level might be a necessary validation of the aforementioned tacit claims. At this point, one may recall ROCHAT's (2010) remark on built-in trust being *biologically* determined. He leaves this remark without elaborating on it – the subsequent chapters of the present paper discuss this topic.

3.3. The role of oxytocin

According to Mario BUNGE (1979), the universe is a world of interconnected systems. Bunge gives an exhaustive account on these systems, first of which is the physical-chemical system encompassing particles from the atoms to biomolecules, such as the DNA. The biological system (or the 'living things') emerges from the physical-chemical system by self-assembly. There are different biotic levels: the cells, components of an organism, an organism, population, ecosystem, and the biosphere. BUNGE claims that in this hierarchy, every level has spontaneously emerged, 'self-assembled from things at the proceeding level' (1979, 85), prebiotic or biotic. The human organism, together with its brain (a neural system) is a sub-system of the social system and is affected by exogenous influences besides the endogenous influences of its biological system. The social system with its (artificial) subsystems – the economical, the cultural, and the political system – is thus based on and emerges in a natural evolutionary process from the physical-chemical, the biological, and the psychological systems, interwoven with environmental influences.

A remark has to be made here. Although Bunge is aware of Michael POLÁNYI's 1968 paper 'Life's Irreducible Structure', he seems to have missed some important points of it (only refers to it while discussing the reductionist, goal-directed 'machinism' of biology), most probably because of his rigid ('emergentist' / 'systemist') materialist stance. He rejected, though, sheer biological determinism and promoted the 'nature *and* nurture' view of human behaviour. Still, POLÁNYI's (1968) argumentation is more elegant in this regard, whose emergentism is different from Bunge's emergentist materialism. One fundamental difference between the two is that Bunge allowed only for an ultimately closed system, whereas Polányi built his theory on

[24] The question arises: could Lagerspetz have avoided such essentialist overtones? If the unchosen, onto-ethical expectation of goodness is taken out of the idea of Lagerspetzean basic trust, it weakens the concept critically. Therefore, Weil's claim is essential, not only in a general sense, but regarding Lagerspetz' idea as well.

a more flexible, stable open-system approach (Paksi 2009). His moderate materialist reductionism notwithstanding, Bunge's heroic endeavour shows the benefits of integrating the insights of seemingly distant fields of science, in this case the life sciences and the social sciences. If treated carefully, discoveries of the life sciences can enrich the understanding of social phenomena.

3.3.1. Nature versus nurture – a few remarks

The 'nature versus nurture' debate has been around since the very dawn of systematic human thought: it can be traced back to Plato and Aristotle. On the origins of human knowledge, Plato claimed that humans are born with innate concepts, i.e. the universal forms (representing the 'nature' side), whereas Aristotle proposed in *De Anima* (2016) that knowledge and behaviour come from experiences, and human reason is a clean tablet, on which nothing is written (i.e. *tabula rasa*, fitting to the 'nurture' accounts). Their views of women's role in society seem a bit controversial, though: Plato argued in *The Republic* (2000) that, provided they get the same education (i.e. 'nurture'), women and men are equally capable of performing the role of guardians in the ideal society, whereas Aristotle stated in *Politics* (1998) that women are inferior to men by nature. Without delving further into the details of this ever-present debate, it has to be made clear that it later developed into a huge cross-disciplinary argument involving philosophy, biology, psychology, sociology, and so on. In the present days, extremist ideas promoting either biological determinism on the nature side (e.g. genes, neuroendocrine systems, etc.) or a 'blank slate' nurture approach of environmental influences (e.g. socialisation) are both considered as incorrect. The predominant view is that human behaviour in general arise from the complex interplay of endogenous (biological, genetic) as well as exogenous (environmental, social, etc.) influences (Bouchard 1994; Cacioppo et al. 2000; Johnson 2007; Riedl & Javor 2012). A proper discussion of the nature vs. nurture debate would fill books, and is not the subject of the present paper, therefore I cannot pursue this matter further. It is important to see, however, that the question of the (metaphorical) roots of trust, as every other aspect of human agency, is inextricably linked to and situated in this most ancient of debates.

3.3.2. Oxytocin and trust

After these introductory remarks, let us get to the point. The answer for the question of how trust can be 'inscribed' in the biological system can be found in a life sciences discovery on certain hormonal processes. That is, human beings have not only the better-known cortisol-based hormonal response to stress, but also the less-known oxytocin-based response.[25] If one looks at the constantly growing body of literature

[25] It should be noted that the aim of the following passages is not to provide a full account on the matter from a life sciences aspect. The inter-disciplinary approach is used only for the sake of deciding the question of learnt versus innate trust. If interested in further details, see for instance Riedl and Javor's (2012) exceptionally thorough literature review.

of the life sciences on human cooperative behaviour, s/he might get the impression that the question on the origins of trust is an already answered one, not worthy of further discussion. Such a rash judgment would miss the fact that no satisfactory analysis integrates the issues and debates of the 'high' theoretical level with the 'low-level' biological findings, thus resolving the debate of cognitive versus non-cognitive trust. The aim of the subsequent passages is just that: to announce, finally, an indisputable 'winner' based on the firm facts instead of – in certain cases quite inapt – reasoning.

Although there are studies concluding that the propensity to socially oriented action is wired in the human being even *before* birth (Castiello et al. 2010, analysing the movements of twin foetuses), suggesting the *prenatal nature of trust*, such a hypothesis would be a too long shot yet. Instead, the present paper sticks to the less obscure role of oxytocin in human cooperation, especially in trusting behaviour. In contrast with the cortisol-based *fight-or-flight* stress response, which entails aggressive and anti-social behaviour among others, the oxytocin-based *tend and befriend* stress response reduces anxiety and fear, promotes cooperation and trust, and induces prosocial behaviour (Taylor et al. 2000; Kirsch et al. 2005; Kosfeld et al. 2005; Zak et al. 2005; Taylor 2006; Krueger et al. 2007; Baumgartner et al. 2008; Varga 2009; Buchanan & Preston 2014; McGonigal 2015) under stress conditions characterised by acute threat (Buchanan & Preston 2014). In spite of the common stereotype (as hypothesised by Taylor (2006), too), it is not only true for females but also for males. In the study of von Dawans and colleagues (2012), acute stress conditions increased prosocial behaviour in men such as trust, trustworthiness, and sharing. According to Buchanan and Preston, this result '[undermines] the assumption that stressed men would revert to anti-social "fight-or-flight" responses and support[s] the possibility of prosocial effects' (2014, 2). Consequently, human beings are proven to be equipped with the ability to react in a trusting manner even at the biological level, regardless of sex. That is to say, the ability to react even to stressful situations prosocially and trusting is *built-in* the human neurobiological system. Humans are born with the possibility of oxytocin-based response. This fact alone proves the innate nature of trust and denies the early learning hypothesis at the same time. The present paper does not assume, though, that the propensity to trust could not change after birth. To avoid misunderstanding, it is important to emphasise that the dichotomy is not about whether (early) learning matters or not with respect to trust. The validation of the innate nature of the possibility to trust is not to be mistaken for a total denial of the importance and role of early learning. What it denies is the popular blank slate account of Erikson, the cognitive theories, and Uslaner, namely, that trust would *only* be a learnt faculty. Consequently, the argumentation proves not more and not less that the *capacity* to trust is *a priori* given, built in the human hormonal system, and not solely a (by-)product of learning processes. It is rather in line with the conclusions of Riedl and Javor's (2012) literature review on the biological perspective of trust research, taking into account not only the hormonal, but also the genetic as well as brain level in their analysis.

3.3.3. Michael Polányi and the limits of the biological explanation

The proof of Rochat's (2010) above-cited biological 'manifestation' of trust is found, validating the non-cognitive approach on firm factual grounds. At this point, the other, not less important part of Rochat's claim has to be examined: trust being biologically *determined*. Many seem to agree with this assumption, such as Paul Zak (2011) – who goes as far as identifying the *molecule* of trust, empathy, and morality in oxytocin –, Kelly McGonigal (2015), or the approach (Kirsch et al. 2005; Kosfeld et al. 2005; Baumgartner et al. 2008, van Ijzendoorn & Bakermans-Kranenburg 2011) presenting external oxytocin administration as a potential means of trust stimulation, and so on.

However, these approaches simplify the highly complex phenomenon of trust. In the recent years, concerning trust, many (Cacioppo et al. 2000; Riedl & Javor 2012), warn against such a simplification of biological determinism. The aforementioned conclusions of the nature versus nurture debate also suggest that true trust cannot be induced merely by meddling with the biological-hormonal system. In line with these conclusions, trust, as all other aspects of human agency, is under the effect of a highly complex interplay of biological, psychological, social, cultural, etc. factors. For the purposes of the present paper, Michael Polányi's (1966) analysis on the *principle of marginal control* and the *boundary conditions* of a system provides the means for a firm argument against biological determinism without a long adventure on the fields of the nature versus nurture debate. The introduction of Polányi's theory in this argument has another additional value: he gives a rather elegant and simple, but hardly refutable account against determinisms of any kind – even against 'integrative determinism' suggesting that trust is a result of 'only' the complex interplay of biological and social etc. factors. For example, Tsur's (2012) paper demonstrates the relevance of Polányi's theory in contemporary debates against biological-neurological determinism.

According to Polányi's principle, the lower levels of operation (of reality) impose restrictions on the higher levels but they do not determine the *entirety* of these upper levels, for their special and unique organising principles cannot be derived from the laws governing the particulars constituent of the lower levels. The organising principles of the upper levels specify the boundary conditions of the system and the control these principles exercise over the particulars forming the lower levels is marginal control. As vocabulary cannot be derived from phonetics, grammar from vocabulary, good style from grammar, neither gives good style the content of a literary work. According to Richard Allen, Polányi argues against objectivist reductionism by explaining that 'a comprehensive entity, or complex performance, exists on at least two levels: that of its subsidiary details and that of the entity or performance itself into which they are organised' (1990, 59–60). Now, it is quite indisputable that trust is a complex performance of a comprehensive entity. By understanding this system of *dual control*, as Allen (1990) calls it, the reason can also be understood why trust cannot be fully determined by hormonal (or cognitive) functions.

Take trust as the *content* and the mode of action of oxytocin as the grammar (the oxytocin molecule is similar to phonetics in this comparison). It is clear from this example that the phenomenon of trust cannot be understood (or created) on the sole basis of oxytocin effects on human brain and behaviour, as neither can a literary work be understood or created on the basis of a mere analysis of its grammar and phonetics. According to Polányi (1966, 41), 'we can see this principle of marginal control operating also in the hierarchy of biotic levels. (. . .) [T]he principle of marginality (. . .) is present alike in artifacts, like machines; in human performances, like speech; and in living functions at all levels.' Polányi marks the predominant view of the biologists claiming that all manifestations of life and living functions are possible to be explained from the laws of inanimate nature a 'patent nonsense' (Polanyi 1966, 37). He recognised that sentience – 'the most striking feature of our existence' (Polanyi 1966, 37) – as an upper level is based directly or indirectly on the 'governors' of the lower levels, i.e. laws of physics and chemistry as well, but not purely on them. These laws do not include any *conception* of sentience. Therefore, 'any system wholly determined by these laws must be insentient. It may be in the interest of science to turn a blind eye on this central fact of the universe, but it is certainly not in the interest of truth' (Polanyi 1966, 37–38). As Harry Prosch summarises, since Polányi's understanding of the notion that being exists in hierarchical levels, '[w]e are no longer faced with the hopeless task of attempting to explain sentience by concepts taken from the insentient' (1986, 131). It is 'rooted in and dependent for their existence upon a lower, insentient level' (Prosch 1986, 131), but it also structures the lower, insentient level. Sentience is not reducible to the physical level. It is *controlled* by the insentient, it might even be *evoked* by the insentient, but is not *determined* by the insentient. Therefore, it is *uncaused* (Prosch 1986).

Humans are, no doubt, among the most developed sentient beings, and trust can be viewed as a sentiment. Trust, at least in certain cases, is something we *feel*, even if most of the time, according to the non-cognitive approaches (and especially to the 'business-as-usual' accounts of Hertzberg and Lagerspetz), we do not even recognise its presence. Consequently, it is impossible to derive the *whole* complex of the sentiment of trust only from the laws of the hormonal function of oxytocin, which is not even fully discovered yet, by the way. Following Polányi's decisive logic, such system would necessarily be 'trustless'. All of the interpretive values emphasised by the non-cognitive trust theories should be absent from such a system. And if affective trust cannot be fully derived from sheer 'sub-sentient' neurohormonal processes, neither can be the phenomenon of the unconscious, default trusting attitude of the non-cognitivists, especially Hertzberg and Lagerspetz – the above-quoted pattern in the weave of human life. Accordingly, without the marginal control of the upper levels, real trusting attitude, let alone morality or prosperous society (as Zak 2011 mistakenly envisages) is theoretically impossible to stem from oxytocin *alone*. Oxytocin-based hormonal reactions cannot be responsible for the entire human phenomenon called *trust*, as a word does not come out of sounds arranged randomly after one another or a sentence out of a tangled mix of words and so on (Polanyi 1966). In

Polányi's words: 'nothing that *ought* to be, can be determined by knowing what *is*' (1966, 41. Italics in the original).[26] Trust is, as Tsur summarises Polányi's principles, subjected to 'bottom-up unpredictability' (2012, 434).

To sum up: given that a lower level can never subject the boundary conditions set by an upper level completely under its control (Polanyi 1966), neither can trust be fully *determined* by oxytocin, although – as it follows from Polányi's above-described principles – trust must undoubtedly obey its biological laws, however obscure and undiscovered these 'laws' may yet be. They do not alone *govern* human trust, but ignoring them would be a mistake. The ambiguous findings on the effect of oxytocin on aggression (Kirsch et al. 2005), and the failure of certain experiments designed to generate trust by external oxytocin administration (Shamay-Tsoory et al. 2009; Declerck et al. 2010; Bartz et al. 2011) clearly illustrate that trust cannot be truly 'triggered' by any ill-considered meddling with human biochemistry. In the same way, the lower level, i.e. the function of oxytocin is necessarily subjected at its boundaries to a marginal control determined by the upper level, the (sentient) human being. Furthermore, these findings are also true for any 'integrative determinism' – no endogenous (internal, biological) or exogenous (external, e.g. social) effect of a 'closed system of worlds' can account for the boundaries.

The research showing the trust-evoking effect of oxytocin-based stress response verifies the hypothesis of the *a priori* inherent nature of our propensity to trust with an important note: they indicate a *plausibility*, not a *necessity*. The hypothesis of biologically determined trust is mistaken, and so is the assumption that trust would be a mere derivate of a learning / socialisation process, however early it may be.

Before proceeding to the next chapter, let me have one final remark. The same argument as the one that was just set forth would be possible to hold for the early learning theories as well as the theories assuming that trust is rooted in experiences and is possible to be derived from them in some (yet obscure) way – see the 'leap of faith' approaches. No lower-level experience can account entirely for the upper-level function of trust, even if we view trust as a choice or a decision. Polányi (1975, 176) is quite clear on the matter that even choices (and decisions, see Prosch 1986) are ' "uncaused", in the sense that there is nothing (. . .) that determines or necessitates that they become precisely what they do become.' As shown above, it is the case with trust, too. The inputs assumed by the cognitive theories to be the 'causes' of trust would result in mere reliance without requiring any obscure 'leaps' of faith. It is not to say that experiences and other cognitive processes cannot at all affect our trusting. In this regard, non-cognitive theories might be mistaken (it might be the subject of a future article).[27] But any attempt on basing trust only on cognitive processes is just

[26] Max Weber (2012) gave an exhaustive explanation on the difference between 'what is' and 'what should be' in his well-known 1904 study on the objectivity of social science.

[27] For instance, according to Allen's reading of Polányi, by studying the lower levels we can acquire knowledge of the limits of trust, too, 'within which it operates or lives, and of how it can break down, go wrong, become ill or die. By being embodied in matter, (. . .) the functioning of organisms, and the thoughts and actions of minds, are put at risk by the lower levels on which they depend' (1990, 63). Experiences can thus

as mistaken as the assumption that trusting would be determined by neurobiological processes alone. The same argument holds for the external influences of socialisation and culture, too. Therefore, keeping these in mind, together with Polányi's above-cited words on the relationship of the known and the existing, any integrative hypothesis assuming that trust would be the 'sole' outcome of the obscure interplay of biological, cognitive, and social factors must be rejected. No determinism is allowed, however nuanced and elaborate it might be. An integrative theory of trust is needed instead, which reconciles cognitive and non-cognitive approaches by building on their advantages and getting rid of their disadvantages, and introducing factors that help avoiding determinism. Such a task cannot be carried out without finding a novel aspect, for the existing frameworks seem to lack the potential.

4. Directions for future research

Although it is a cliché but it is true: rebuilding a collapsed building must start from the foundations. Tiny, sometimes confusing but important signs across the trust literature mark the way. To mention but a few, Erikson (1977), Luhmann (1979), Baier (1986), Dunn (1988), Good (1988), Giddens (1990), Yamagishi and Yamagishi (1994), Sztompka (1999), Marková (2004), Bauman (2006), Marková and colleagues (2006), and Kohn (2008) all give us more or less (usually less) detailed pointers on the dynamic relationship of trust and fear / anxiety. The neurological literature on trust gives the same impression. Is it not suspicious that the above-discussed oxytocin-based stress-response is an alternative to the cortisol-based, which is accompanied by fear, anxiety, and anti-social behaviour? My guess is that it is most suspicious, indeed.

Anthony Giddens (1991), in a few short passages that no trust theorist seems to have taken notice of so far, drew attention to Paul Tillich's (1952) ontology of courage. In this all-important writing[28], Tillich was concerned with the problem of how the human being deals with existential anxiety: the threat of mortality and non-being. Although Giddens misunderstood Tillich in assuming that trust is the source of Tillich's *courage to be*, future trust research has to start with investigating the fundamental existential questions of being and nonbeing with a special attention to existential (death) anxiety (the unavoidable and unending component of human life) and its relationship with trust. This path can help us make sense of such vague and unelaborate assumptions as Möllering's (2006; 2013) who suspects that trust can

set the limits of trust. Although they alone cannot account for trust, knowing of their *corruption* can help us understand the breach of trust (not forgetting the experience-resistant nature of real trust). Perhaps the non-cognitive theories are not so far off by assuming that becoming *aware* of trust and the possible reasons (or lack of reasons) thereof might be a token of the fact that our trusting is already corrupted in a way or another . . . This issue, like many others, also falls out of the scope of the present paper, but I hope I will be able to elaborate it in a future article.

[28] It served, for instance, as a major basis of such highly regarded psychological works as Yalom's well-known *Existential Psychotherapy* (1980).

form *in spite of* / *despite* insufficient information or knowledge – but neither he nor anyone else has ever pursued this suspicion far enough.

MÖLLERING's (2006, 191) well-known agenda should then be rephrased by replacing leap of faith with 'despiteness': unless we contribute how the '*despiteness of trust*' is possible, we miss the point of trust and explain anything but trust. Advancing on this path may also enable distinguishing trust from naïveté, an issue that neither the cognitive nor the non-cognitive approaches address properly.

5. Conclusion

By introducing the oxytocin-related discoveries of the life sciences, the present paper settled the debate between the 'early learning' theories and the '*a priorists*'. Trust is a built-in, innate human faculty, not a capacity picked up as a result of early childhood experiences with the primary caregiver. The possibility to trust is inscribed in the neurobiological system right from the start – humans are *able* to have the oxytocin-based hormonal response since birth (or even earlier). To avoid falling into the pitfall of reductionism, the paper rejected the idea of (biologically) determined trust on the basis of Michael POLÁNYI's (1966) principles of marginal control and boundary conditions. Theorising trust, at least the upper ('sentient') levels of trusting remains the task of the committed representatives of human and social sciences. The present paper contributes to the challenge of a proper future conceptualisation of trust by making the basics clear: metaphorically speaking, how deep the roots of trust dive in the human being.

The general conclusion supports the initial assumption of the introduction: the human is indeed a trusting being, below and at the same time above any early infantile experience or lack of experience. The faculty and thus the possibility and opportunity to trust is always present. As for *how* we are able to truly 'trigger', actualise this opportunity, that shall be the subject of a future article.

References

Allen, R. (1990) *Thinkers of Our Time: Polanyi* (London: Claridge).

Aristotle (1998) *Politics*, trans. C.D.C. Reeve (Indianapolis: Hackett).

Aristotle (2016) *De Anima*, trans. C. Shields (New York: Clarendon & Oxford UP).

Baier, A. (1986) 'Trust and Antitrust', *Ethics* 96:1, 231–60 (https://doi.org/10.1086/292745).

Balázs Z. (2002) 'A bizalom fogalma', *Századvég* 7:2, 27–48.

Balázs Z. (2008) 'Politikai bizalmi válság', *Politikatudományi Szemle* 17:1, 113–29.

Bartz, J., D. Simeon, H. Hamilton, S. Kim, S. Crystal, A. Braun, V. Vicens & E. Hollander (2011) 'Oxytocin Can Hinder Trust and Cooperation in Borderline Personality Disorder', *Social Cognitive and Affective Neuroscience* 6:5, 556–63 (https://doi.org/10.1093/scan/nsq085).

Bauman, Z. (2006) *Liquid Fear* (Cambridge: Polity Press).

Baumgartner, T., M. Heinrichs, A. Vonlanthen, U. Fischbacher & E. Fehr (2008) 'Oxytocin Shapes the Neural Circuitry of Trust and Trust Adaptation in Humans', *Neuron* 58, 639–50 (https://doi.org/10.1016/j.neuron.2008.04.009).

Becker, L.C. (1996) 'Trust as Noncognitive Security about Motives', *Ethics* 107:1, 43–61 (https://doi.org/10.1086/233696).

Berg, N. & G. Gigerenzer (2010) 'As-If Behavioral Economics: Neoclassical Economics in Disguise?', *History of Economic Ideas* 18:1, 133–65 (https://doi.org/10.1400/140334).

Bernstein, J.M. (2011) 'Trust: On the Real But Almost Always Unnoticed, Ever-Changing Foundation of Ethical Life', *Metaphilosophy* 42:4, 395–416 (https://doi.org/10.1111/j.1467-9973.2011.01709.x).

Blau, P.M. (1964) *Exchange and Power in Social Life* (New York: Wiley).

Bok, S. (1978) *Lying: Moral Choice in Public and Private Life* (New York: Pantheon).

Bouchard, T.J. Jr. (1994) 'Genes, Environment, and Personality', *Science* 264, 1700–01 (https://doi.org/10.1126/science.8209250).

Buchanan, T.W. & S.D. Preston (2014) 'Stress Leads to Prosocial Actions in Immediate Need Situations', *Frontiers in Behavioral Neuroscience* 8:5, 1–6 (https://doi.org/10.3389/fnbeh.2014.00005).

Bunge, M. (1979) *Treatise on Basic Philosophy Volume 4. Ontology II: A World of Systems* (Dordrecht: D. Reidel).

Cacioppo, J.T., G.G. Bernston, J.F. Sheridan & M.K. McClintock (2000) 'Multilevel Integrative Analyses of Human Behavior: Social Neuroscience and the Complementing Nature of Social and Biological Approaches', *Psychological Bulletin* 126:6, 829–43 (https://doi.org/10.1037/0033-2909.126.6.829).

Castiello, U., C. Becchio, S. Zoia, C. Nelini, L. Sartori, L. Blason, G. D'Ottavio, M. Bulgheroni & V. Gallese (2010) 'Wired to Be Social: The Ontogeny of Human Interaction' *PLoS ONE* 5:10, e13199, retrieved 01 Oct 2018 from https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0013199&type=printable (https://doi.org/10.1371/journal.pone.0013199).

Charmaz, K. (2006) *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis* (Thousand Oaks: SAGE).

Coleman, J.S. (1990) *Foundations of Social Theory* (Cambridge, Mass. & London: Harvard UP).

Cook, K. & A. Gerbasi (2009) 'Trust' in P. Bearman & P. Hedström, eds., *The Oxford Handbook of Analytical Sociology* (New York: Oxford UP) 218–41 (https://doi.org/10.1093/oxfordhb/9780199215362.013.10).

Dasgupta, P. (1988) 'Trust as a Commodity' in D. Gambetta, ed., *Trust: Making and Breaking Cooperative Relations* (Oxford: Blackwell) 49–72.

Declerck, C.H., C. Boone & T. Kiyonari (2010) 'Oxytocin and Cooperation Under Conditions of Uncertainty: The Modulating Role of Incentives and Social Information', *Hormones and Behavior* 57:3, 368–74 (https://doi.org/10.1016/j.yhbeh.2010.01.006).
Delhey, J. & K. Newton (2003) 'Who Trusts? The Origins of Social Trust in Seven Societies', *European Societies* 5:2, 93–137 (https://doi.org/10.1080/1461669032000072256).
Duncan, G. (2019) *The Problem of Political Trust: A Conceptual Reformulation* (New York: Routledge) (https://doi.org/10.4324/9781351061469).
Dunn, J. (1988) 'Trust and Political Agency' in D. Gambetta, ed., *Trust: Making and Breaking Cooperative Relations*. (Oxford: Blackwell) 73–93.
Emerson, R.M. (1976) 'Social Exchange Theory', *Annual Review of Sociology* 2, 335–62 (https://doi.org/10.1146/annurev.so.02.080176.002003).
Erikson, E.H. (1977) *Childhood and Society* (London: Paladin Grafton).
Finch, H.L. (1971) *Wittgenstein – The Early Philosophy: An Exposition of the "Tractatus"* (Atlantic Highlands: Humanities Press).
Finch, H.L. (1999) *Simone Weil and the Intellect of Grace* (New York: Continuum).
Friedman, M. (1953) *Essays in Positive Economics* (Chicago: The University of Chicago Press).
Fukuyama, F. (1995) *Trust: The Social Virtues and the Creation of Prosperity* (New York: Free Press).
Gambetta, D. (1988) 'Can We Trust Trust?' in D. Gambetta, ed., *Trust: Making and Breaking Cooperative Relations*. (Oxford: Blackwell) 213–38.
Giddens, A. (1990) *The Consequences of Modernity*. (Stanford: Stanford UP).
Giddens, A. (1991) *Modernity and Self-Identity: Self and Society in the Late Modern Age* (Cambridge: Polity Press).
Glaser, B.G. & A.L. Strauss (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research* (New Brunswick & London: AldineTransaction).
Good, D. (1988) 'Individuals, Interpersonal Relations, and Trust' in D. Gambetta, ed., *Trust: Making and Breaking Cooperative Relations*. (Oxford: Blackwell) 31–48.
Gouldner, A.W. (1960) 'The Norm of Reciprocity: A Preliminary Statement', *American Sociological Review* 25:2, 161–78 (https://doi.org/10.2307/2092623).
Govier, T. (1993) 'Self-Trust, Autonomy, and Self-Esteem', *Hypatia* 8:1, 99–120 (https://doi.org/10.1111/j.1527-2001.1993.tb00630.x).
Greenhalgh, T. & R. Peacock (2005) 'Effectiveness and Efficiency of Search Methods in Systematic Reviews of Complex Evidence: Audit of Primary Sources', *British Medical Journal* 331, 1064–65 (https://doi.org/10.1136/bmj.38636.593461.68).
Hardin, R. (1993) 'The Street-Level Epistemology of Trust', *Politics Society* 21:4, 505–29 (https://doi.org/10.1515/auk-1992-0204).
Hardin, R. (2006) *Trust* (Cambridge: Polity Press).
Hertzberg, L. (1988) 'On the Attitude of Trust', *Inquiry: An Interdisciplinary Journal of Philosophy* 31:3, 307–22 (https://doi.org/10.1080/00201748808602157).
Hertzberg, L. (2010) 'On Being Trusted' in A. Grøn & C. Welz, eds., *Trust, Sociality, Selfhood.* (Tübingen: Mohr Siebeck) 193–204.
Hirschman, A.O. (1984) 'Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse', *American Economic Review* 74:2, 88–96 (https://doi.org/10.1017/S0266267100001863).
Hollis, M. (1998) *Trust Within Reason* (Cambridge: Cambridge UP) (https://doi.org/10.1017/CBO9780511612244).
Homans, G.C. (1961) *Social Behavior: Its Elementary Forms* (New York: Harcourt, Brave & World).
Johnson, W. (2007) 'Genetic and Environmental Influences on Behavior: Capturing All the Interplay', *Psychological Review* 114:2, 423–40 (https://doi.org/10.1037/0033-295X.114.2.423).

Jones, K. (1996) 'Trust as an Affective Attitude', *Ethics* 107:1, 4–25 (https://doi.org/10.1086/233694).

Kierkegaard, S. (1983) 'Fear and Trembling' in S. Kierkegaard, *Fear and Trembling; Repetition*, trans & eds. H.V. Hong & E.V. Hong (Princeton: Princeton UP) 1–123.

Kirsch, P., C. Esslinger, Q. Chen, D. Mier, S. Lis, S. Siddhanti, H. Gruppe, V.S. Mattay, B. Gallhofer & A. Meyer-Lindenberg (2005) 'Oxytocin Modulates Neural Circuitry for Social Cognition and Fear in Humans', *The Journal of Neuroscience* 25:49, 11489–93 (https://doi.org/10.1523/JNEUROSCI.3984-05.2005).

Kohn, M. (2008) *Trust. Self-Interest and the Common Good* (New York: Oxford UP).

Kosfeld, M., M. Heinrichs, P.J. Zak, U. Fischbacher & E. Fehr (2005) 'Oxytocin Increases Trust in Humans', *Nature* 435, 673–76 (https://doi.org/10.1038/nature03701).

Krueger, F., K. McCabe, J. Moll, N. Kriegeskorte, R. Zahn, M. Strenziok, A. Heinecke & J. Grafman (2007) 'Neural Correlates of Trust', *PNAS* (*Proceedings of the National Academy of Sciences of the United States of America*) 104:50, 20084–89 (https://doi.org/10.1073/pnas.0710103104).

Lagerspetz, O. (1998) *Trust: The Tacit Demand* (Dordrecht: Kluwer) (https://doi.org/ 10.1007/978-94-015-8986-4).

Lagerspetz, O. (2015) *Trust, Ethics and Human Reason* (London: Bloomsbury).

Lewis, D.J. & A. Weigert (1985) 'Trust as a Social Reality', *Social Forces* 6:34, 967–85 (https://doi.org/10.1093/sf/63.4.967).

Li, P. P. (2015) 'Trust as a Leap of Hope for Transaction Value: A Two-Way Street Above and Beyond Trust Propensity and Expected Trustworthiness' in B.H. Bornstein & A.J. Tomkins, eds., *Motivating Cooperation and Compliance with Authority: The Role of Institutional Trust* (New York: Springer) 37–53 (https://doi.org/10.1007/978-3-319-16151-8_3).

Løgstrup, K.E. (1997) *The Ethical Demand* (Notre Dame: University of Notre Dame Press).

Luhmann, N. (1979) *Trust and Power* (Chichester: Wiley).

Luhmann, N. (1988) 'Familiarity, Confidence, Trust' in D. Gambetta, ed, *Trust: Making and Breaking Cooperative Relations* (Oxford: Blackwell) 94–107.

Marková, I. (2004) 'Introduction: Trust/Risk and Trust/Fear' in I. Marková, ed., *Trust and Democratic Transition in Post-Communist Europe* (Oxford: Oxford UP) 1–23.

Marková, I., P. Linell & A. Gillespie (2007) 'Trust and Distrust in Society' in I. Marková & A. Gillespie, eds., *Trust and Distrust: Sociocultural Perspectives* (Charlotte: Information Age Publishing) 3–27.

Mayer, R.C., J.H. Davis & F.D. Schoorman (1995) 'An Integrative Model of Organizational Trust', *The Academy of Management Review* 20:3, 709–34 (https://doi.org/10.5465/amr.2007.24348410).

McGonigal, K. (2015) *The Upside of Stress: Why Stress is Good for You, and How to Get Good at It* (New York: Avery).

Misztal, B. (1996) *Trust in Modern Societies* (Cambridge: Polity Press).

Möllering, G. (2006) *Trust: Reason, Routine, Reflexivity* (Oxford: Elsevier).

Möllering, G. (2013) 'Trust Without Knowledge? Comment on Hardin, 'Government without Trust'', *Journal of Trust Research* 3:1, 53–58 (https://doi.org/10.1080/21515581.2013.771504).

Mújdricza F. & Földvári M. (2018) 'A kognitív kérdőívtesztelés módszertana: A kognitív interjúk elemzése', *Statisztikai Szemle* 96:6, 545–74 (https://doi.org/10.20311/stat2018.06.hu0545).

Nishina, K., H. Takagishi, F. Alan, M. Inoue-Murayama, H. Takahashi, M. Sakagami & T. Yamagishi (2018) 'Association of the Oxytocin Receptor Gene With Attitudinal Trust in Men: Role of the Amygdala Volume', *Social Cognitive and Affective Neuroscience* 13:10, 1091–97 (https://doi.org/10.1093/scan/nsy075).

Offe, C. (1999) 'How Can We Trust Our Fellow Citizens?' in M.E. Warren, ed., *Democracy and Trust* (Cambridge: Cambridge UP) 42–87 (https://doi.org/10.1017/CBO9780511659959.003).
Örkény A. & Székelyi M. (2009) 'The Role of Trust in the Social Integration of Immigrants', *Demográfia* 52:5, 124–47.
Paksi D. (2009) Az evolúció és az emergencia fogalmának értelmezése Polányi Mihály filozófiájában (PhD diss., Budapesti Műszaki és Gazdaságtudományi Egyetem, Budapest) retrieved 20 Sept 2018 from https://repozitorium.omikk.bme.hu/bitstream/handle/10890/920/ertekezes.pdf?sequence=1.
Pettit, P. (1995) 'The Cunning of Trust', *Philosopy and Public Affairs* 24:3, 202–25 (https://doi.org/10.1111/j.1088-4963.1995.tb00029.x).
Plato (2000) *The Republic*, trans T. Griffith (Cambridge: Cambridge UP).
Polányi, M. (1966) *The Tacit Dimension* (Chicago & London: The University of Chicago Press).
Polányi, M. (1968) 'Life's Irreducible Structure', *Science* 160:3834, 1308–12 (https://doi.org/10.1126/science.160.3834.1308).
Polányi, M. (1975) *Meaning* (Chicago & London: The University of Chicago Press).
Prosch, H. (1986) *Michael Polanyi. A Critical Exposition* (Albany: State University of New York Press).
Putnam, R.D. (1993) M*aking Democracy Work: Civic Traditions in Modern Italy* (Princeton: Princeton UP).
Putnam, R.D. (2000) *Bowling Alone: The Collapse and Revival of American Community.* (New York: Simon & Schuster) (https://doi.org/10.1145/358916.361990).
Randolph, J.J. (2009) 'A Guide to Writing the Dissertation Literature Review', *Practical Assessment, Research & Evaluation* 14:13 retrieved 12 June 2016 from http://pareonline.net/getvn.asp?v=14&n=13.
Riedl, R. & A. Javor (2012) 'The Biology of Trust: Integrating Evidence From Genetics, Endocrinology, and Functional Brain Imaging', *Journal of Neuroscience, Psychology, and Economics* 5:2, 63–91 (https://doi.org/10.1037/a0026318).
Rochat, P. (2010) 'Trust in Early Development' in A. Grøn & C. Welz, eds., *Trust, Sociality, Selfhood* (Tübingen: Mohr Siebeck) 31–43.
Rousseau, D.M., S.B. Sitkin, R.S. Burt & C. Camerer (1998) 'Not So Different After All: A Cross-Discipline View of Trust', *Academy Management Review* 23:3, 393–404 (https://doi.org/10.5465/amr.1998.926617).
Shamay-Tsoory, S.G., M. Fischer, J. Dvash, H. Harari, N. Perach-Bloom & Y. Levkovitz (2009) 'Intranasal Administration of Oxytocin Increases Envy and Schadenfreude (Gloating)', *Biological Psychiatry* 66:9, 864–70 (https://doi.org/10.1016/j.biopsych.2009.06.009).
Stern, R. (2017) ''Trust is Basic': Løgstrup on the Priority of Trust' in P. Faulkner & T. Simpson, eds., *The Philosophy of Trust* (New York: Oxford UP) 272–93 (https://doi.org/10.1093/acprof:oso/9780198732549.003.0016).
Sztompka, P. (1999) *Trust: A Sociological Theory* (New York: Cambridge UP).
Taylor, S.E. (2006) 'Tend and Befriend: Biobehavioral Bases of Affiliation Under Stress', *Current Directions in Psychological Science* 15:6, 273–77 (https://doi.org/10.1111/j.1467-8721.2006.00451.x).
Taylor, S.E., L.C. Klein, B.P. Lewis, T.L. Gruenewald, R.A.R. Gurung & J.A. Updegraff (2000) 'Biobehavioral Responses to Stress in Females: Tend-and-Befriend, Not Fight-or-Flight', *Psychological Review* 107:3, 411–29 (https://doi.org/10.1037/0033-295X.107.3.411).
Thibaut, J.W. & H.H. Kelley (1959) *The Social Psychology of Groups* (New York: Wiley).
Tillich, P. (1952) *The Courage to Be* (New Haven & London: Yale UP).
Tillich, P. (2009) *Dynamics of Faith* (New York: HarperOne).

Tsur, R. (2012) 'The Neurological Fallacy', *Pragmatics & Cognition* 20:3, 429–66 (https://doi.org/10.1075/pc.20.3.01tsu).

Uslaner, E.M. (2002) *The Moral Foundations of Trust* (New York: Cambridge UP).

van Ijzendoorn, M.H. & M.J. Bakermans-Kranenburg (2011) 'A Sniff of Trust: Meta-Analysis of the Effects of Intranasal Oxytocin Administration on Face Recognition, Trust to In-Group, and Trust to Out-Group', *Psychoneuroendocrinology* 37:3, 438–43 (https://doi.org/10.1016/j.psyneuen.2011.07.008).

Varga, K. (2009) 'Szexualitás, szülés, kötődés: az oxitocin pszichoemotív hatásai' in Bagdy E, Demetrovics Zs. & Pilling J., eds., *Polihistória: Köszöntők és tanulmányok Buda Béla 70. születésnapja alkalmából* (Budapest: Akadémiai Kiadó) 447–74.

von Dawans, B., U. Fischbacher, C. Kirschbaum, E. Fehr & M. Heinrichs (2012) 'The Social Dimension of Stress Reactivity: Acute Stress Increases Prosocial Behaviour in Humans', *Psychological Science* 23:6, 651–60 (https://doi.org/10.1177/0956797611431576).

Weber, M. (2012) 'The "Objectivity" of Knowledge in Social Science and Social Policy' in H.H. Bruun & S. Whimster, eds., *MAX WEBER: Collected Methodological Writings* (London & New York: Routledge) 100–38.

Weil, S. (1957) 'La Personne et le Sacré' in S. Weil, *Écrits de Londres et dernières lettres* (Paris: Gallimard) 11–44.

Weil, S. (2005) 'Human Personality' in S. Miles, ed., *Simone Weil: An Anthology* (London: Penguin) 69–98.

Williams, B. (1988) 'Formal Structures and Social Reality' in D. Gambetta, ed., *Trust: Making and Breaking Cooperative Relations* (Oxford: Blackwell) 3–13.

Wittgenstein, L. (1961) *Notebooks 1914-1916*, trans. G.E.M. Anscombe, eds. G.H. von Wright & G.E.M. Anscombe (New York & Evanston: Harper & Row).

Wittgenstein, L. (1969) *On Certainty*, trans. D. Paul & G.E.M. Anscombe, eds. G.E.M. Anscombe & G.H. von Wright (Oxford: Basil Blackwell).

Yalom, I.D. (1980) *Existential Psychotherapy* (New York: Basic Books).

Yamagishi, T. (2011) *Trust: The Evolutionary Game of Mind and Society* (New York: Springer) (https://doi.org/10.1007/978-4-431-53936-0).

Yamagishi, T. & M. Yamagishi (1994) 'Trust and Commitment in the United States and Japan', *Motivation and Emotion* 18:2, 129–66 (https://doi.org/10.1007/BF02249397).

Zak, P. (2011) 'Trust, Morality – And Oxytocin? TEDtalk' retrieved 22 April 2018 from http://www.ted.com/talks/paul_zak_trust_morality_and_oxytocin.

Zak, P.J., R. Kurzban & W.T. Matzner (2005) 'Oxytocin is associated with human trustworthiness', *Hormones and Behavior* 48, 522–27 (https://doi.org/10.1016/j.yhbeh.2005.07.009).