

Please cite as:

Simon, J. (2013). Trust. In: Pritchard, D. (Ed.): Oxford Bibliographies in Philosophy. New York: Oxford University Press. Available online at:  
<http://www.oxfordbibliographies.com/view/document/obo-9780195396577/obo-9780195396577-0157.xml>

Philosophy

Entry ID: 9780195396577-0143

Citation Style: Humanities

Version Date: uploaddate

## **TRUST**

Judith Simon

University of Vienna, Department of Philosophy

Karlsruhe Institute of Technology, Institute for Technology Assessment and Systems Analysis

### **Introduction**

General Overviews

Reference Works and Bibliographies

Anthologies

Accounts of Trust

Trust and Trustworthiness

Biases and Injustices in Trusting and Assessing Trustworthiness

Trust without evidence

Beyond Interpersonal Trust

Self-Trust

Trust in Society, Institutions, and Governments

Trust and Knowledge

Trust and Testimony

Trust in Science, Trust within Science

Trust in Socio-Technical Systems

E-Trust

Trust in Wikipedia

## **INTRODUCTION**

Imagine a world without trust. Without trust even the simplest actions in our lifeworld would hardly ever happen: We would never enter a taxi, never pay with coin or believe in what our doctor says. We would neither know when and where we are born and might even still believe that the sun rotates around the earth. Trust is an essential trait of our social life and our relation to our environment. Given the pervasiveness of trust in our lives, it comes as a surprise that trust has only very recently started to receive attention in philosophy. Apart from some early consideration on trust amongst friends, on trust in god as well as some contributions regarding the role of trust in society by Hobbes, Locke and Hume, trust emerged as a topic of philosophical interest only in the last decades of the 20<sup>th</sup> century. As pervasive trust appears as a phenomenon, as elusive it seems as a concept. What is trust? Is it a belief, an expectation, an attitude or an emotion? Can trust be willed or can I merely decide to act *as if* I trusted? Moreover, while the intrinsic as well as the instrumental value of trust for cooperation and social life is almost commonsensical, trust

always carries the risk of being unwarranted. Trusting those who are not worthy of our trust can lead to exploitation and betrayal. Yet, not trusting those who would be trustworthy can also be a mistake and cause harm. It has been particularly feminist scholars, who have emphasizes the janus-faced nature of trust. How trust is defined and characterized depends strongly on the examples chosen. It makes a difference whether we analyze trust relations between children and their parents, between humans of equal power, between friends, lovers or strangers. Trust in other persons differs from trust in groups, trust in a specific representative of the state differs from trust in more abstract entities such as governments, democracy or society. Technology is yet another important if neglected patient of trust as well as a mediator of many human-to-human trust relations in our contemporary world. Instead of distinguishing proper and improper uses of the term trust, we should carefully attend to these different meanings of the word "trust" to provide a rich and multi-facetted perspective on this complex and important phenomenon.

## GENERAL OVERVIEWS

Kohn's 2008 essay on trust is a concise and well-written starting point to get an overview on the topic of trust, while Govier 1997 offers a particularly deep and nuanced analysis of social trust in its manifold forms and appearances. Both Hardin 2006 and O'Neill 2002 focus specifically on public trust, respectively the lack of it. Baier's 1986 paper is probably the single-most influential article on trust in philosophy and a must-read. (See also \*Accounts of Trust\*). Gambetta's 1988 important paper serves as an introduction to game-theoretic accounts of trust. As a concluding chapter of Gambetta's important anthology it also serves as a pointer to other papers of this anthology (See also \*Anthologies\*). Hollis 1998 proposes a communitarian account of trust explicitly against game-theoretic considerations. Finally, Luhmann's 1979 account of trust, although rooted in sociology, has been highly influential in philosophy and is therefore a must-read.

Baier, Annette C. "Trust and Antitrust." *Ethics* 96.2 (1986): 231-60.

One of the most influential philosophical papers on trust. Criticizes contractualist fixation in philosophy and emphasizes the relevance of power inequalities for trust. Her focus on the distinction between trust and reliance remains a dominant theme ever since.

Gambetta, Diego. "Can We Trust Trust?" In *Trust. Making and Breaking Cooperative Relations*, edited by Diego Gambetta, 213-37. Oxford: Basil Blackwell, 1988.

This concluding essay of Gambetta's seminal anthology serves as a good introduction to cognitive accounts of trust highlighting the relationship between trust and cooperation.

Govier, Trudy. *Social Trust and Human Communities*. Montreal: McGill-Queen's University Press, 1997.

Detailed examination of the manifold role of trust in social life by a prolific writer on trust. Explores a wide range of issue from trust in testimony to trust in civil society.

Hardin, Russell. *Trust, Key Concepts in the Social Sciences*. Cambridge: Polity, 2006.

Overview on trust by probably the most prolific writer on trust. Focus on trust in public life. Attacks empirical research on trust as being misguided.

Hollis, Martin. *Trust within Reason*. Cambridge: Cambridge University Press, 1998.

Challenges purely game-theoretic and predictive accounts of trust and defends a communitarian account of trust. Introduces a principle of generalized reciprocity.

Kohn, Marek. *Trust. Self-Interest and the Common Good*. Oxford: Oxford University Press, 2008.  
Brief and well-written essayistic primer into the complexities around trust. Good starting point and easy to read.

Luhmann, Niklas. English Translation published in: Luhmann, Niklas. *Trust and Power*. Toronto: Wiley, 1979.

Starting point for modern examinations of trust in sociology, highly influential also in philosophy. Focus on the function of trust to reduce complexity and uncertainty in social systems. Originally published as *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*. Stuttgart: Enke, 1968.

O'Neill, Onora. \*A Question of Trust[<http://assets.cambridge.org/97805215/29969/sample/9780521529969ws.pdf>]\*. Cambridge: Cambridge University Press, 2002.

Challenges the proclaimed contemporary crisis of trust and accuses the media for fostering public distrust. Criticizes increased amount of bureaucratic controls and panacea of formalized accountability procedures, proposes that we need to restore trust instead.

## REFERENCE WORKS & BIBLIOGRAPHIES

McLeod 2006 delivers an excellent and in-depth starting point into the topic of trust. For a very brief primer into the topic confer Jones 1998. While the **\*\*bibliography on trust\*\*** provided by Philpapers is an excellent and permanently updated resource for philosophical literature on trust, the **\*\*Trust-Wiki\*\*** serves as a good basis for those interested in how trust is approached in neighbouring disciplines.

Jones, Karen. "Trust." In *Routledge Encyclopedia of Philosophy*, edited by Edward Craig, 466-70. London: Routledge, 1998.

In this concise overview, Jones outlines shared assumptions about trust, distinguishes the most important accounts of trust, stresses the question of the justification of trust as well as the differences and relations between interpersonal and organizational trust.

McLeod, Carolyn. "Trust." In *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta, 2006.

By outlining discussions around the nature of trust and trustworthiness, the epistemology of trust, the value of trust, and relationship between trust and will, McLeod maps current debates and links to relevant literature. Very good starting point into the topic of trust and freely available.

Bibliography on Trust in Philpapers: <http://philpapers.org/browse/trust/>

Philpapers delivers a constantly updated bibliography on many aspects of trust in philosophy. Edited by Edward Hinchman.

Trust-Wiki. [http://t3.istc.cnr.it/trustwiki/index.php/Trust\\_across\\_Disciplines](http://t3.istc.cnr.it/trustwiki/index.php/Trust_across_Disciplines)

Useful starting point to explore accounts of trust in different disciplines. Hosted by the Institute of Cognitive Sciences and Technologies (ISTC) in Rome.

## ANTHOLOGIES

Gambetta's 1988 anthology is not only the first, but also the most important anthology on trust including contributions by Niklas Luhman, Bernard Williams and many others. Apart from Gambetta's anthology, the Russell Sage Series on Trust is the most important reference point for interdisciplinary research on trust. Although many contributions are rather rooted in economics, social or political science, they offer many important sources for philosophical analysis of very different aspects of trust. Cook's 2003 anthology on trust in society addresses the social origins and consequences of trust, while Braithwaite and Levi 2003 focuses particularly on the state as an agent and patient of trust. Ostrom and Walker's 2005 collection on trust and reciprocity offers an overview on experimental research on trust. The topic of distrust, while inherent in many accounts of trust, is explicitly addressed in Hardin 2004 as well as more recently in Kramer and Cook 2009. The role of trust in the online world has recently stirred a lot of interest also amongst philosophers. While the anthology on e-trust by Cook et al. 2009 has a decidedly economic focus, the collection by Ess and Thorseth 2011 offers a more nuanced ethical analysis of different trust relations in virtual worlds.

Braithwaite, Valerie, and Margaret Levi, eds. *Trust and Governance*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi. Vol. I, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2003.

Anthology focusing on the role of trust in the state. Numerous significant contributions shedding light on the basis for trusting the state and its agents, the concept of a trustworthy state, the role of trust for representative democracy and the concept of trust responsiveness.

Cook, Karen S., ed. *Trust in Society*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi. Vol. II, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2003.

Important anthology comprising of a highly interdisciplinary set of articles focusing on the role of trust in very different aspects of social life.

Cook, Karen S., Chris Snijders, Vincent Buskens, and Coye Cheshire, eds. *Etrust. Forming Relationships in the Online World*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2009.

Anthology on the emerging topic of e-trust, albeit with a narrow focus on the role of trust and reputation for commercial interaction online. (See \*E-trust\* for accounts of e-trust beyond economic interactions)

Ess, Charles, and May Thorseth, eds. *Trust in Virtual Worlds: Contemporary Perspectives*. New York: Peter Lang, 2011.

Useful collection of papers on the ethics of trust in online environments. (See also \*E-Trust\*)

Gambetta, Diego, ed. *\*Trust. Making and Breaking Cooperative Relations*[<http://www.sociology.ox.ac.uk/papers/trustbook.html>]\*. Oxford: Basil Blackwell, 1988.

The single-most important and influential anthology of trust edited by one of its most prolific theoreticians. Numerous important contributions from different disciplines. Rational-choice perspective with a focus on the instrumental value of trust for co-operation.

Hardin, Russell, ed. *Distrust*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi. Vol. VIII, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2004.

Anthology comprising of interdisciplinary contributions on the concept and the implications of distrust in politics and civil society.

Kramer, Roderick M., and Karen S. Cook, eds. *Trust and Distrust in Organizations*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2009.

Assesses the role of trust and distrust in a variety of different settings often characterized through power differences. Various case studies on trust and distrust relationships between employers and employee, doctors and patients, political leaders and their administration, as well as in geographically dispersed or virtual teams.

Ostrom, Elinow, and James Walker, eds. *Trust and Reciprocity*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi. Vol. VI, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2005.

Influential and useful review of behavioral research on trust and reciprocity. Contributors from different disciplines, but strong focus trust research in economics.

## ACCOUNTS OF TRUST

A contentious issue in theories of trust concerns the question whether trust is a belief, an expectation, an attitude or an emotion. Depending on their answer to this question different theoreticians have proposed cognitive, will-based or affective accounts of trust. Cognitive accounts, such as the one proposed by Gambetta 1988 consider trust to be a form of belief, a particular expectation with regard to the likely behavior of someone else. Sometimes also labeled risk-assessment or rational-choice accounts of trust these relate trust to cooperation and are often based upon Coleman's 1990 seminal work on the relationship between individual rational choice and collective action. Critics challenging these purely cognitive accounts have argued that they fail to distinguish trust from mere reliance. Yet this distinction appears crucial and warranted through our emotional reactions: in cases of misplaced trust we feel betrayed and not merely disappointed, which would be the adequate response to failed reliance. Accordingly, different theoreticians proposed various factors X to distinguish trust from mere reliance. Hardin 2002 argues that it is *encapsulated interest* that distinguishes trust from mere reliance, while Baier, a strong opponent of cognitive accounts (Baier 1986 & Baier 1991), proposes a *will-based account* of trust arguing that trust depends not merely upon expectations, but on beliefs about the competence and the good-will of the trustee. Holton 1994 proposes a *participant stance* as what distinguishes trust from mere reliance, while Jones 1996 and Lahno 2001 both challenge that trust is a form of belief altogether and propose *affective accounts* of trust depicting trust as an attitude and emotion.

Baier, Annette C. "Trust and Antitrust." *Ethics* 96.2 (1986): 231-60.

In this seminal paper, Baier offers a will-based account of trust and takes trust to be reliance upon competence and good-will of others. Addressing the potential power differences in trust relations, she criticizes moral philosophy's obsession with the prisoner's dilemma as a reductionist account of trust relations.

Baier, Annette C. "Trust and Its Vulnerabilities & Sustaining Trust." Paper presented at the Tanner Lectures on Human Values, Salt Lake City, 1991. Available online at: [tannerlectures.utah.edu/lectures/documents/baier92.pdf](http://tannerlectures.utah.edu/lectures/documents/baier92.pdf)

In the Tanner Lectures Baier sharpens her account of trust. Emphasizing that in trusting we accept being vulnerable. In the first lecture, she focuses the dangers of trust, while in lecture 2 she addresses the problem of providing general rules for trust.

Coleman, J. S. *Foundations of Social Theory*. Boston: Harvard University Press, 1990.

Conceptual basis of many cognitive accounts of trust and a landmark in sociological theory. Combines individual rational choice with a sociological conception of collective action.

Gambetta, Diego. "Can We Trust Trust?" In *Trust. Making and Breaking Cooperative Relations*, edited by Diego Gambetta, 213-37. Oxford: Basil Blackwell, 1988.

Frequently cited paper by one of the most important proponents of a rational account of trust. Defines trust and distrust as subjective probabilities with which an agent assesses that another agent will perform a particular action. Concluding essay of his important anthology. (See also \*Anthologies\*)

Hardin, Russell, ed. *Trust and Trustworthiness*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi. Vol. IV, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2002.

Proposes the encapsulated-interest view of trust, arguing that while trust involves expectation, this expectation must be grounded in the trustee's concern with the truster's interest.

Holton, Richard. "Deciding to Trust, Coming to Believe." *Australasian Journal of Philosophy* 72.1 (1994): 63-76.

Challenging Baier's 1986 account, Holton argues that a) trust can be willed, i.e. that we can sometimes decide to trust and b) that what distinguishes trust from mere reliance is not goodwill, but a participant stance towards the trustee.

Jones, Karen. "Trust as an Affective Attitude." *Ethics* 107 (1996): 4-25.

Defends an affective account of trust as an attitude of optimism in the goodwill and competence of a trustee who is moved favorable by our trust. This optimism that is not based upon beliefs about the others trustworthiness, but in an affectively loaded way of seeing the trustee.

Lahno, Bernd. "On the Emotional Character of Trust." *Ethical Theory and Moral Practice* 4.2 (2001): 171-89.

Argues that trust is an emotion, to some extent independent from rational deliberation. Trust as opposed to reliance requires a participant attitude and shared values and norms between truster and trustee.

## TRUST AND TRUSTWORTHINESS

Trust and trustworthiness are two concepts that are naturally connected. Despite the fact that trust appears to be a positively charged term, the relationship between trust and trustworthiness marks its ambivalence. Given the dangers of misplaced trust, of exploitation, lies and betrayal, ideally, we should not simply trust, but trust those who are worthy of our trust. Trustworthiness as a characteristic of the trustee should be of as vital concern to philosophers as trust is. Yet, the topic of trustworthiness remains often implicit and has received less attention than the concept of trust. The following contributions are notable exceptions in that they either focus specifically on trustworthiness or on the relationship between trust and trustworthiness. In a chapter of his influential book, Hardin 2002 assesses the dispositions, motivations and constraints of trustworthiness based upon an earlier paper on trustworthiness (Hardin 1996). Potter 2002 and Daukas 2006 both offer virtue theoretical accounts of trustworthiness, while Pettit 1995 introduces the notion of trust-reponsiveness as a rather more neutral alternative to virtue-theoretical accounts of trustworthiness. Moreover, coming from a feminist perspective, Potter 2002 and Daukas 2006 both emphasize the role of power in trust relations and the social dimension of credibility assessments. Since trustworthiness is often not directly observable, social cues are often used as proxies for trustworthiness as is demonstrated also in Gambetta and Hamill 2005's case studies on how taxi drivers in New York and Belfast assess the trustworthiness of their prospective customers. That social cues are often biased is addressed in the section on \*The Darker Side of Trust & Trustworthiness: Authority, Credibility & Bias\*.

Daukas, Nancy. "Epistemic Trust and Social Location." *Episteme* 3.1-2 (2006): 109-24.

In this feminist account epistemic trustworthiness is characterized as a social epistemic virtue that a) depends upon appropriate attitudes towards others and oneself as epistemic agents and b) requires character traits and skills that are developed over time in social practices. Daukas proposes an epistemic principle of charity. (See also \*The Darker Side of Trust & Trustworthiness: Authority, Credibility & Bias\*).

Gambetta, Diego, and Heather Hamill. *Streetwise. How Taxi Drivers Establish Their Customer's Trustworthiness*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi. Vol. X, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2005.

Intriguing empirical study on how taxi drivers in Belfast and New York assess the trustworthiness of their prospective clients by screening for signs that correlate reliably with trust- or distrust warranting properties of those clients. Relates game theory to signaling theory.

Hardin, Russell. "Trustworthiness." *Ethics* 107 (1996): 26-42.

Early paper on trustworthiness by one of the most prolific writers on trust.

Hardin, Russell. *Trust and Trustworthiness*. Edited by Karen S. Cook, Russell Hardin and Margaret Levi. Vol. IV, Russell Sage Foundation Series on Trust. New York: Russell Sage Foundation, 2002.

Hardin assesses the dispositions, motivations and constraints of trustworthiness in chapter 2 of this important book, which also introduces his account of trust as encapsulated interest and serves as a good overview on important topics, such as distrust, the epistemology of trust as well as trust in government and society.

Pettit, Philip. "The Cunning of Trust." *Philosophy and Public Affairs* 24 (1995): 202-25.

Distinguishes trust-reponsiveness from trustworthiness as a more neutral form of trust-reliability based upon people's love of standing in the eyes of others. Highlights the implications of this distinction for institutional design.

Potter, Nancy Nyquist. *How Can I Be Trusted? A Virtue Theory of Trustworthiness*. Lanham: Rowman & Littlefield, 2002.

Develops an account of trustworthiness around the virtue of giving uptake. Decidedly feminist perspective with a focus on power asymmetries. Draws on case studies in crisis counseling and teaching, targeting also practitioners.

### **Biases and Injustices in Trusting and Assessing Trustworthiness**

We should only trust those who are trustworthy. However, following this rule would require us to reliably identify those who are worthy of our trust which is not necessarily the case. First, we can trust those who are not trustworthy, the classic problem of trust, which is inherent in many accounts highlighting the downsides and dangers of trust especially in epistemological accounts of trust. (See also \*Trust & Testimony\*). The less obvious problem is, that we may not trust those who are trustworthy. Call this the beta-error of credibility assessment: we distrust, when we could and should have trusted. This epistemically and ethically problematic behaviour relates to biases in our credibility assessment of others and it has been feminist philosophers in particular, who have highlighted these problems. Fricker's 2007 book "Epistemic Injustice" is probably the most discussed piece in contemporary philosophy focusing on such failures in credibility assessment. Origi 2012 and Marsh 2011 both respond to Fricker by attempting to expand Fricker's analyses. While Scheman 2001 and Daukas 2006 both criticize the improper use of social indicators for credibility assessments, Alcoff 2001 argues that there are instances in which social identity may indeed be epistemically relevant. Rolin 2002 explicitly addresses the problems of credibility assessment in science (See also \*Trust in Science, Trust within Science\*).

Alcoff, Linda Martin. "On Judging Epistemic Credibility: Is Social Identity Relevant?" In *Engendering Rationalities*, edited by Nancy Tuana and Sandra Morgen, 53-80. Albany: SUNY Press, 2001.

Challenges interchangeableness model of epistemic agency and argues that social identity can sometimes be epistemically relevant for assessing epistemic authority and credibility.

Daukas, Nancy. "Epistemic Trust and Social Location." *Episteme* 3.1-2 (2006): 109-24.

Proposes the principle of epistemic charity to counter the ethical as much as epistemological problem of epistemic exclusion due to unwarranted withholding of epistemic authority. Characterizes epistemic trustworthiness as a social epistemic virtue. (See also \*Trust and Trustworthiness\*)

Fricker, Miranda. *Epistemic Injustice. Power and the Ethics of Knowing*. Oxford: Oxford University Press, 2007.

In this highly influential work, Fricker introduces the notions of testimonial and hermeneutic injustice to describe wrongs in the attribution of credibility. Relates epistemology to ethics and stresses inequalities of power in the credibility economy.

Marsh, Gerald. "Trust, Testimony, and Prejudice in the Credibility Economy." *Hypatia* 26.2 (2011): 280-93.



Departing from Fricker's notion of epistemic injustice, Marsh argues for a broader notion of trust injustice. The detrimental effects of trust injustice on the self-respect and self-worth of its victims demand the development of trust justice as a virtue to counter such trust injustices.

Origi, Gloria. "Epistemic Injustice and Epistemic Trust." *Social Epistemology* 26.2 (2012): 221-35.

Broadens the scope of testimonial injustice in response to Fricker 2007. Demands epistemic responsibility and vigilance to become aware of own cognitive heuristics and biases in information filtering.

Rolin, Kristina. "Gender and Trust in Science." *Hypatia* 17.4 (2002): 95.

Argues that credibility may not reliably reflect trustworthiness because trustworthiness is not transparent. While the latter is a characteristic of potential trustees, the former refers to the truster's perception of it.

Scheman, Naomi. "Epistemology Resuscitated: Objectivity as Trustworthiness." In *Engendering Rationalities*, edited by Nancy Tuana and Sandra Morgen, 23-52. Albany: SUNY Press, 2001.

Argues that while epistemic dependence is unavoidable, epistemic trust can be more or less justified. Following Shapin 1994 (cited under \*Trust in Science, Trust within Science\*) she stresses the relationship between social status and credibility attribution. Puts emphasis on normative implications for countering the trust-eroding effects of social, political, and economic injustice.

### **Trust without evidence**

If we need to have positive evidence for someone's trustworthiness before trusting them, how can new trust relationships be established in the first place? Moreover, is it not possible that trust can elicit trustworthiness? Trust, it seems, does not have to be justified by prior evidence. We can sometimes make a leap into the future and postpone the evaluation of trustworthiness. Both Möllering 2006 and McGeer 2008 argue that it can be rational to trust even in the absence of evidence of trustworthiness. Uslaner 2002 even argues that we have a moral commandment to treat strangers as if they were trustworthy. This performative dimension of trust has already been noted in one of the earliest papers on trust by Horsburgh 1960, who introduces the concept of therapeutic trust to label trust which aims at eliciting or increasing the trustworthiness of those trusted.

Horsburgh, H. J. N. "The Ethics of Trust." *Philosophical Quarterly* 10, (1960): 343-354.

Early text on the ethics of interpersonal trust. Introduces the concept of therapeutic trust – as opposed to innocent and guilty trust - to describe trust which aims at increasing the trustworthiness of those in whom it is reposed.

McGeer, Victoria. "Trust, Hope, and Empowerment." *Australasian Journal of Philosophy* 86.2 (2008): 237-54.

Argues that hope as a distinctive state of mind plays a crucial role for initiating and sustaining trust and that hopeful trust is not irrational although this does not imply that it necessarily pays off.

Möllering, Guido. *Trust: Reason, Routine, Reflexivity*. Oxford: Elsevier, 2006.

Emphasizes the performative aspects of trust. In order to enter trust relationships we have to suspend feelings of vulnerability and uncertainty. Instead we take a leap of faith and act as if we trusted.

Uslaner, Eric M. *The Moral Foundations of Trust*. Cambridge: Cambridge University Press, 2002. Challenges purely rational accounts of trust based upon game theory and argue that trust is important because of its a moral dimension. Distinguishing strategic trust as confidence in acquaintances from moralistic trust in strangers, he takes the latter as a moral commandment to treat strangers as if they were trustworthy.

## **BEYOND INTERPERSONAL TRUST**

Since definitions of trust are contingent to the specificities of the trust relation in question, we need to take a look also at trust relations apart from those between two individuals. Two forms of trust beyond interpersonal trust are trust in oneself as well as trust in society, in institutions and specifically trust in governments.

### **Self-Trust**

In recent years, philosophers have started addressing self-trust as a very particular trust relation. While both Lehrer 1997 and Foley 2001 put the notion of self-trust at the foundation of their epistemological theories, Govier 1993 assesses the intrinsic value of self-trust as well as its instrumental value for personal autonomy and self-esteem.

Foley, Richard. *Intellectual Trust in Oneself and Others*. Cambridge: Cambridge University Press, 2001.

Puts the notion of intellectual trust at the center of epistemological inquiry. Argues that an intellectual leap of faith is needed to get any philosophical inquiry off the ground if we accept that there are no epistemic foundations, no guarantees for the accuracy of our beliefs. Argues that self-trust precedes trust in others and can justify the latter.

Govier, Trudy. "Self-Trust, Autonomy, and Self-Esteem." *Hypatia* 8.1 (1993): 99-120.

Explores the concept of self-trust as a necessary condition of personal autonomy and self-respect.

Lehrer, Keith. *Self-Trust. A Study of Reason, Knowledge, and Autonomy*. Oxford: Clarendon Press, 1997.

Argues for a mathematical loop of trustworthiness and proposes self-trust as the basis of our trustworthiness on which reason, knowledge, and wisdom are grounded.

### **Trust in Society, Institutions, and Governments**

Social trust and trust in politics has emerged as an important topic in recent years and accordingly manifold are the accounts proposed and case studies analyzed in political philosophy and ethics as well as in economics, social and political science. In this section only some of the most prominent and influential accounts of trust beyond persons will be portrayed. However, more material on trust in governance or society can be found in the sections on anthologies (\*see Anthologies\*) and on overview (\*see General Overviews\*). In his seminal book Fukuyama 1995 argues for the economic and social value of trust by showing that poor economic performance and its social implications result from a lack of trust in certain countries. Szomptka's 1999 influential account of trust is based upon his analyses on the role of trust for the emergence of a

post-communist social order in Poland. In agreement to the benefits of trust, Seligman 1997 diagnoses a decline in such trust in America. In applied ethics, more specific institutions have been investigated. While O'Neill 2002 explores problems of trust and distrust in health care and politics, Meijboom 2008 focuses on trust problems in the food and agricultural sector. Skyrms 2008 challenges the basic assumption that trust in society may be needed for cooperation using simulations based upon rational-choice theories. Opposing such rational-choice models, Becker 1996 proposes a non-cognitive account of trust in politics.

Becker, Lawrence C. "Trust as Noncognitive Security About Motives." *Ethics* 107.1 (1996): 43-61.  
Against the dominance of cognitive accounts of trust, Becker proposes a noncognitive account of trust in governance. He distinguishes three forms of trust, i.e. credulity, reliance and security, the last one being the most problematic.

Fukuyama, Francis. *Trust. The Social Virtues and the Creation of Prosperity*. New York: Free Press Paperbacks, 1995.

Seminal book on trust and its relation to social and economic prosperity. Focus on instrumental value of trust for economies, comparison of high-trust and low-trust societies.

Meijboom, Franck L. B. *Problems of Trust: A Question of Trustworthiness*. Utrecht: Utrecht University Press, 2008.

Examines public trust in the European agricultural and food sector, attempting to disentangle trust from a risk-discourse. Emphasizes the need for trustworthiness and institutional integrity to strengthen the relationships between food producers and consumers.

O'Neill, Onora. *Autonomy and Trust in Bioethics*. Cambridge: Cambridge University Press, 2002.  
Explores problems of diminished trust in health care and health politics. Criticizes obsessive focus on individual autonomy and informed consent in bioethics.

Seligman, Adam B. *The Problem of Trust*. Princeton: Princeton University Press, 1997.

Important book on the relationship between trust and civil society offering historical and religious perspective on trust. Takes trust to be a form of belief in the goodwill of the other under conditions of opaqueness of their intentions and calculations.

Skyrms, Brian. "Trust, Risk, and the Social Contract." *Synthese* 160 (2008): 21-25.

Based upon rational choice and evolutionary game theory, it is argued that trust is not needed for cooperation, but merely a modest capacity for learning.

Sztompka, Piotr. *Trust. A Sociological Theory*. Cambridge: Cambridge University Press, 1999.

Influential account of trust based upon empirical research. Examines the emergence of a postcommunist social order in Poland in the late 1990ies.

## **TRUST AND KNOWLEDGE**

In his classic article "The role of trust in knowledge", Hardwig 1991 asserts that for most epistemologists trusting and knowing are deeply antithetical, when we trust, we don't know and when we know, we don't need to trust. Any observation of epistemic practices in science - as much as in everyday life as Webb 1993 notes - should have taught them better: Without trust we

would neither know basic facts of our life, such as when and where we are born, nor could we have achieved advanced scientific knowledge. Hence, although the topos of epistemic autonomy is still powerful, the topic of trust and its relevance for knowledge has started to receive increasing attention also in epistemology and philosophy of science.

Hardwig, John. "The Role of Trust in Knowledge." *The Journal of Philosophy* 88.12 (1991): 693-708.

Seminal paper by Hardwig often regarded as the starting point for inquiries into the role of trust in for scientific knowledge. A must-read.

Webb, Mark Owen. "Why I Know About as Much as You: A Reply to Hardwig." *The Journal of Philosophy* 90.5 (1993): 260-70.

In his reponse to Hardwig 1991, Webb argues that not only science rests on trust, but also our everyday epistemic practices. Acknowledging how little we would know if we depended solely on our own immediate experience, he argues that our trust in knowledge is grounded in the necessity for trust in language acquisition and use.

### **Trust and Testimony**

Within epistemology, analyses of trust are closely related to the topic of testimony.

Adler 1994 offers a good starting point to the topic and raises many important questions concerning the relationship between trust and knowledge. Origgi 2004 provides a useful overview over different reductionist and anti-reductionist accounts of testimony and proposes a non-reductionist account herself that stresses the pragmatics of communication. Hinchman 2005 also defends a non-reductionist accounts focusing on the act of telling as an invitation to trust. Fricker 1995 defends a reductionist position, arguing that to be justified in our testimonial beliefs, we need to monitor for signs of trustworthiness. Faulkner's 2011 recent book summarizes his theses on trust and serves as a pointer to his earlier articles, such as Faulker 2007, where he proposes an affective account of trust. For both McMyler 2011 and Keren 2007, the concept of epistemic authority is central to understand testimony and the role of trust for knowledge. Readers interested in this topic may want to take a look at the bibliography on *\*Epistemology of Testimony*[obo-9780195396577-0119.xml ]\*).

Adler, Jonathan. "Testimony, Trust, Knowing." *The Journal of Philosophy* 91.5 (1994): 264-75.

Excellent overview over many important issues regarding the relationship between trust and testimony, such as: the need to distinguish whether knowledge originates in testimony or is sustained by it, the importance of corroboration through background knowledge, the power of constraints to be truthful and reliable, the context-sensitivity and embeddedness of trusting, the difference between trustworthiness as a state or a trait (see *\*Trust and Trustworthiness\**).

Faulkner, Paul "On Telling and Trusting." *Mind* 116.464 (2007): 875-902.

Argues that our reasons for accepting testimony are not supplied by beliefs about the testifiers trustworthiness or empirical evidence but rather by an attitude of affective trust.

Faulkner, Paul. *Knowledge on Trust*. Oxford: Oxford University Press, 2011.

Faulkner's novel account of trust aims at reconciling reductionist and non-reductionist theories of testimony. Emphasizes internalized norms of conversational trust that are

socially shared and sustained by emotions. Not an easy read but rewarding for those interested in the epistemology of testimony.

Fricker, Elizabeth. "Telling and Trusting: Reductionism and Anti-Reductionism in the Epistemology of Testimony." *Mind* 104.414 (1995): 393-411.

Fricker defends a reductionist position in this classic paper on trust and testimony.  
Considers monitoring for signs of trustworthiness necessary to justify testimonial belief.

Hinchman, Edward S. "Telling as Inviting to Trust." *Philosophy and Phenomenological Research* 70.3 (2005): 562-87.

Proposes an anti-reductionist account of testimony by arguing that unlike mere assertions, acts of telling give reasons directly by inviting the hearer's trust.

Keren, Arnon. "Epistemic Authority, Testimony and the Transmission of Knowledge." *Episteme* 4.3 (2007): 368-81.

Argues that to trust a speaker is to grant her epistemic authority. Taking her word means to accept her judgment instead of relying on other reasons available.

McMyler, Benjamin. *Testimony, Trust, and Authority*. Oxford: Oxford University Press, 2011.

Rejecting the notion of epistemic autonomy, McMyler argues that testimonial knowledge is epistemically dependent on second-personal relations towards others and involves trusting them for the truth. Focus on interpersonal relations of trust, responsibility and authority.

Origgi, Gloria. "Is Trust an Epistemological Notion?" *Episteme* 1.1 (2004): 1-12.

Gives a useful overview over different reductionist and non-reductionist positions.  
Defends a non-reductionist account of epistemic trust by emphasizing the pragmatics of verbal comprehension.

### **Trust in Science, Trust within Science**

The title of this section is meant to denote two related, but distinct topics: on the one hand – as trust within science, it refers to the necessity of scientists to trust each other. On the other hand, trust in science also refers to debates about public trust in science, to deference to authority and trust in experts. Hardwig's 1991 paper entitled "The Role of Trust in Knowledge" is the most influential article on trust in science. In an earlier paper, Hardwig 1985 already challenges epistemic individualism, arguing that deference to authority is sometimes rational for experts as much as for lay people. While for Hardwig this epistemic interdependence calls for moral trust, Blais 1987 uses game-theoretic considerations to argue that strategic trust may suffice to explain cooperation in science. Focusing on the novice/expert problem, Goldman 2001 also challenges Hardwig by arguing that trust in experts need not be blind and due to the availability of various strategies by which even novices can evaluate experts. Rolin 2002 finally criticizes Hardwig from a feminist perspective for conflating trustworthiness and credibility and for neglecting biases in credibility assessment. The relationship between the social and the epistemic status of testifiers is also central in Shapin's 1994 highly recommendable analyses on the role of Gentlemen's testimony in pre-modern science. Drawing upon these historical insights from 17th century England, Shapin 1995 also comments on the role of experts in our contemporary worlds. Finally, Whyte and Crease 2010 provide interesting case studies on trust and distrust between scientific experts and citizens.

Blais, Michael. "Epistemic Tit for Tat." *The Journal of Philosophy* 84.7 (1987): 363-75.

Challenging Hardwig 1985, Blais uses game-theoretic considerations to argue that strategic rather than moral trustworthiness suffices to explain collaborative behavior in science. Without denying moral virtues in science, he concludes that it may not be necessary to trust the moral character on one's peers. Rather one can rely on mechanisms to detect and punish defectors in the cooperation game.

Goldman, Alvin Ira. "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63.1 (2001): 85-110.

Goldman criticizes Hardwig for unnecessarily promoting blind trust. According to him even novices can assess the relative expertise of different experts through different epistemic strategies.

Hardwig, John. "Epistemic Dependence." *The Journal of Philosophy* 82.7 (1985): 335-49.

Hardwig challenges epistemic individualism and argues that rationality sometimes consists in deferring to epistemic authority and trusting others who have more expertise. Moreover, given the collective nature of most epistemic practices in science, he draws two conclusions unpalatable to most epistemic individualists: either one can know without possessing evidence or groups must be accepted as epistemic agents.

Hardwig, John. "The Role of Trust in Knowledge." *The Journal of Philosophy* 88.12 (1991): 693-708.

This seminal paper by Hardwig is often regarded as the starting point for inquiring the role of trust in science and a must-read. Challenging epistemic individualism, he claims that moderns knowers cannot be independent due to the collective nature of science. Since mechanism to detect defectors are never perfect, scientists have to trust in the character of their peers, i.e. in their competence, honesty and adequate epistemic self-assessment.

Rolin, Kristina. "Gender and Trust in Science." *Hypatia* 17.4 (2002): 95.

Criticizes Hardwig 1991 from a feminist perspective for not taking into account the problem that credibility may not reliably reflect trustworthiness. Asks the important question under what social and institutional conditions credibility reflects trustworthiness and nothing but trustworthiness. Stresses the need for reliable community practices to ensure the trustworthiness of scientific testimony. (See also \*The Darker Side of Trust & Trustworthiness: Authority, Credibility & Bias\*).

Shapin, Steven. *A Social History of Truth: Civility and Science in Seventeenth-Century England*. Chicago: University of Chicago Press, 1994.

Insightful historical analysis on the role of trust in the testimony of gentlemen in pre-modern science. Since the social status of being a gentlemen served as a proxy for the epistemic status of their testimony, Shapin's work has been widely received in feminist epistemology. (See also \*The Darker Side of Trust & Trustworthiness: Authority, Credibility & Bias\*).

Shapin, Steven. "Trust, Honesty, and the Authority of Science." In *Society's Choices: Social and Ethical Decision Making in Biomedicine*, edited by R. E. Bulger, E. M. Bobby and H. V. Fineberg, 388-408. Washington: National Academy Press, 1995.

Based upon historical analyses on the role of gentlemen's testimony in 17th century England in Shapin 1994, Shapin draws conclusions concerning the role of experts in our contemporary worlds. Instead of installing "ethical experts" in fields such as biomedicine, a culture of virtue needs to be revived so that the public can trust scientists who deserve this trust.

Whyte, Kyle Powys, and Robert P. Crease. "Trust, Expertise, and the Philosophy of Science." *Synthese* 177 (2010): 411-25.

Delivers interesting case studies about trust and distrust between lay people and experts in contemporary science. Promotes the utilization of local knowledge and a new role of philosophers of science as mediators and translators in cases of unjustified mutual distrust.

## **TRUST IN SOCIO-TECHNICAL SYSTEMS**

Although often neglected in standard accounts of trust in philosophy, the relationship between trust and technology has emerged as an increasingly important topic for philosophical inquiry. While this section encompasses contributions on trust in different technologies, respectively socio-technical systems, the next two sections on \*E-Trust\* and \*Trust in Wikipedia\* focus explicitly on information and communication technologies. A special issue edited by Taddeo offers numerous valuable contributions to the topic of trust in technology, her introduction serves as a pointer (Taddeo 2010). While Kiran and Verbeek 2010 promote an understanding of trust as confidence, Nickel et al. 2010 changes the perspective by focusing on the trustworthiness of technologies. The case of robots provides a particularly interesting case for trust in technology. Pagallo's 2010 paper serves as a primer to the important link between trust and responsibility in legal philosophy, while Coecklbergh 2011 proposes a phenomenological-social perspective on trust in robots. Carusi's 2008 contribution finally points to a very different conception of trust and technology: assessing the use of visualizations in research she shows how technology can serve as a medium to establish interpersonal trust.

Carusi, Annamaria. "Scientific Visualisations and Aesthetic Grounds for Trust." *Ethics and Information Technology* 10 (2008): 243-54.

Analyzes problems of trust among practitioners in computationally mediated research practices. Argues that visualizations contribute to an intersubjectively shared framework of agreement that can serve as a basis for trust.

Coeckelbergh, Mark. "Can We Trust Robots?" *Ethics and Information Technology* 14 (2011): 53-60.

Assesses the utility of different contractarian-individualist approaches to capture trust in robots and proposes a phenomenological-social alternative.

Kiran, Asle, and Peter-Paul Verbeek. "Trusting Our Selves to Technology." *Knowledge, Technology & Policy* 23.3 (2010): 409-27.

Insightful approach to re-describe the relationship between humans and technology as mutually constitutive. Proposes an understanding of trust in technology as confidence, as opposed to trust as mere reliance in extensionist theories of technologies or as suspicion in risk-assessment and precautionary approaches towards technology. Case studies on telemonitoring in healthcare and prosthetics.

Nickel, Philip, Maarten Franssen, and Peter Kroes. "Can We Make Sense of the Notion of Trustworthy Technology?" *Knowledge, Technology & Policy* 23.3 (2010): 429-44.

Provides a useful overview over different account of interpersonal trust and trustworthiness and compares them with definitions of reliability in engineering to assess whether and if so how trustworthiness can be meaningfully attributed to technology. Concludes that only a thin notion of trustworthiness can be attributed to technology. Useful concept of technology that takes socio-technical systems into account. (See also \*Trust and Trustworthiness\*)

Pagallo, Ugo. "Robotrust and Legal Responsibility." *Knowledge, Technology & Policy* 23.3-4 (2010): 367-79.

Ties the debates around trust in technology to issues of legal responsibility in the production and use of robots. Puts emphasis on new challenges in determining liability due to autonomous systems.

Taddeo, Mariarosaria. "Trust in Technology: A Distinctive and a Problematic Relation." *Knowledge, Technology & Policy* 24.3-4 (2010): 283-86.

Introduction to a special issue on trust and technology edited by Taddeo with numerous valuable contributions.

## **E-Trust**

The topic of E-trust as conceived here comprises of two different branches. Interpersonal trust online and trust in multi-agent systems. Ess 2010, Weckert 2005, Nissenbaum 2001 and Pettit 2004 address the former by asking whether, and if so how and to what extent, trust between human agents is possible in an online environment. Regarding trust in artificial agents or multi-agent systems, Taddeo 2010, Castelfranchi & Falcone 2010, Buechner & Tavani 2010 assess how trust can be conceived or modeled in artificial agents, while Grodzinsky et al 2010 focus on trust relations between human and artificial agents. While these accounts have a strong links towards computer science and artificial intelligence, their insights are nonetheless highly relevant for philosophical accounts of trust. For further contributions on e-trust see also \*Anthologies\*.

Buechner, Jeff, and Herman Tavani. "Trust and Multi-Agent Systems: Applying the "Diffuse, Default Model" of Trust to Experiments Involving Artificial Agents." *Ethics and Information Technology* (2010): 1-13.

The authors apply Walker's notions of default trust and default, diffuse trust in the context of e-trust and argue that experiments with artificial agents can further philosophical understanding of trust.

Castelfranchi, Cristiano, and Rino Falcone. *Trust Theory: A Socio-Cognitive and Computational Model*. Edited by Michael Woolridge, Wiley Series in Agent Technology. Chichester: John Wiley & Sons 2010.

Introduction to formal-based modeling of trust and its application in agent-based systems by two of the leading figures in the field. Their socio-cognitive model is put into an interdisciplinary context and has received recognition in philosophical accounts of e-trust.

Ess, Charles. "Trust and New Communication Technologies: Vicious Circles, Virtuous Circles, Possible Futures." *Knowledge, Technology & Policy* 23.3-4 (2010): 287-305.



Delivers a useful taxonomy of different philosophical accounts of trust. Argues that online trust is possible, despite various caveats and objections raised in virtue ethical and phenomenological accounts.

Grodzinsky, F., K. Miller, and M. Wolf. "Developing Artificial Agents Worthy of Trust: "Would You Buy a Used Car from This Artificial Agent?" Ethics and Information Technology (2010): 1-11.

Develops an object-oriented model for e-trust distinguishing three different philosophically relevant trust relations concerning artificial agents: trust of humans in artificial agents, trust amongst artificial agents and trust of artificial agents in humans.

Nissenbaum, Helen. "Securing Trust Online: Wisdom or Oxymoron." Boston University Law Review 81:3 (2001): 635-64.

Early and influential paper criticizing the vision of trustworthiness as security, or trust through security. Proposal to aim at nourishing instead of securing trust online.

Pettit, Philip. "Trust, Reliance and the Internet." Analyse und Kritik 26, (2004): 108-121.

Distinguishes trust from reliance and argues that at least amongst people not known otherwise to each other the internet offers merely opportunities for reliance, but not for trust.

Taddeo, Mariarosaria. "Modelling Trust in Artificial Agents, a First Step toward the Analysis of E-Trust." Minds and Machines 20.2 (2010): 243-57.

Non-psychological account of e-trust based upon a Kantian regulative ideal of a rational agent by one of the most important contributors to the topic of e-trust.

Weckert, John. "Trust in Cyberspace." In The Impact of the Internet on Our Moral Lives, edited by Robert J. Cavalier, 99-120. Albany: University of New York Press, 2005.

Account of e-trust that seeks to ground e-trust in an agent's attitude. Based upon Annette Baier's claim that trust is not only cognitive, but at the same time an affective and conative phenomenon.

### **Trust in Wikipedia**

Given its prominence as a major source of information, Wikipedia has recently attracted some attention by philosopher. Both Tollefsen 2009 and Magnus 2009 assess Wikipedia as a source of testimony asking whether and if so how readers are justified in accepting content provided by Wikipedia as testimony (see also \*Trust and Testimony\*). De Laat 2012 emphasised the role of editorial policies to modulate levels of trust in contributors, while Simon 2010 stresses the relevance of transparency for epistemic trust in the use of Wikipedia and recommender systems.

de Laat, Paul B. "Open Source Production of Encyclopedias: Editorial Policies at the Intersection of Organizational and Epistemological Trust." Social Epistemology 26.1 (2012): 71-103.

de Laat shows how trust in contributors is pre-formatted very differently through the editorial policies of six online encyclopedias. Compares the levels of discretion, the role of experts and reputation, he concludes that both non-reductionist and reductionist views of testimony have been implemented in editorial strategies.

Magnus, P. D. "On Trusting Wikipedia." Episteme 6.1 (2009): 74-90.

Applies five common strategies to assess the reliability of testifiers, such as plausibility assessments, appeal to authority, calibration and sampling, to Wikipedia. Concluding that these strategies do not work properly for Wikipedia, Magnus encourages the development of new epistemic methods and strategies for these new contexts.

Simon, Judith. "The Entanglement of Trust and Knowledge on the Web." *Ethics and Information Technology* 12.4 (2010): 343-55.

Considers trust in Wikipedia and recommender systems to be a form of procedural trust and emphasizes the need for transparency as a crucial requirement for epistemic responsibility and vigilant behavior. Proposes a new role of philosophers to assess and guide systems design.

Tollefsen, Deborah Perron. "Wikipedia and the Epistemology of Testimony." *Episteme* 6.1 (2009): 8-24.

Departing from a non-summative account of group testimony, Tollefsen asks whether Wikipedia is a source of testimony and if so what its exact nature is. Treating Wikipedia as an epistemic child, Tollefsen stresses the corroboration of testimony through background knowledge for the assessment of Wikipedia entries.