Journal of Neuroscience, Psychology, and Economics
2012, Vol. 5, No. 2, 63–91

1937-321X/11/$12.00 DOI: 10.1037/a0026318

The Biology of Trust: Integrating Evidence From Genetics, Endocrinology, and Functional Brain Imaging

René Riedl
University of Linz, Austria

Andrija Javor
Linz General Hospital, Austria

Trust is among the most important factors in human life, as it pervades almost all domains of society. Although behavioral research has revealed a number of insights into the nature of trust, as well as its antecedents and consequences, an increasing number of scholars have begun to investigate the topic from a biological perspective to gain a deeper understanding. These biological investigations into trust have been carried out on three levels of analysis: genes, endocrinology, and the brain. Based on these three levels, we present a review of the literature on the biology of trust. Moreover, we integrate our findings into a conceptual framework which unifies the three levels of analysis, and we also link the biological levels to trust behavior. The results show that trust behavior is at least moderately genetically predetermined. Moreover, trust behavior is associated with specific hormones, in particular oxytocin, as well as specific brain structures, which are located in the basal ganglia, limbic system, and the frontal cortex. Based on these results, we discuss both methodological and thematic implications.

Keywords: trust, biology, hormones, fMRI, oxytocin

Determining whom to trust and whom not to trust has been crucial since the early days of ancient civilizations. Trust in individuals who turned out not to be trustworthy resulted in death in many situations from the Stone Age to the Middle Ages. However, although in today's world the possible consequences of trusting untrustworthy people are not as directly related to survival as they were in former times, breached trust may result in severe negative feelings and/or economic consequences—imagine, for example, the spouse who is cheated in marriage or an online shopper whose product, although already paid for, is not delivered.

This article was published Online First November 28, 2011.

René Riedl, Department of Business Informatics–Information Engineering, University of Linz, Linz, Austria; Andrija Javor, Department of Neurology and Psychiatry, Linz General Hospital, Linz, Austria.

We thank Bernd Weber for his work in providing guidance on ways to improve this article. We also acknowledge the excellent comments during the review process. Finally, we are grateful to the participants of the Gmunden Retreat on NeuroIS 2011 (see www.NeuroIS.org) for their thoughtful comments on several parts of this review.

Correspondence concerning this article should be addressed to Dr. René Riedl, University of Linz, Department of Business Informatics–Information Engineering, Altenberger Strasse 69, 4040 Linz, Austria. E-mail: rene.riedl@jku.at

Economic research has shown that low trust, as the result of repeatedly breached trust, leads to a low rate of investments, which in turn impedes new businesses and employment (Zak & Fakhar, 2006). As a consequence, low trust countries are typically poor countries, and trust is, therefore, among the strongest predictors of poverty identified by economists (Zak & Knack, 2001). Overall, trust is among the most important factors in human life, as it pervades almost all domains of society, ranging from love, friendship, economic collaborations, medicine, and politics, to the steadily increasing number of human interactions in anonymous virtual Internet environments (Falk & Kosfeld, 2006; Füllbrunn, Richwien, & Sadrieh, 2011; Gefen, Benbasat, & Pavlou, 2008; Lee & Lin, 2009; Luhmann, 1979; Nguyen et al., 2009; Smith, 2010).

Against the background of the ubiquity of trust in human societies, both in traditional and virtual environments, a vast amount of research in several scientific disciplines has investigated the nature of trust, as well as its antecedents (e.g., facial expressions in personal interactions among humans or well-designed user interfaces in computer-mediated interactions) and consequences (e.g., information disclosure or cooper-

ative behavior). Psychologists, sociologists, economists, as well as management and information systems researchers, among scholars in other academic fields, have investigated trust, both theoretically and empirically (Gefen et al., 2008; Rousseau, Sitkin, Burt, & Camerer, 1998; Seppänen, Blomqvist, & Sundqvist, 2007; Swan, Bowers, & Richardson, 1999).

Despite the large variety of disciplines that have studied the phenomenon, there is a surprising consensus regarding the conceptualization of trust. Drawing upon the work of Fishbein and Ajzen (1975) and Coleman (1990), many scholars describe trust as a *behavior* which makes one party, the trustor, vulnerable to the actions of another party, the trustee. This behavior of a trustor is influenced by his or her *beliefs* about the trustee's trustworthiness, which in turn affect his or her *attitudes* toward the trustee and subsequent *behavioral intentions*. Moreover, trust behavior in a specific situation is also influenced by the trustor's general level of trust (i.e., trust disposition), as well as by an individual's risk preferences.

Important characteristics of a trustee are ability, benevolence, and integrity (Mayer, Davis, & Schoorman, 1995). If a trustor believes in the trustworthiness of a trustee, he or she believes that the trustee (a) has skills and competencies that are important for the relationship (ability), (b) means well toward the trustor aside from an egocentric profit motive (benevolence), and (c) adheres to a set of principles that the trustor finds acceptable (integrity). The importance of each characteristic may vary as a function of the inner states of a trustor (e.g., high risk perceptions regarding possible betrayal may turn one's attention to benevolence rather than ability) and context factors such as the importance of the trust decision for survival (e.g., in tandem skydiving, the ability of the instructor is expected to be more important than benevolence).

Figure 1 illustrates the structure of a trust situation. A person, the trustor, who meets an unknown individual, the trustee, for the first time, will form his or her beliefs about the individual's trustworthiness on the basis of stimuli (e.g., appearance, facial expression, or self-control of the trustee; Adolphs et al., 2005; Righetti & Finkenauer, 2011; Todorov, 2008; Winston, Strange, O'Doherty, & Dolan, 2002), which in turn affects the attitude toward the individual, subsequent behavioral intentions, and ultimately actual behavior (e.g., trust/distrust and approach/avoidance). Moreover, the trustor's disposition to trust, as well as his or her risk preferences, may moderate each linkage in this causal chain (e.g., Fehr, 2009a). (We

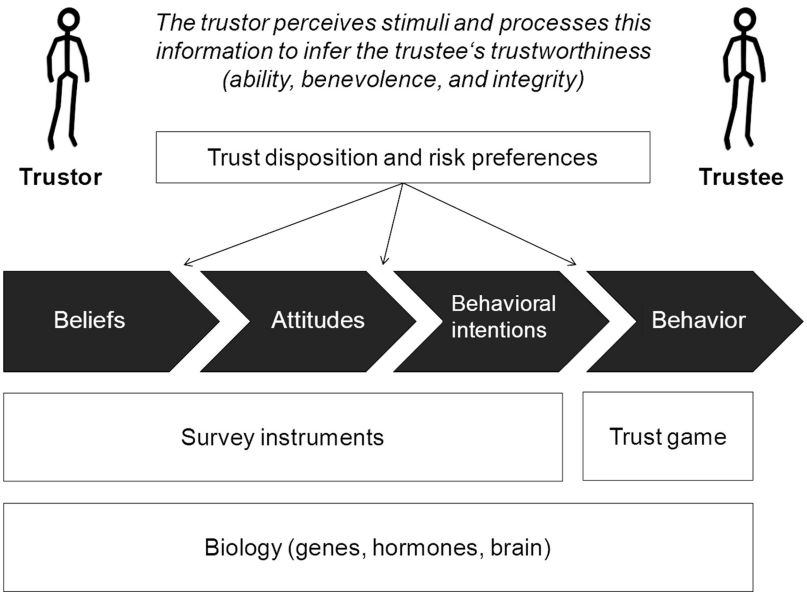

Figure 1. Structure of a trust situation.

describe the lower part in Figure 1 in the next section.)

Although behavioral trust research has revealed a number of insights into the nature of trust, as well as its antecedents and consequences (Rousseau et al., 1998; Seppänen et al., 2007; Swan et al., 1999), scholars have also started to investigate the topic from a biological perspective during the past decade. These biological investigations into trust can be classified into three groups, namely genetics (e.g., Cesarini et al., 2008), endocrinology (e.g., Kosfeld, Heinrichs, Zak, Fischbacher, & Fehr, 2005), and brain functionality (e.g., King-Casas et al., 2005). One reason for this development toward biologically oriented trust research is the availability of powerful methods such as functional MRI (fMRI). A further reason is the insight that all human behavior that varies among individuals is associated, at least partly, with biological factors, in particular those related to the nervous system (e.g., Cacioppo, Berntson, Sheridan, & McClintock, 2000; Turkheimer, 1998). Hence, the consideration of biological factors is indispensible to explain a large proportion of the variance in human trust behavior (Fehr & Camerer, 2007).

In the present article, we review the literature that has investigated trust from a biological perspective. Extensive investigation revealed a large number of studies on this topic. So far, a vast amount of brain imaging studies, mostly using fMRI, has revealed several brain regions associated with trust. Moreover, endocrinological studies show that a number of hormones affect trust. Finally, recent gene-based research has demonstrated that at least a moderate degree of human trust behavior is genetically predetermined.

Due to the large number of biologically oriented studies on trust that we identified, a more complete understanding of the research findings would be possible if the studies were integrated into a unifying framework. Therefore, in the present article we discuss existing research on the biological foundations of trust within a conceptual framework that distinguishes three levels of analysis, namely genetic, hormonal, and at the brain level. Also, we discuss how these biological factors influence trust behavior.

The remainder of this article is structured as follows: In section 1, we highlight important foundations of trust research. First, we present a model which we use to structure our review. Second, we discuss the measurement of trust. In section 2, we outline the biology of trust along three levels of analysis (genes, hormones, and the brain), and we make explicit that the influence of the genetic and hormonal levels on trust behavior is mediated by the brain level (i.e., activation in specific brain areas). Then, in section 3, we integrate the findings of the three levels of analysis into a conceptual framework. Finally, in section 4, we summarize our findings, present implications for future research, and provide concluding comments.

1. Foundations of Trust

Biology and Trust Behavior: A Model

The question of how *nature* and *nurture* contribute to the manifestation of human behavior, such as trust behavior, has been one of the most fundamental research issues in a number of scientific disciplines, particularly in psychology. Though there have been extremists who have believed that either the biological influences of nature (e.g., genes) or the environmental influences of nurture (e.g., socialization) predominate, today most scientists agree that both are important and neither is deterministic (Johnson, 2007). Empirical evidence substantiates the notion that human behavior is the result of the complex interplay between both biological and environmental factors (e.g., Bouchard, 1994; Cacioppo et al., 2000).

Figure 2 illustrates the relations between behavior, biology, and environment. It is shown that human behavior is influenced by (a) biological factors (genes, hormones, and the brain) and (b) environmental factors (e.g., socialization, culture, experience, and task demands). While human behavior is directly observable (e.g., one person trusts another one and therefore exhibits approach behavior), biological factors are typically not (e.g., activation in brain regions associated with trust is not directly observable). Moreover, the interrelationship between biological and environmental factors is illustrated by the gray arrows in Figure 2.

In this article, we review the biologically oriented literature on trust. We discuss the relation between biological factors (genes, hormones, and the brain) and trust behavior. Moreover, we outline the relations among the three biological factors in the trust context. Research

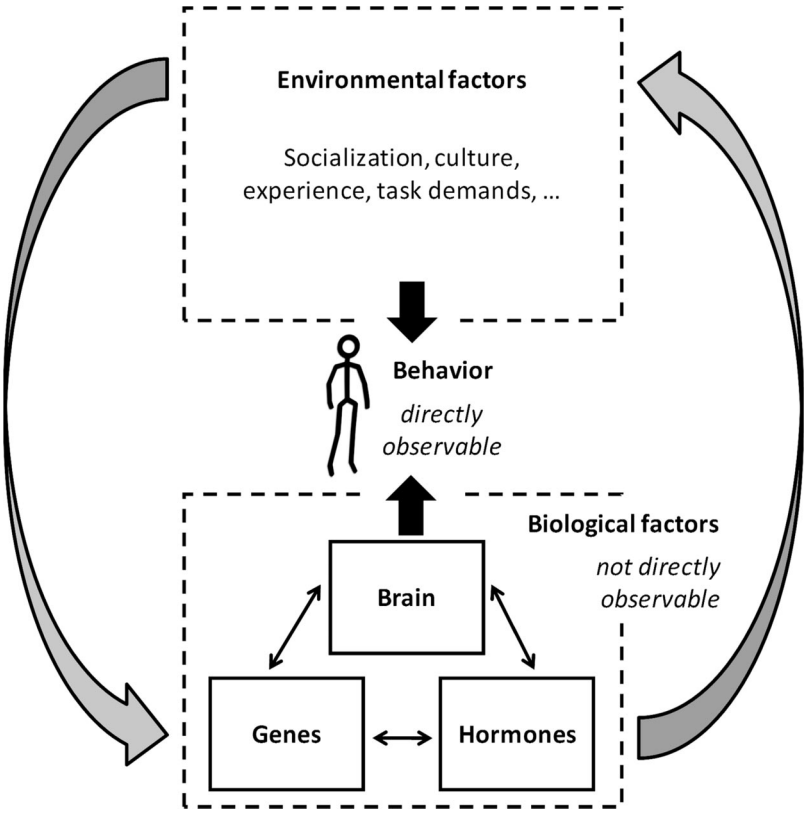

Figure 2. Model of behavior, biology, and environment.

(e.g., Cacioppo et al., 2000) indicates that genes, hormones, and the brain are interrelated components. For example, genes influence the production and release of hormones and the anatomy of the brain and its processing mechanisms. Moreover, hormones influence the activation of brain regions because certain regions have receptors for certain hormones. Finally, to state another example, the brain influences hormones because the brain regulates the production and release of hormones. The interrelationships among the biological factors are illustrated by the arrows in Figure 2. It is important to note that the influence of genes and hormones on behavior is mediated by activation in specific brain areas.

Note that in this article we do not discuss the direct influence of environmental factors on trust behavior (Zak & Fakhar, 2006). That is, we do not discuss topics such as the influence of socialization and culture on trust behavior. Such a discussion could be a fruitful topic for future studies (see, e.g., Bjornskov, 2007 and Welch et al., 2005).

The Measurement of Trust: Survey Instruments and The Trust Game

Biological research on trust is based on measurements that are often different from the conceptualizations used in the behavioral sciences (e.g., those used in psychology, economics, management science, and information systems research). Hence, to fully understand the biological foundations of trust, as well as their implications for trust behavior, it is important to outline the major conceptual differences.

Based on extensive reviews of the literature, research (McKnight & Chervany, 2001; Seppänen et al., 2007; Swan et al., 1999) has found that trust is often conceptualized as a belief, attitude, intention, or behavior (see Figure 1).

Such conceptualizations are used to develop survey instruments which make possible the measurement of trust. For example, if trust is conceptualized as a *belief*, the construct is operationalized along the characteristics of a trustee (e.g., ability, benevolence, and integrity), and three items to measure benevolence based on a Likert-type scale are, for example: (a) "I believe that the trustee would act in my best interest," (b) "If I required help, the trustee would do his or her best to help me," and (c) "The trustee is interested in my well-being, not just his or her own" (McKnight, Choudhury, & Kacmar, 2002, p. 355). If trust is conceptualized as an *attitude* or *behavioral intention*, other items, though similar to the three mentioned, are used to measure trust (McKnight et al., 2002).

Most behavioral studies measure trust by means of survey instruments. Thus, many studies do not measure actual trust behavior. Rather, antecedents of trust behavior (i.e., beliefs in the trustworthiness of a trustee, attitudes toward the trustee, and behavioral intentions) are measured. However, if trust is conceptualized as a behavior, it could be more advantageous to measure *actual behavior* (e.g., cooperative behavior) in a trust situation (e.g., in an economic game) rather than by asking people about their beliefs, attitudes, or intentions in a hypothetical trust setting (Fehr, 2009a).

Figure 1 (the lower part) shows that survey instruments have been used to measure trust beliefs, attitudes, and intentions, whereas the so-called trust game, an economic game, was developed to measure actual trust behavior. Moreover, it is illustrated in Figure 1 that all components of a trust situation (from perceptions of stimuli such as a trustee's facial expression to actual behavior) are based on biological factors, namely genes, hormones, and the brain (Cesarini et al., 2008; King-Casas et al., 2005; Kosfeld et al., 2005).

The *trust game* was developed to measure both trust and trustworthiness as actual behavior of players in an economic exchange game (Berg, Dickhaut, & McCabe, 1995; Camerer & Weigelt, 1988; McCabe, Rigdon, & Smith, 2003; McCabe & Smith, 2000).[1] In the mainstream version of the trust game, decision maker 1 (the trustor, DM 1) has an initial endowment of x monetary units (e.g., \$10, see Figure 3). First, DM 1 decides whether to share his or her endowment (e.g., \$5 for each player, then the game ends, see left path in Figure 3) or to send (a part of) it to DM 2 (the trustee) (right path in Figure 3). DM 2 observes DM 1's action and, if money was sent, decides whether to keep the amount or share (some of) it with DM 1. The experimenter triples DM 1's transfer, so that both players are better off collectively if DM 1 transfers money and DM 2 sends back a sufficient amount. Hence, in the example illustrated in Figure 3, DM 2 has two possibilities: Either to share the money (i.e., each player gets \$15) or to keep all the money (i.e., DM 1 gets \$0, DM 2 gets \$30).

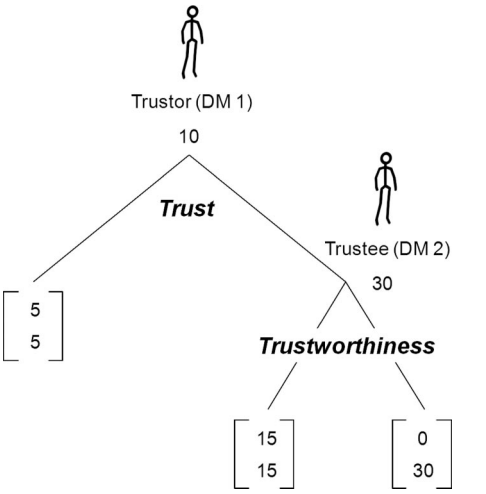

Figure 3. Structure of the trust game. *Notes:* The upper value in the square brackets indicates the trustor's payoff (decision maker 1, DM 1), the lower value the trustee's payoff (DM 2). The amounts of money and the payoff matrix in the figure are used as an example.

This game mimics a sequential economic exchange in the absence of contract enforcement institutions (Fehr & Camerer, 2007). If the game is played on a one-shot basis, DM 2 has a strong incentive to keep all the money and repay none to DM 1. If DM 1 anticipates this behavior, however, there is little reason to transfer. Consequently, if DM 1 transferred no money, then a chance for higher mutual gain would be lost. In the trust game, the amount sent by DM 1 is used as a behavioral measure for trust, and DM 2's transfer back is used as a behavioral measure for trustworthiness (see Figure 3).

[1] For a review of economic games used in neurobiological research, see Krueger, Grafman, and McCabe (2008).

The trust game has been used in many studies on the genetic, hormonal, and neurological foundations of human trust behavior. We review these investigations in the following, complemented by studies using other methods (e.g., surveys). Methodological discussions about the trust game can be found, for example, in Kugler, Connolly, and Kausel (2009) and Vilares, Dam, and Kording (2011).

2. The Biology of Trust

Genetic Level of Analysis

Survey research (Dohmen, Falk, Huffman, & Sunde, 2006) indicates that the willingness to trust other people is similar between parents and their children. This raises the question whether genetic or environmental influences, or both in conjunction, are responsible for this result. Genetic and environmental influences on behavior can be investigated either through quantitative or molecular genetic studies. Quantitative genetic methods, in contrast to molecular genetic methods, provide estimates of the relative magnitudes of omnibus genetic and environmental influences without the necessity of specifying the actual DNA sequences or environmental circumstances that provide those influences (Johnson, 2007).

To investigate whether humans are endowed with genetic variation that could account for individual differences in trust behavior, one study (Cesarini et al., 2008) applied a quantitative method based on the trust game (see Figure 3). Because monozygotic (MZ) twins share the same genes, whereas the genes of dizygotic (DZ) twins are only imperfectly correlated, MZ twins should exhibit a higher correlation in their behavior than DZ twins if genetic differences help explain the variance of trust behavior. In fact, the study of Cesarini et al. found a heritability estimate of (a) trust between 10% (Swedish subjects) and 20% (U.S. subjects) and (b) trustworthiness between 17% (U.S. subjects) and 18% (Swedish subjects). Considering these findings, Cesarini et al. (2008) concluded: "These results show that we consistently found a significant proportion of variance in trust is due to heritability" (p. 3723).

Another recent twin study (Sturgis et al., 2010) also investigated the genetic and environmental basis of trust. In contrast to the Cesarini et al. (2008) study, trust was not measured as actual behavior (trust game). Rather, a questionnaire consisting of four items was used to capture the participants' trust beliefs, namely (a) "I believe that most people are basically well-intentioned," (b) "I believe that most people will take advantage of you if you let them," (c) "I think that most of the people I deal with are honest and trustworthy," and (d) "My first reaction is to trust people." The results show that "the majority of the variance in a multiitem trust scale is accounted for by an additive genetic factor . . . the environmental influences experienced in common by sibling pairs have no discernable effect . . . [these] findings problematise the widely held view that the development of social trust occurs through a process of familial socialization at an early stage of the life course" (p. 205).

Scientific evidence demonstrates that nasally administered oxytocin (OXT, a hormone that also acts as a neurotransmitter in the brain) increases trust in humans (Kosfeld et al., 2005), highlighting the importance of this hormone for cooperative behavior (details on trust hormones are reported in the next section). Genetic research on OXT focuses on the oxytocin receptor (OXTR) gene, which (a) is located on chromosome 3p25 (Inoue et al., 1994) and (b) shows multiple variations (so-called single nucleotide polymorphisms, SNPs). These SNPs are related to differences in human perception and behavior (e.g., empathy and stress reactivity, Rodrigues, Saslow, Garcia, John, & Keltner, 2007, or autism, Wermter et al., 2010). Therefore, they have been hypothesized to play an important role in explaining individual differences in trust behavior (Reuter et al., 2009). To test this hypothesis, Reuter et al. conducted a trust game experiment with participants whose OXTR gene was screened. The results show that participants who have a particular variant of the OXTR gene exhibit more trust taking than those who exhibit the alternative variant of the gene. Reuter et al. (2009) concluded: "Our results indicate that individual differences in the proclivity to trust are influenced by variations in the OXTR gene" (p. 21).

Another study (Israel et al., 2009) also demonstrated that genetic polymorphisms for the OXTR gene are associated with human prosocial decision making behavior. In line with this finding, an investigation by Tost et al. (2010),

using a multimodal imaging intermediate phenotype approach, demonstrated that a genetic variant in the OXTR gene linked to social functions predicts individual differences in brain structure and functioning, as well as personality (e.g., reward dependency, a construct associated with trust; Delgado, Frank, & Phelps, 2005).

It is important to note that a recent study found *no* significant association between nine SNPs of the OXTR gene and behavior in the trust game (Apicella et al., 2010). However, Apicella et al. write that "[g]iven that our research design only allows us to statistically reject moderate to large effect sizes, the results reported here are not inconsistent with the results of hormonal association studies involving OXT in trust and generosity and do not necessarily rule out a role for OXTR polymorphisms in explaining phenotypic variation" (p. 7).

In addition to research on the OXTR gene, there is evidence that variants in the length of a promoter region in the arginine vasopressin 1a receptor (AVPR1a) gene predict altruistic behavior. Because altruism is defined as a selfless concern for the welfare of other individuals, and hence is strongly related to one component of trustworthiness, namely benevolence (Mayer et al., 1995), this research provides additional evidence that human trust behavior is to some extent genetically predetermined. The findings of a recent review (Skuse & Gallagher, 2011) on the genetic influences on social cognition, a concept closely related to trust (Fehr, 2009b; Krueger et al., 2007), support this notion. In particular, this review identifies genetic variation in the receptors associated with OXT and AVP as a major determinant of differences in human trust behavior.

Hormonal Level of Analysis

Gene expression is not only essential for the development of brain structures and hormones (Harris, Vernon, & Boomsma, 1998; Rushton & Ankney, 2007), but is itself influenced by hormones (Harlan, 1988). Hormone regulation occurs in the hypothalamus, which in turn affects activation in the pituitary gland (Harris, 1948). These findings support our unifying framework on the biology of trust which describes genes, hormones, and brain processes as interrelated factors (see Figure 2).

In the following, we review the literature that has investigated the effects of various hormones and neurotransmitters on human trust behavior. Specifically, we discuss oxytocin, arginine vasopressin, cortisol, dopamine, testosterone, estrogen, and serotonin.

Oxytocin. Oxytocin (OXT), a neuropeptide, has been shown to influence human social behavior (Heinrichs & Domes, 2008; Heinrichs, von Dawans, & Domes, 2009). In particular, OXT plays an important role in *pro*social behaviors such as parturition, lactation, maternal attachment, and pair bonding (Donaldson & Young, 2008). Hence, OXT is associated with approach and human trust behavior.

Anatomically, OXT is synthesized in magnocellular neurons of the paraventricular and supraoptic nuclei of the hypothalamus. Moreover, it is processed from its precursor form, together with the carrier protein, along the axonal projection to the posterior pituitary, from which the peptide is secreted into the systemic circulation. Also, OXT is distributed throughout the central nervous system from smaller parvocellular neurons, influencing many neurobehavioral functions (Heinrichs, Baumgartner, Kirschbaum, & Ehlert, 2003).

Based on empirical evidence which shows that OXT plays a prominent role in facilitating various social behaviors of animals (Insel & Young, 2001), Zak and colleagues formulated the hypothesis that OXT also affects human trust behavior. Using the trust game, Zak, Kurzban, and Matzneret (2004, 2005b) found two results: First, OXT levels are, on average, 41% higher in those subjects (DM 2 in Figure 3) who receive a monetary transfer that reflects an intention of trust relative to a random monetary transfer of the same amount. Second, subjects who receive an intentional trust signal return, on average, 53% of the amount they received to the gaming partner, whereas in the random monetary condition the mean amount returned is only 18%. Thus, the perception of a signal of trust increases OXT levels, which in turn cause trustworthy behavior (i.e., the reciprocation of trust). In other words, the studies by Zak and colleagues show that when people are trusted, their brains release OXT, which in turn predicts trustworthiness. Thus, trusting behaviors depend significantly on the endogenous release of OXT.

These results suggest that trust put in another individual is likely to pay off, namely via reci-

procity. But why do humans reciprocate trust even when they interact with strangers who they will never meet again? Recent explanations to understand this phenomenon are based on the concept of evolutionary psychology. This theorizing (Lieberman & Eisenberger, 2009) is based on the observation that "our emotional responses to [complex social] events rely on much of the same neural circuitry that underlies the simplest physical pains and pleasures" (p. 890) and "the brain may treat abstract social experiences [such as trust] and concrete physical experiences as more similar than is generally assumed" (p. 891).

This observation can be explained by two evolutionary phenomena (Lieberman & Eisenberger, 2009): First, mammalian newborns depend on other humans for survival, because they are relatively immature at birth. Hence, the survival of a newborn, and thus the survival of mankind in general, depends on the social bond between a newborn and other humans, which implies trust. Second, the division of basic human activities among group members (e.g., one individual takes care of food acquisition, another one of protection from adversarial groups, and another one of care for offspring) has turned out to guarantee survival, while a failure to divide the various human activities did not. This division of activities among group members, however, is associated with cooperative behavior, which also implies trust. Therefore, trust contributes to survival from an evolutionary perspective. In line with this argumentation, a recent study by De Dreu et al. (2010) found that intranasal administration of OXT modulates trust toward members of one's own group, but not toward members of competing out-groups.

In another trust game experiment, Kosfeld et al. (2005) administered—via nasal spray—either OXT or a placebo to subjects (DM 1 in Figure 3). As the results show, OXT considerably increases the trustors' trust into the trustees. Out of the 29 subjects, 13 (45%) in the OXT group showed the maximal trust level (i.e., they sent all the money), whereas only six of the 29 subjects (21%) in the placebo group showed maximal trust. Thus, the Kosfeld et al. experiment shows that exogenously administered OXT increases trust.

This result, however, allows for a number of causal interpretations (e.g., Fehr, Fischbacher, & Kosfeld, 2005). First, did OXT increase the trustors' trust by creating more optimistic beliefs about the trustworthiness of the trustees? Second, did OXT make trustors more generous? Third, did it make trustors more risk/ambiguity-seeking? Or fourth, did OXT influence betrayal aversion? Because the experiment of Kosfeld et al. (2005) controlled for these alternative explanations, it was possible to find out that OXT helps to overcome betrayal aversion (Bohnet, Greig, Herrmann, & Zeckhauser, 2008; Damasio, 2005).[2] Thus, it can be theorized that OXT affects the neurobiological mechanisms that underlie social preferences (Fehr, 2009a, 2009b). However, although the effect of OXT on generosity was ruled out in the Kosfeld et al. (2005) study, other investigations (Barraza & Zak, 2009; Zak, Stanton, & Ahmadi, 2007) identified a positive correlation between OXT and generosity. Considering this, current research suggests that OXT affects both betrayal aversion and generosity.

When OXT levels are increased, people become more trusting (e.g., Kosfeld et al., 2005). Mikolajczak et al. (2010), however, raised the following question: Does OXT increase people's trust in anybody, or can contextual cues of unreliability override the effects of OXT? The results of their experiment show that people given OXT rather than placebo trust others more to the extent that trustees were displayed in a neutral or positive fashion. However, when the trustees were displayed in a negative fashion, OXT did not influence trust. This result suggests that the effect of OXT on trust may be moderated by the perceived risk inherent to the interaction (Mikolajczak et al., 2010).

Another study (Domes, Heinrichs, Michel, Berger, & Herpertz, 2007) investigated the effect of intranasally administered OXT on mind-reading ability. Mind-reading is the ability to infer the internal states of other actors (e.g., intentions, thoughts, and feelings) to predict their behavior (also known as theory of mind, TOM), and the underlying inference process is commonly referred to as mentalizing (Frith & Frith, 2003; Premack & Woodruff, 1978; Singer, 2009). The study by Domes et al. shows

[2] Aimone and Houser (2011) show that groups including betrayal-averse agents may achieve higher levels of reciprocity and more profitable social exchange than groups lacking betrayal aversion. These findings provide evidence on the benefits of betrayal aversion in trust situations.

that OXT improves the ability to infer the mental state of others from social cues of the eye region.

In general, mentalizing is a fundamental cognitive process in trust situations because the decision to trust involves thinking about an interaction partner's intentions, thoughts, and feelings to infer his or her trustworthiness (Fehr, 2009b; Krueger et al., 2007). Thus, in addition to its effects on betrayal aversion and generosity, OXT also has an effect on mentalizing. This further substantiates the prominent role of OXT for the establishment of interpersonal trust.

Arginine vasopressin. Arginine vasopressin (AVP), another neuropeptide, has also been shown to influence human social behavior (Heinrichs & Domes, 2008; Heinrichs et al., 2009). AVP is associated with a variety of male-typical social behaviors including aggression, territoriality, and enhanced stress-responsiveness (Donaldson & Young, 2008; Heinrichs & Domes, 2008). Consequently, AVP is hypothesized to be associated with avoidance and distrust behavior rather than with approach and trust behavior.

It is of importance for the neurobiological foundations of trust that research has already investigated the effects of nasally administered AVP on human facial responses related to social communication. One study (Thompson, Gupta, Miller, Mills, & Orr, 2004), for example, was designed to determine if AVP administration would influence cognitive, autonomic, as well as somatic responses to social stimuli important for agonistic communication in humans (agonistic is used as a synonym for aggressive in this context). Specifically, the study tested the effects of intranasal AVP administration on attention toward emotionally expressive facial expressions, as well as on heart rate, skin conductance, and corrugator supercilii electromyograms (corrugator EMG) in response to these social stimuli. The results show that AVP did not affect attention toward emotionally neutral, happy, and angry facial expressions. Also, AVP did not influence autonomic arousal in response to the three facial expressions. However, AVP did selectively enhance the corrugator EMG responses evoked by emotionally neutral facial expressions, making them similar in magnitude to responses evoked by angry facial expressions in control subjects.

In their discussion of the results, Thompson et al. (2004) argue that "[b]ecause this muscle group is involved in agonistic communication, these results suggest that AVP may influence aggression in human males by biasing individuals to respond to emotionally ambiguous social stimuli as if they were threatening/aggressive" (p. 35). Considering this and similar research findings (e.g., it was found that AVP decreases perceptions of the friendliness of faces, Thompson, George, Walton, Orr, & Benson, 2006), AVP can be considered as the adversary of OXT. Thus, AVP is associated with distrust, while OXT is associated with trust. In line with this reasoning, research has also shown that AVP is related to increased vigilance and anxiety (Carter, 2007; Murgatroyd et al., 2004).

Cortisol. Glucocorticoids, and cortisol (COR) in particular, are secreted by the adrenal cortex. Hence, they are also referred to as corticosteroids. These corticosteroids have been shown to have a number of neural and behavioral effects which are important in trust situations, such as learning and memory. Research has found, for example, that stress-induced COR elevation impairs social memory (Takahashi et al., 2004), in particular person recognition memory based on facial information (Rimmele et al., 2009), which is crucial for trust decisions in face-to-face settings (Winston et al., 2002). Moreover, psychosocial constructs such as optimism, self-mastery, self-esteem, and extraversion are associated with low activation in fear-related brain areas, which in turn results in attenuated COR responses to stress (Taylor et al., 2008).

Other studies (Heinrichs et al., 2003; Takahashi et al., 2005) investigated the direct relationship between trust and COR elevation induced by a social stress test. The results revealed a significant negative correlation between the two factors. That is, subjects with higher degrees of trust disposition have reduced social stress-induced COR response. Because (a) the stress hormone COR is synthesized in response to activation in the amygdala (Tillfors et al., 2001), and (b) lesions of the central nucleus of the amygdala decrease stress-induced hormone release (Prewitt & Herman, 1994), amygdala activation in healthy people is associated with distrust (Adolphs, Tranel, & Damasio, 1998; Dimoka, 2010; Takahashi et al., 2005).

There is also a link between neuropeptides and corticosteroids. Intranasal administration of AVP, for example, has been shown to increase

salivary COR levels in a social stress test (de Winter et al., 2003; Ebstein et al., 2009; Patchev & Almeida, 1995), while intranasal administered OXT reduced the COR response to social stress (Ditzen et al., 2009; Heinrichs et al., 2003).

Dopamine. There is evidence for a relation between hypothalamo-pituitary-adrenal (HPA) axis responses and the dopaminergic system (Rostene et al., 1995). It has been reported that selective destruction of mesencephalic dopaminergic neurons in rats leads to decreased levels of corticosterone (Casolini et al., 1993). Also, it has been shown that COR regulates dopamine (DOP; Barrot et al., 2000).

DOP is the main neurotransmitter associated with reward processing (Ikemoto & Panksepp, 1999; Schultz, 2000, 2002, 2006; Schultz, Dayan, & Montague, 1997). Because the goal of trusting another individual is to realize a reward (e.g., through a beneficial cooperation), it is no surprise that OXT has been shown to enhance DOP levels (Shahrokh, Zhang, Diorio, Gratton, & Meaney, 2010; Zak, 2011). Altogether, the important role of reward related neurobiological mechanisms in trust situations is stressed in multiple studies (e.g., King-Casas et al., 2005; Krueger et al., 2007)—Zak et al. (2005b) write in this context: "[T]he findings here should be considered in light of the fMRI study of Rilling et al. (2002), who show significant activity in ventromedial regions rich in dopamine receptors during cooperative behaviors. In the prairie vole, the nucleus accumbens is dense in OT [oxytocin] receptors . . . and OT appears to be critical for linking social signals to ventromedial reward circuits" (p. 526). In line with this argumentation, a recent review (Baskerville & Douglas, 2010) indicates the existence of a positive interaction between DOP and OXT.

Testosterone. Sex hormones also play a key role in human social behavior. Perception of distrust is correlated with an increase of dihydrotestosterone, a metabolite of the male sex hormone testosterone (TES) (Zak, Borja, Matzner, & Kurzban, 2005a). Moreover, elevated TES also causes people to behave antisocially (Zak et al., 2009). Based on the knowledge that TES is associated with competition and dominance, and therefore considered as an inhibitor of sociality, one recent study (Bos, Terburg, & Van Honk, 2010) tested whether TES has antagonistic characteristics with OXT. The results of the study show that compared with the placebo, TES significantly decreased interpersonal trust. Hence, the results confirm the predicted antagonistic characteristics with OXT. Moreover, additional analyses of the data revealed that this effect is determined by those who give trust easily, so-called naïve humans. Bos et al. (2010) conclude that "testosterone adaptively increases social vigilance in these trusting individuals [in particular the naïve humans] to better prepare them for competition over status and valued resources" (p. 1). Further support for the theorizing of Bos et al. (2010) comes from Johnson and Breedlove (2010) who argue that TES "may reduce interpersonal trust by acting on vasopressinergic neurons in the amygdala to increase communication to brainstem systems that activate fearful responses" (p. 11149).

An animal study (Arsenijevic & Tribollet, 1998), however, demonstrated that TES treatment of aging rats can restore normal adult levels of OXT receptor binding in specific brain areas. Thus, this study suggests a positive correlation between TES and OXT, thereby challenging the results of other research findings.

In another study (Zethraeus et al., 2009) that investigated the effects of sex hormones (e.g., TES) on various forms of social behaviors (e.g., trust, altruism, fairness), subjects were allocated to four weeks of treatment with sex hormones or a placebo. At the end of the treatment period, the subjects participated in experiments that were designed to measure the various forms of social behaviors. In essence, no significant effect of TES on any of the social behaviors was found. This result and findings of similar studies (e.g., Eisenegger, Naef, Snozzi, Heinrichs, & Fehr, 2010; Sanchez-Pages and Turiegano, 2010) substantiate the notion that the relation between TES and OXT is still not well understood. However, because studies with human subjects, in contrast to animal studies, found a negative relation between TES and OXT, we lend more weight to this theoretical perspective. Nevertheless, future studies should address the relation between TES and OXT to gain further insights.

Estrogen. With respect to the female sex hormone estrogen (EST), it has been reported that EST facilitates OXT uptake by facilitating receptor binding and increasing the number of OXT receptors (Verbalis, 1999). In line with

this finding, one study (Zak & Fakhar, 2006) analyzing data on biological, social, and environmental factors associated with hormone levels for a sample of 41 countries found that particular environmental conditions in some nations may be conducive to higher trust levels. In particular, this study found that nations whose citizens consume more food containing phytoestrogens (i.e., a diverse group of naturally occurring nonsteroidal plant compounds which may have similar effects as endogenous EST) exhibit higher levels of trust disposition. Thus, EST and OXT are positively correlated (Zak & Fakhar, 2006).

Serotonin. Serotonin (SER), a neurotransmitter mainly produced in the raphe nuclei, has multiple behavioral functions, in particular those related to social status, sexual behavior, as well as aggression and mood regulation. There is evidence that SER deficiency is associated with violent and aggressive behavior, while high levels of AVP facilitate aggressive behavior (Ferris et al., 1997). Thus, SER antagonizes AVP activity in the brain (Ferris & Delville, 1994).

SER depletion has been found to result in a higher rejection rate of unfair offers in an economic game (Crockett, Clark, Tabibnia, Lieberman, & Robbins, 2008), thereby indicating the role of SER for the control of negative emotions which also play an important role in trust situations (Winston et al., 2002). In line with this result, one study (Wood, Rilling, Sanfey, Bhagwager, & Rogers, 2006) found that subjects deprived of L-tryptophan, and consequently lower SER levels in the brain, are less cooperative in a repeated economic game than control subjects. In line with this finding, a recent article by Zak (2011) indicates that OXT release causes synaptic SER to rise, which in turn leads to calmness and positive mood by binding to 5-HT1A (SER) receptors in the temporal and prefrontal cortices.

Depression is also known to be the result of a serotonin deficiency. Also, depressive patients have been shown to have higher serum COR levels (Cowen, 2002). This suggests a link between SER and COR. Stress is known to activate the serotonergic system (McKittrick, Blanchard, Hardy, & Blanchard, 2009). Hence, considering the prominent role of COR for stress perceptions, a negative correlation between SER and COR with a positive feedback loop seems to exist. In line with this finding, research found that touch is associated with increased levels of SER (Field, Grizzle, Scafidi, & Schanberg, 1996), which have been demonstrated to reduce stress reactivity in humans (Hanley & Van de Kar, 2003). Finally, one investigation (Skuse & Gallagher, 2011) indicates that SER desensitizes the AVP receptor, which could increase avoidance and distrust behavior.

Summary of hormonal level of analysis. Figure 4 summarizes the results of our hormone review. The following list reflects the most important findings:

- Estrogen (EST), oxytocin (OXT), dompamine (DOP), and serotonin (SER) are typically associated with approach behavior and trust, whereas arginine vasopressin (AVP), cortisol (COR), and testosterone (TES) are more strongly associated with avoidance behavior and distrust.
- OXT, in contrast to the other hormones and neurotransmitters, seems to play a more central role for human trust behavior, as this substance has connections to five other substances, whereas the other substances have fewer connections.
- OXT is mainly influenced by sex hormones (i.e., EST and TES).
- The hormones and neurotransmitters make possible inferences regarding the psychological processes associated with a trust decision. In particular, reward processing (DOP), fear/uncertainty/stress processing and memory (COR), as well as mentalizing and processing of social preferences such as betrayal aversion and generosity (OXT) are important psychological processes involved in interpersonal trust situations.

Brain Level of Analysis

Genes influence the anatomy of the brain and its processing mechanisms (Rushton & Ankney, 2007; Rushton & Jensen, 2008). Moreover, the release of hormones is controlled by the brain, and, in turn, hormones influence the activation of particular brain regions via corresponding receptors (Loup et al., 1991). Consequently, a model of the biological determinants of human trust behavior has to consider the brain with its multifarious structures. In the following, we

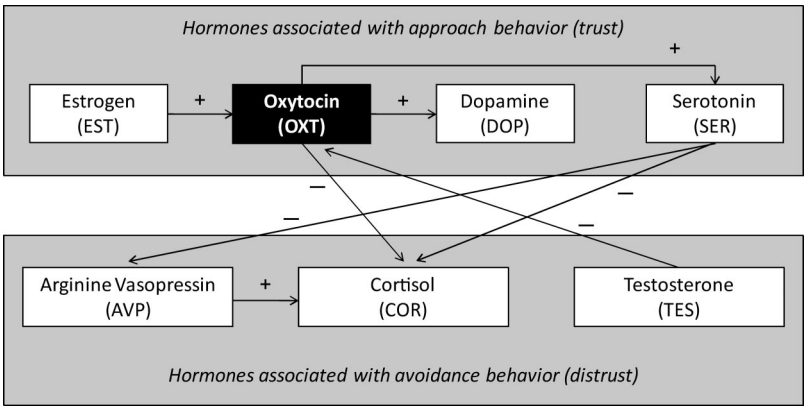

Figure 4. Relations among important trust-relevant hormones and neurotransmitters. *Notes:* The black box indicates that oxytocin (OXT) is the hub in the system of trust- and distrust-related hormones.

present empirical evidence showing that human trust behavior is associated with activation in specific brain areas.

A detailed look at the fMRI literature reveals that three different experimental paradigms have been used so far to study the neural correlates of trust. First, several studies have used the *trust game* as a paradigm. In this article, we discuss the following trust game studies: Baumgartner, Heinrichs, Vonlanthen, Fischbacher, and Fehr (2008); King-Casas et al. (2005); Delgado et al. (2005), and Krueger et al. (2007). Second, one study (Winston et al., 2002) used *human faces* with a varying degree of trustworthiness as stimuli. Third, two studies (Dimoka, 2010; Riedl, Hubert, & Kenning, 2010) used *eBay Web sites* as stimulus material in order to take into account the increasing trend toward computer-mediated interactions among humans. We structure the following section along these three experimental paradigms.

Trust game paradigm. One trust game experiment (Baumgartner et al., 2008) combined intranasal administration of OXT and fMRI. The results of the study revealed that subjects who had received OXT showed no change in their trust behavior after they learned that their trust had been breached several times, while subjects receiving a placebo decreased their trust. This result is in line with the findings of other hormone studies (Kosfeld et al., 2005; Zak et al., 2004, 2005b). Moreover, as the brain scans show, this difference in trust adaptation is associated with a specific reduction in activation in the amygdala, the midbrain regions, and the dorsal striatum in subjects receiving OXT, suggesting that neural systems mediating fear processing (amygdala and midbrain regions) and behavioral adaptations to feedback information (caudate nucleus, which is a part of the striatum) modulate the effect of OXT on trust (Baumgartner et al., 2008; Kirsch et al., 2005).

In addition to their main results, Baumgartner et al. (2008) list a number of additional brain regions activated in the trust game, namely putamen (a part of the striatum), thalamus, insular cortex, anterior cingulate cortex (ACC), inferior temporal gyrus, precuneus/posterior cingulum, superior parietal gyrus, postcentral gyrus, and the red nucleus. At least four regions on this list are hypothesized to play a crucial role in the neural implementation of mental processes that are important in trust situations.

First, activation in the striatum and thalamus is associated with reward processing and reward anticipation (Komura et al., 2001; O'Doherty et al., 2004; Schultz, 2006; Tricomi, Delgado, & Fiez, 2004). Second, activation in the insular cortex is related to perception of faces that express negative emotions (Phillips et al., 1997, 1998). Moreover, insular activation is associated with perception and processing of uncertainty, risk, and ambiguity (Critchley, Mathias, & Dolan, 2001; Huettel, Stowe, Gordon, Warner, & Platt, 2006; Kuhnen & Knutson, 2005; Krain, Wilson, Arbuckle, Castellanos, & Mil-

ham, 2006). Third, activation in the ACC is associated with cognitive conflict (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Botvinick, Cohen, & Carter, 2004). Considering these research findings, it can be hypothesized that striatum and thalamus activation plays a significant role in trust perceptions, whereas activation in the insular cortex is crucial for distrust perceptions. ACC activation, in contrast, is not trust or distrust specific. Rather, it is related to the comparison of the potential benefits and risks associated with a trust decision, thereby leading to cognitive conflict.

Using a multiround version of the trust game, King-Casas et al. (2005) found that reciprocity expressed by one player strongly predicts future trust expressed by the other player. Examining the trustees' brain scans, the study found that activity in the caudate nucleus was greatest when the trustor showed benevolent behavior, and most subdued when the trustor showed malevolent behavior. Moreover, activity in the caudate rose and fell with changes in the amount of money trustees returned to their trustors on the subsequent round, suggesting that this brain circuit receives or computes information about both the fairness of a social partner's decision and the intention to repay that decision with trust. Thus, activation in the caudate nucleus signals trust and/or learning of a partner's trustworthiness (King-Casas et al., 2005; Miller, 2005). In addition to caudate activity, the study revealed further brain regions activated in the trust game, namely the thalamus, ACC, inferior frontal sulcus, superior frontal sulcus, inferior/superior colliculi, and middle cingulate cortex.

In another fMRI study (Delgado et al., 2005), participants played a trust game with three hypothetical partners depicted as having moral character that was good, bad, or neutral. As the results show, participants chose to be more cooperative with the morally good partner. Moreover, caudate nucleus activation differentiated between positive and negative feedback, but only for the neutral partner. Notably, it did not do so for the good and bad partners, suggesting that prior social and moral perceptions can diminish reliance on feedback mechanisms in the caudate nucleus. Overall, the study provides evidence that moral and social perceptions can modulate neural mechanisms associated with feedback and reward processing and cognitive control in trust situations (Delgado et al., 2005; Singer, Kiebel, Winston, Dolan, & Frith, 2004).

In addition to the caudate, Delgado et al. (2005) found that the ventral portions of the striatum also play an important role in trust situations, because they are crucial for making predictions and anticipating the outcome of risky decisions. Finally, the study of Delgado et al. (2005) revealed high activation of the insular cortex and the cingulate cortex in trust situations—the former brain structure is important for the perception and processing of negative emotions (e.g., uncertainty), whereas the latter is important for the processing of cognitive conflict. This result is in line with the findings reported in the Baumgartner et al. (2008) study.

In a hyperfunctional MRI study (Krueger et al., 2007), two strangers interacted with one another in a sequential reciprocal trust game, while their brains were simultaneously scanned. The findings of the study suggest that the paracingulate cortex is critically involved in building a trust relationship by inferring another person's intentions to predict subsequent behavior (Bhatt & Camerer, 2005). In addition to the paracingulate cortex—which was identified as a mentalizing brain area not only by Krueger et al. (2007), but also by Gallagher, Jack, Roepstorff, and Frith (2002) and McCabe, Houser, Ryan, Smith, and Trouard (2001), the medial prefrontal cortex has also been found to be involved in thinking about other people's mental states (McCabe et al., 2001; Siegal & Varley, 2002; Stuss, Gallup, & Alexander, 2001). In particular, it is activated when making decisions and choices based on calculative expectations of what others will do (Bhatt & Camerer, 2005). Moreover, a near-infrared spectroscopy (NIRS) study (Yanagisawa et al., 2011) found that activity in the right ventrolateral prefrontal cortex (rVLPFC) is positively correlated with general trust. Because the level of general trust was negatively correlated with self-reported social pain during a social exclusion situation in this experiment, Yanagisawa et al. conclude that "rVLPFC activity mediated the relationship between general trust levels and social pain" (p. 190). In line with these findings, a recent review by

Rilling and Sanfey (2011) highlights the role of the prefrontal cortex in social decision-making.

Finally, a further brain region associated with mentalizing is the posterior superior temporal sulcus (Frith & Frith, 2003). This brain region is of particular importance in perceiving and processing biological motion of other species (Frith & Frith, 2010). Thus, in real-life situations (in contrast to static contexts like the trust game), this brain region is also likely to be activated in interpersonal trust situations.

The study of Krueger et al. (2007) revealed further brain regions associated with trust behavior: the ventral tegmental area, a region linked to the evaluation of expected and realized reward (Schultz et al., 1997), and the septal area, together with the adjoining hypothalamus, a limbic region that has been demonstrated to modulate various aspects of social behavior including social memory and learning (Numan, 2000). Because trust situations do not only exist on the basis of one-shot interpersonal interactions, but also on the basis of repeated interactions, memory and learning are crucial mental processes involved in human trust (King-Casas et al., 2005). Also, it has been shown that the septal area plays a crucial role in the release of several trust-related neuropeptides such as OXT. In addition, this area contains receptors for OXT and other neuropeptides (Loup et al., 1991; Powell & Rorie, 1967), and it is strongly connected with the hippocampus, which also has OXT receptors (Heinrichs, Meinlschmidt, Wippich, Ehlert, & Hellhammer, 2004), substantiating the notion that the septal area is an important brain region associated with human trust.

Human faces as a paradigm. An fMRI experiment conducted by Winston et al. (2002) found that the amygdala and insula were significantly more activated when subjects viewed faces that they rated as most untrustworthy, suggesting that these two brain regions associate perception of a face with an emotional response of distrust (Adolphs, 2002; Winston et al., 2002). Moreover, the study found activation in mentalizing brain areas (superior temporal sulcus). Based on these results, Winston et al. (2002) concluded that "social judgments about faces reflect a combination of brain responses that are stimulus driven, in the case of the amygdala, and driven by processes relating to inferences concerning the intentionality of others" (p. 281).

With respect to amygdala activation, three case studies (Adolphs et al., 1998, 2005) found that patients with complete bilateral amygdala damage judged other people to look more trustworthy and more approachable than did normal viewers or patients with brain damage in other areas, providing additional empirical evidence that the amygdala is associated with distrust. Other empirical studies found that this region is also activated by viewing faces that show certain emotional expressions, notably fear (Morris et al., 1996), and that it is activated when subjects view faces of people from another race (Hart et al., 2000; Phelps et al., 2000). In line with these findings, a recent lesion study (Koscik & Tranel, 2011) found that participants with unilateral damage to the amygdala displayed increased benevolent behavior in a multiround trust game, and specifically, they increased trust in response to betrayals. Thus, research indicates unambiguously that the amygdala is necessary for expressing normal interpersonal trust.

In a comment on the Winston et al. (2002) article, Adolphs (2002) explains that a detailed look at the stimulus material and the experimental designs reveals that "expressions of anger or sadness were negatively correlated with trustworthiness ratings, and happiness was positively correlated" (p. 193). Thus, the trustworthiness assigned to a particular face seems to depend on the emotion conveyed by the face (Todorov, 2008; Todorov, Baron, & Oosterhof, 2008).

However, such a conclusion cannot explain why different viewers of faces assign differential trustworthiness ratings. Adolphs (2002) also offers an explanation for this. In essence, he argues that trustworthiness ratings are a function of both the features of the stimuli (e.g., the emotion conveyed by a face) and the personalities and autobiographies of the viewers. Bearing this in mind, future research should not only manipulate the features of the stimuli. Rather, future studies should also recruit groups of subjects with varying personalities. For example, because autistic people have deficits in mentalizing (Baron-Cohen, Leslie, & Frith, 1985; Frith & Frith, 2010), which in turn has been shown to be associated with OXT levels (Domes et al., 2007) and activation in specific brain regions

such as the medial prefrontal cortex (Frith & Frith, 2006), a comparison of behavior and corresponding brain activation patterns and hormone levels between autistic people versus healthy controls in trust game experiments (or other experimental paradigms) could reveal new insights into the neurobiological mechanisms of the formation of trustworthiness judgments.

It is important to outline that it has not only been demonstrated in a face evaluation task (Winston et al., 2002) that automatic and unconscious information processing plays a crucial role in trust decisions. Based on their investigation in which they administered OXT to fMRI subjects, Baumgartner et al. (2008) for example, write: "[D]ifferences in brain activation between placebo and OT [oxytocin] subjects were only observed in subcortical structures as the amygdala, the midbrain, and the striatum. Those brain structures have each been associated with automatic and intuitive . . . or even unconscious processes" (p. 646). Prominent scholars such as Damasio (2005) and Zak (2003) support the importance of automatic and unconscious information processing in trust situations; the latter, for example, writes: "[T]rust is not a calculative activity . . . but a visceral sense that one has that a person can be trusted or not" (p. 21). Altogether, these findings indicate that a considerable proportion of neural processing in trust situations seems to be associated with automatic, emotional, and unconscious rather than deliberate, neutral, and conscious information processing.

Web sites as a Paradigm. So far, we have discussed the neurobiological foundations of trust based on the assumption that a trustor and a trustee interact with each other in a face-to-face setting (see Figure 1). During the past thousands of years, interaction among humans took place in such face-to-face environments. Today, however, increasingly more interactions among humans are computer-mediated. Information and communication technologies, in particular the Internet, have dramatically changed the context in which interactions among humans take place. In today's society, approximately two billion people use the Internet for communication and interaction, and this number is increasing on a daily basis (Internet World Stats, 2010).

As a result of this shift from face-to-face interaction to Internet-based communication and transactions, traditional human-human interactions are becoming increasingly more often computer-mediated. The use of computers and the Internet, however, usually increases a perception of complexity and uncertainty in both social and economic exchange (Pavlou, Liang, & Xue, 2007). Thus, trust as a construct is even more important in Internet environments than in traditional contexts in establishing interaction and cooperation among humans. A major reason for the increasing complexity and uncertainty in Internet environments is that during an interaction, a partner's face, which is typically not observable in computer-mediated interaction, serves the interpersonal function of allowing one person to predict another's personality traits and behavior (Knutson, 1996; Oosterhof & Todorov, 2008). This, in turn, has been demonstrated to apply in particular to trustworthiness predictions (Todorov, 2008; Winston et al., 2002). Hence, prediction of another person's personality traits and behavior is more difficult in computer-mediated interaction than in traditional face-to-face interaction.

Against this background, recent neuroimaging studies have begun to investigate the neural correlates of trust in computer-mediated interactions among humans. One fMRI study (Dimoka, 2010) used Web sites of eBay feedback profiles with varying levels of trustworthiness (manipulation based on the ratio between positive, neutral, and negative feedback comments) as stimuli to trigger brain activation (feedback profiles can be used by buyers to evaluate the quality of a transaction they have conducted with a seller; hence, sellers have an incentive to act in a trustworthy manner to build up a good reputation). In essence, the study found that trust is associated with brain areas linked to anticipating rewards (caudate nucleus), predicting the behavior of others (anterior paracingulate cortex), and calculating uncertainty (orbitofrontal cortex). Distrust, in contrast to trust, was found to be associated with brain areas linked to intense negative emotions (amygdala) and fear of loss (insular cortex).

Another brain imaging study (Riedl et al., 2010) used Internet offers of eBay sellers as stimulus material. Specifically, Toulmin's (1958) model of argumentation was applied to develop product description texts with varying

degrees of trustworthiness. The results of the study show that the processing of trustworthy eBay offers activates reward processing areas (striatum and thalamus) and mentalizing areas (prefrontal regions and cingulate cortex). In contrast, the processing of untrustworthy eBay offers activated regions associated with perception of uncertainty, in particular the insular cortex.

Summary of brain level of analysis. Taken together, the fMRI literature on human trust behavior has identified a number of crucial brain regions. We summarize these brain regions in Table 1. A marked cell in the table indicates that a study has found this brain region activated in a trust-related task. Table 1 shows that the neural correlates of trust are not so different at all across the three different experimental paradigms (trust game, human faces, eBay Web sites). Thus, our review provides evidence that the neural correlates of trust are relatively independent of the specific trust situation.

In Table 1, we also assign the brain regions identified in our review to five classes of mental processes that are crucial in trust situations:

- *Reward Processing* is associated with activation in the striatum and thalamus.
- *Uncertainty, risk, ambiguity, and fear processing* is associated with activation in the amygdala, insular cortex, as well as the hippocampus and parahippocampus gyrus.
- *Memory* is associated with activation in the amygdala, as well as hippocampus and parahippocampus gyrus.
- *Processing of cognitive conflict* is associated with activation in the cingulate cortex (in particular in the ACC).
- *Mentalizing and deliberate thinking* are two cognitive processes that are associated with activation in the frontal cortex.

A look at the five classes of mental processes reveals that the first class (reward) is of positive valence. Thus, activation in the corresponding brain regions is hypothesized to result in trust, approach, and cooperative behavior. In contrast, the second class (uncertainty, risk, ambiguity, fear) is of negative valence. Thus, activation in the corresponding brain regions is hypothesized to result in distrust, avoidance, and withdraw behavior. The other three classes (memory, cognitive conflict, as well as mentalizing and deliberate thinking) are neither of positive nor negative valence.

3. The Biology of Trust: Toward a Unified View

In section 2, we reviewed the biologically oriented literature on trust along three levels of analysis (genes, hormones, and the brain). The major goal of this discussion was to outline the relation between trust behavior and each of the three biological factors. However, as represented by the arrows in Figure 2, the three biological factors are also interrelated. In section 2, we only touched on these interrelationships. In this section, we make the relations more explicit to derive a unified view on the biology of trust. This unified view is illustrated in Figure 5 in the form of a conceptual model.

In the following, we discuss our conceptual model of the biology of trust (see Figure 5). Every human has a certain genetic predisposition that also concerns the oxytocin receptor (OXTR) gene and the arginine vasopressin 1a receptor (AVPR1a) gene. The characteristics of these genes affect two hormones, namely OXT (Cesarini et al., 2008; Reuter et al., 2009) and AVP (Knafo et al., 2008). While high levels of OXT lead to trust behavior, high levels of AVP typically result in distrust behavior (Heinrichs & Domes, 2008; Heinrichs et al., 2009). Hence, OXT antagonizes AVP (Donaldson & Young, 2008).

As illustrated in our model, OXT is influenced by sex hormones. First, the female sex hormone EST facilitates OXT uptake (Verbalis, 1999). Thus, EST positively affects OXT. Second, the male sex hormone TES has antagonistic characteristics with OXT (Bos et al., 2010). Thus, TES negatively affects OXT. (Note that the relation between TES and OXT is currently less well understood than the relation between EST and OXT, see section Hormonal Level of Analysis.)

OXT has been shown to enhance DOP levels, which play a crucial role for reward processing (Schultz, 2000, 2002, 2006; Schultz et al., 1997). This positive relation between the two hormones is accompanied by a corresponding relation on the cognitive and behavioral levels, because the ultimate goal of trusting another

Table 1
FMRI Studies On Human Trust Behavior and Associated Brain Regions

	Baumgartner et al. (2008)	Delgado et al. (2005)	King-Casas et al. (2005)	Krueger et al. (2007)	Winston et al. (2002)	Dimoka (2010)	Riedl et al. (2010)
Sample size (female/male)	49 (0/49)	12 (5/7)	96 (n.a.)	44 (22/22)	14 (6/8)	15 (6/9)	20 (10/10)
Experimental paradigm (stimulus)	Trust game	Trust game	Trust game	Trust game	Faces	eBay websites	eBay websites
Mental processes/brain regions							
Reward							
Striatum	●	●	●			●	●
Thalamus	●		●		●		●
Uncertainty, risk, ambiguity, fear/Memory							
Amygdala*	●				●	●	
Insular cortex	●	●			●	●	●
Hippocampus, parahippocampus gyrus*		●			●		●
Cognitive conflict							
Cingulate cortex	●	●	●	●		●	●
Mentalizing and deliberate thinking							
Frontal cortex				●		●	●

Note. The table lists brain regions that are mentioned in at least three of the seven studies. The brain regions are categorized along classes of mental processes that are important in trust situations: (a) reward processing, (b) uncertainty, risk, ambiguity, and fear processing, (c) processing of cognitive conflict, and (d) mentalizing and deliberate thinking. Note that we do not claim that a particular brain region is exclusively associated with one mental process. Rather, the relation between brain regions and mental processes is best described by a "many-to-many-concept." That is, activation in a specific brain region might be associated with several mental processes, and a specific mental process might be associated with activation in several brain regions (Price and Friston 2005). The asterisk (*) indicates that the amygdala, together with the hippocampus and parahippocampus gyrus, is also a crucial brain structure for emotion-based memory and learning, which are of particular importance for trust decisions independent of the experimental paradigm. Finally, note that although five out of the seven studies used a mixed-gender sample (with the exception of the study by Baumgartner et al., and with information unavailable in King-Casas et al.), only the Riedl et al. study has a focus on gender differences in brain activation in trust decisions.

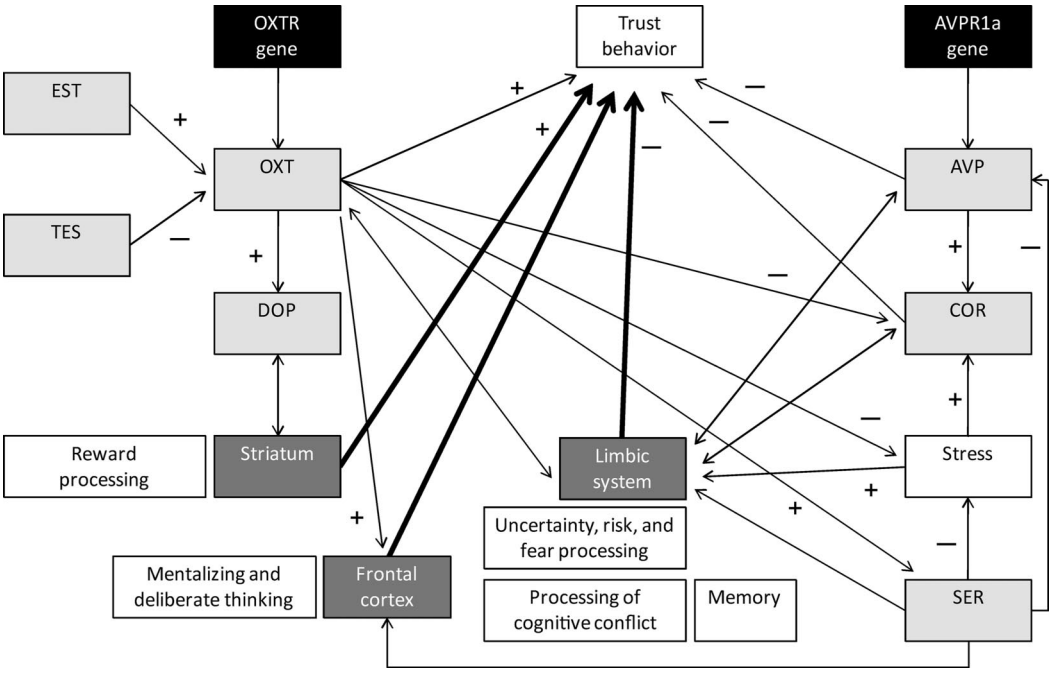

Figure 5. Conceptual Model of the Biology of Trust. *Notes:* AVP: Arginine Vasopressin, AVPR1a: Arginine Vasopressin 1a Receptor, COR: Cortisol, DOP: Dopamine, EST: Estrogen, OXT: Oxytocin, OXTR: Oxytocin Receptor, SER: Serotonin, TES: Testosterone. Black boxes indicate a genetic level of analysis. Light gray boxes indicate a hormonal level of analysis. Dark gray boxes indicate a brain level of analysis. White boxes indicate a behavioral level of analysis (trust behavior and stress), as well as cognitive constructs (reward processing; uncertainty, risk, and fear processing; processing of cognitive conflict; mentalizing and deliberate thinking). Plus (+) and minus (−) signs indicate the hypothesized valence of the relation (positive or negative). Direction of the arrows indicates the hypothesized causality. A two sided-arrow indicates interplay among two variables. The arrows in bold illustrate that the influences of genes and hormones on trust behavior are mediated by activation in specific brain areas. Thus, a direct link between a hormone and trust behavior indicates that research has investigated this relation without considering brain activity as a mediator (e.g., OXT→Trust behavior, Kosfeld et al., 2005). Anatomical connections between the brain areas (striatum, limbic system, and frontal cortex) are not illustrated to facilitate clarity of the model.

person (OXT) is the realization of a reward (DOP; Shahrokh et al., 2010). In line with this argumentation, a recent study (Theodoridou, Rowe, Penton-Voak, & Rogers, 2009) found that nasally administered OXT increases perceived facial trustworthiness and attractiveness of unfamiliar people, and this in turn may positively affect reward perceptions.

DOP is correlated with reward processing in the brain, and DOP is released by a specific part in the striatum, the nucleus accumbens (Schultz et al., 1997). Hence, an interplay between the striatum and DOP does exist. In line with our theorizing (i.e., OXT→DOP↔Striatum), a number of brain imaging experiments found activation in the striatum, and also in other reward areas like the thalamus, in trust-related tasks (see Table 1). Therefore, one major conclusion is that activation in the striatum typically leads to trust behavior (Zak, 2011).

As illustrated in Figure 5, OXT positively affects the ability to infer the internal states of other people—for example their intentions, thoughts, and feelings—to predict their behavior (Domes et al., 2007). This mentalizing process, together with deliberate thinking, is associated with activation in the frontal cortex, in particular in the medial frontal cortex (Amodio & Frith, 2006; Stuss et al., 2001). Importantly, mentalizing is a fundamental

cognitive process in trust situations because the decision to trust implies thinking about an interaction partner's intentions to infer his or her trustworthiness (Fehr, 2009b; Krueger et al., 2007). Thus, mentalizing and corresponding activity in the frontal cortex, respectively, influences trust behavior.

OXT is also related to the limbic system. OXT is synthesized in the paraventricular nucleus and supraoptic nucleus of the hypothalamus (Zak & Fakhar, 2006). Moreover, limbic regions associated with emotions and memory (e.g., amygdala, hippocampus, and hypothalamus) have an accumulation of OXT receptors (Barberis & Tribollet, 1996; Insel, 1997; Landgraf & Neumann, 2004; Verbalis, 1999). Also, it has been found that OXT acts on limbic regions, in particular the amygdala, by inhibiting excitatory information from these regions to brainstem sites mediating autonomic fear responses (Debiec, 2005; Huber, Pierre, & Ron, 2005; Petrovic, Kalisch, Singer, & Dolan, 2008). In this context, Baumgartner et al. (2008) recently wrote: "[OXT] reduces fear responses during the trust game by reducing activation in the amygdala and connected brainstem effector sites, which in turn enhances subjects' ability to trust in situations characterized by the risk of betrayal" (p. 645).

In addition to brain structures associated with the processing of negative emotions like fear (e.g., amygdala), as well as uncertainty and risk (e.g., insular cortex), a certain limbic region has been shown to process cognitive conflict, namely the ACC (Botvinick et al., 1999, 2004). Because emotion-related brain structures like the amygdala have abundant connections to the ACC, it can be hypothesized that automatic and typically unconscious information processing, in particular in the amygdala, may be followed by the processing of cognitive conflict in the ACC as a result of balancing the rapid visceral perceptions with more deliberate thoughts (Riedl et al., 2010).

Our review reveals a crucial role for the amygdala and hippocampus (plus parahippocampus gyrus) in trust decisions across all three experimental paradigms (see the asterisks in Table 1). Both limbic structures are important for emotion perception and processing. However, in addition to this general role in emotion perception and processing, both the amygdala and hippocampus were also found to be specifically associated with emotion-based memory (Gazzaniga, Ivry, & Mangun, 2009). Although the exact neural mechanisms are not yet fully understood, a basic understanding of the role of the amygdala and hippocampus in emotion-based memory does exist. We describe this mechanism in the following.

Because emotion is a "fuzzy" construct (Phelps & La Bar, 2006, p. 422), brain research has begun to anatomize it. One prominent categorization is based on the factors *valence* (pleasant-unpleasant or good-bad) and *arousal* (the intensity of the internal emotional response, high-low). By using this categorization, scholars can achieve "a more concrete assessment of the emotional reactions elicited by stimuli" (Gazzaniga et al., 2009, p. 367). With respect to memory, established categorizations refer to *duration* (sensory memory: milliseconds to seconds; short-term memory: seconds to minutes; long-term memory: hours to years) and *content* (declarative memory: knowledge to which we have conscious access such as the recall of events in our lives; nondeclarative memory: knowledge to which we do not have conscious access such as motor and cognitive skills; Gazzaniga et al., 2009).

Research has found that amygdala activity is typically associated with negatively valenced emotions (in particular fear) and with high levels of arousal (Phelps & LeDoux, 2005). Moreover, it has been found that the amygdala plays a crucial role in the neural implementation of declarative memory, independent of how a particular stimulus in the world has been linked to potentially aversive consequences (i.e., via instruction, observation, or experience). Specifically, it has also been found that information storage in the hippocampus is mediated by the amygdala (Gazzaniga et al., 2009), and arousal, rather than the valence of emotion, has been demonstrated to mediate memory (McGaugh, 2004). Thus, arousing events are not forgotten as quickly as nonarousing events (Kleinsmith & Kaplan, 1963).

Against this background, the results presented in Table 1 can be interpreted as follows: Independently of the specific trust situation (i.e., playing the trust game, as well as viewing human faces or Web sites with varying degrees of trustworthiness), the amygdala processes the arousal associated with a particular stimulus, thereby mediating the storage of information concerning the trustworthiness of a stimulus in the hippocampus. A high level of arousal at the storage of a stimulus, both with positive and

negative valence, positively affects the capacity for remembering it in a future trust situation.

Figure 5 also shows an interplay between the limbic system and COR. It is an established fact that social stress leads to increased amygdala-HPA activation, which in turn elevates COR levels (Takahashi et al., 2005). Moreover, COR affects the limbic system because it influences hippocampus activity. Specifically, if levels of arousal and stress become too high, resulting in pronounced COR elevation, it is possible that social memory becomes impaired because glucocorticoid receptor (GR, Type II receptor)-mediated neuronal pathways result in suppressed synaptic potentiation (Takahashi, 2005). Thus, the relation between COR and synaptic plasticity (which is important for memory) is hypothesized to have an inverted-U shape (de Kloet et al., 1999; Kirschbaum & Hellhammer, 1994).

With respect to the relation between OXT and COR (stress), it was found that intranasally administered OXT significantly reduces salivary COR levels (Ditzen et al., 2009). In line with this finding, another study (Takahashi et al., 2005) reports a negative correlation between interpersonal trust and COR. That is, subjects with higher degrees of interpersonal trust have lower levels of neuroendocrine response (i.e., COR) to social stress (measured with the Trier Social Stress Test). Additional support for these findings comes from research which reports that subjects with social phobia and anxiety disorders, who typically have an exaggerated HPA response to social stressors, exhibit high levels of interpersonal distrust (Furlan, DeMartinis, Schweizer, Rickels, & Lucki, 2001; Takahashi et al., 2005).

We have already outlined that high levels of AVP typically result in distrust behavior (Heinrichs & Domes, 2008; Heinrichs et al., 2009). A possible pathway for this effect could exist via the limbic system (Debiec, 2005). One study (Murgatroyd et al., 2004), for example, has shown that elevated AVP expression in the hypothalamic paraventricular nucleus is associated with increased anxiety levels, which in turn are related to amygdala activity. In line with this finding, it is reported that AVP (a) is correlated with increased vigilance, anxiety, arousal, and activation (Heinrichs & Domes, 2008), and (b) enhances the encoding of both happy and angry social information to make this more memorable (Guastella, Kenyon, Alvares, Carson, & Hickie, 2010). Because all these four states, as well as memory, are related to activity in limbic structures, this finding further supports the notion of an interplay between AVP and limbic brain regions. Finally, because both COR and AVP negatively affect trust behavior, it is no surprise that research has also found a direct positive relation between these two hormones. Specifically, intranasal administration of AVP has been shown to increase salivary COR levels in a social stress test (de Winter et al., 2003; Ebstein et al., 2009; Patchev & Almeida, 1995).

The final hormone which we address in our conceptual model (see Figure 5) is SER. It has been shown that SER depletion leads to a higher rejection rate of unfair offers in a two-person economic game (Crockett et al., 2008), indicating the importance of SER for the control of negative emotions that are processed and controlled both in the limbic system and in the frontal cortex (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003). Importantly, such negative emotions play a significant role in trust situations (Winston et al., 2002). In line with this result, research (Wood et al., 2006) has found that people with lower SER levels in the brain are less cooperative in economic game playing than control subjects. Moreover, a negative correlation between SER and stress is likely to exist because research has found that touch is associated with increased levels of SER (Field et al., 1996), which have been demonstrated to reduce stress reactivity in humans (Hanley & Van de Kar, 2003). Finally, it is reported that SER desensitizes the AVP receptor (Skuse & Gallagher, 2011), which could reduce approach and trust behavior.

4. Summary, Implications, and Conclusion

In the present article, we have reviewed the literature on the biology of trust. We structured our discussion along a model, which integrates three biological levels of analysis: genes, hormones, and the brain (see Figure 2). Altogether, our review indicates that at least a moderate degree of human trust behavior is genetically predetermined, and several hormones (in particular OXT, see Figure 4), as well as specific brain areas (which are mainly located in the basal ganglia, the limbic system, and the frontal cortex, see Table 1), are strongly associated with human trust behavior. Importantly, a major contribution of the present article is the devel-

opment of a conceptual framework which unifies the levels of analysis (see Figure 5).

Based on the results of our review, possible avenues for future trust research have emerged. These implications for future research can be grouped into two classes: methodological and thematic.

From a methodological perspective, one major conclusion that can be drawn from our review is that most studies have so far addressed only one level of analysis. Despite the value of such investigations, it is clear that studies integrating two or even all three biological levels of analysis could result in a deeper understanding of the complex relations among the biological determinants of trust behavior. One study (Baumgartner et al., 2008), for example, has already combined administration of OXT and fMRI scanning, thereby triangulating across two different levels of analysis (hormone and brain). As a consequence, this study revealed intriguing insights into the interplay between OXT and brain areas in the basal ganglia (striatum) and limbic system (amygdala). Drawing upon this pioneering study, we call for more investigations that combine the various levels of analysis.

On the brain level of analysis, we chose to focus on functional brain imaging studies, because these studies dominate the literature in terms of quantity. Other methods (e.g., electroencephalography, EEG), in contrast, have been used to a much smaller extent to reveal the neurobiological mechanisms underlying trust. Against this background, future trust research could shift its focus to EEG and other neuroscience (e.g., diffusion tensor imaging) and neurophysiological (e.g., galvanic skin response) tools.

One study (Boudreau, McCubbins, & Coulson, 2008), for example, already investigated human trust behavior by means of EEG. This study signifies the potential of the complementary use of different neuroscience tools to shed light on the neurobiological mechanisms underlying human trust behavior. EEG enables a temporal resolution of milliseconds and this tool can, therefore, easily detect the time course of neural activity. However, since its spatial resolution is limited, triangulation with other tools is essential—fMRI studies, in particular, have already revealed many insights into the neural mechanisms of trust behavior. However, a major implication of our review is that future investigations could use tools which are currently not well represented in studies on the biological foundations of trust (e.g., EEG). Accordingly, research results should be embraced only after they are corroborated by more than one method.

In addition to these methodological implications, our review also reveals thematic avenues for future studies, of which we address three important ones in the following.

We discussed that increasingly more interactions among humans are taking place online. Thus, traditional face-to-face interactions are declining, while computer-mediated interactions are increasing. Two recent fMRI studies (Dimoka, 2010; Riedl et al., 2010) have already addressed this development and investigated the neural correlates of trust while the participants viewed eBay Web sites. On the Internet, however, an increasing number of users have started to represent themselves as computer generated virtual characters, so-called *avatars* (Bainbridge, 2007).[3] Thus, the question arises whether the neurobiological mechanisms underlying human trust behavior in traditional face-to-face settings resemble those in avatar-based online environments.

In one fMRI study (Moser et al., 2007), participants performed facial emotion recognition tasks based on human and avatar stimuli. Although the neural responses were significantly stronger to human faces in face-sensitive structures (in particular in the fusiform gyrus), robust amygdala activation was found in response to both human and avatar emotional faces. Other studies using different techniques to capture neurobiological reactions to avatars and avatar-like robots (heart rate, electrodermal activity, electromyographic activity, and electroencephalography) found similar results (Dubal, Foucher, Jouvent, & Nadel, 2011; Garau, Slater, Pertaub, & Razzaque, 2005; Weyers, Mühlberger, Hefele, & Pauli, 2006). Considering this, current research suggests that avatars have the potential to elicit strong emotional reactions in humans, one of the major foundations for trust, as well as distrust, to emerge (Damasio, 2005; Winston et al., 2002; Zak, 2003). Building on these existing studies, we call for bio-

[3] For example, Second Life, one of the worldwide largest virtual worlds in which users interact with each other through avatars, has currently approximately 24 million user accounts (see www.secondlife.com, September 2011).

logical research with the purpose of revealing the differences and similarities in the biology of trust in traditional face-to-face settings and computer-mediated (avatar-based) environments. Such research efforts will not only reveal theoretical insights into the neurobiology of trust, but also indicate findings which might be relevant for policymakers.

Another finding of our review is that research has focused on the trustor rather than the trustee (see Figure 1). Thus, existing knowledge on the biology of trust and corresponding behavior mainly pertains to one specific party in a trust situation—exceptions are, for example, Cesarini et al. (2008), Zak et al. (2004, 2005b), and King-Casas et al. (2005). Considering this research gap, we call for more investigations that focus on the trustee, as well as the direct interaction between the trustor and trustee (e.g., via hyperscanning).

A final conclusion of our review is that to date, moderator variables such as gender and age have hardly been addressed in research on the biology of trust. With respect to gender, one fMRI study (Riedl et al., 2010) found brain activation differences between men and women when processing trustworthy and untrustworthy eBay offers. These results indicate that well-known gender differences in computer and Internet behavior are correlated with differences in brain functionality. Moreover, behavioral research (e.g., Sutter & Kocher, 2007) found that trust increases almost linearly from early childhood to adulthood. Because the human brain develops across the entire life span, it would be worthwhile to see what insight future research might reveal as to how these behavioral changes are associated with neurobiological changes.

In our concluding statement, we would like to direct the readers' attention to the fact that this review is focused on the *biology* of trust. We have already pointed out that trust behavior is not only influenced by such biological factors, rather, environmental factors, in particular those related to the *social dimension of interaction among humans* (e.g., socialization and culture), also affect human trust behavior (see Figure 2). In this context, Cacioppo et al. (2000, p. 829) write: "All human behavior, at some level, is biological, but this is not to say that biological reductionism yields a simple, singular, or satisfactory explanation for complex behaviors or that molecular forms of representation provide the only or best level of analysis for understanding human behavior . . . constructs such as those developed by the social sciences provide a means of understanding highly complex activity without needing to specify each individual action of the simplest components, thereby offering an efficient means of describing the behavior of a complex system." Against this background, we argue that in order to obtain a comprehensive understanding of human trust behavior, it would be advantageous to integrate the biological and the social perspectives in future research.

References

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433,* 68–72. doi:10.1038/nature03086

Adolphs, R., Tranel, D., & Damasio, A. (1998). The human amygdala in social judgement. *Nature, 393,* 470–474. doi:10.1038/30982

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion Neurobiology, 12,* 169–177. doi:10.1016/S0959-4388(02)00301-X

Aimone, J. A., & Houser, D. (2011). Beneficial betrayal aversion. *PLoS ONE, 6,* e17725. doi:10.1371/journal.pone.0017725

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7,* 268–277. doi:10.1038/nrn1884

Apicella, C. L., Cesarini, D., Johannesson, M., Dawes, C. T., Lichtenstein, P., Wallace, B., . . . Westberg, L. (2010). No association between oxytocin receptor (OXTR) gene polymorphisms and experimentally elicited social preferences. *PLoS ONE, 5,* e11153. doi:10.1371/journal.pone.0011153

Arsenijevic, Y., & Tribollet, E. (1998). Region-specific effect of testosterone on oxytocin receptor binding in the brain of the aged rat. *Brain Research, 785,* 167–170. doi:10.1016/S0006-8993(97)01429-7

Bainbridge, W. S. (2007). The scientific research potential of virtual worlds. *Science, 317,* 472–476. doi:10.1126/science.1146930

Barberis, C., & Tribollet, E. (1996). Vasopressin and oxytocin receptors in the central nervous system. *Critical Reviews in Neurobiology, 10,* 119–154.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21,* 37–46. doi:10.1016/0010-0277(85)90022-8

Barraza, J. A., & Zak, P. J. (2009). Empathy toward strangers triggers oxytocin release and subsequent generosity. *Annals of the New York Academy of Sci-*

ences, 1167, 182–189. doi:10.1111/j.1749-6632.2009.04504.x

Barrot, M., Marinelli, M., Abrous, D. N., Rouge-Pont, F., Le Moal, M., & Piazza, P. V. (2000). The dopaminergic hyper-responsiveness of the shell of the nucleus accumbens is hormone-dependent. *European Journal of Neuroscience, 12,* 973–979. doi:10.1046/j.1460-9568.2000.00996.x

Baskerville, T. A., & Douglas, A. J. (2010). Dopamine and oxytocin interactions underlying behaviors: Potential contributions to behavioral disorders. *CNS Neuroscience & Therapeutics, 16,* e92–e123. doi:10.1111/j.1755-5949.2010.00154.x

Baumgartner, T., Heinrichs, M., Vonlanthen, A., Fischbacher, U., & Fehr, E. (2008). Oxytocin shapes the neural circuitry of trust and trust adaptation in humans. *Neuron, 58,* 639–650. doi:10.1016/j.neuron.2008.04.009

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social-history. *Games and Economic Behavior, 10,* 122–142. doi:10.1006/game.1995.1027

Bhatt, M., & Camerer, C. F. (2005). Self-referential thinking and equilibrium as states of mind in games: FMRI evidence. *Games and Economic Behavior, 52,* 424–459. doi:10.1016/j.geb.2005.03.007

Bjornskov, C. (2007). Determinants of generalized trust: A cross-county comparison. *Public Choice, 130,* 1–21. doi:10.1007/s11127-006-9069-1

Bohnet, I., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review, 98,* 294–310. doi:10.1257/aer.98.1.294

Bos, P. A., Terburg, D., & Van Honk, J. (2010). Testosterone decreases trust in socially naïve humans. *Proceedings of the National Academy of Sciences, 107,* 9991–9995. doi:10.1073/pnas.0911700107

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8,* 539–546. doi:10.1016/j.tics.2004.10.003

Botvinick, M. M., Nystrom, L., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402,* 179–181. doi:10.1038/46035

Bouchard, T. J., Jr. (1994). Genes, environment, and personality. *Science, 264,* 1700–1701. doi:10.1126/science.8209250

Boudreau, C., McCubbins, M. D., & Coulson, S. (2008). Knowing when to trust others: An ERP study of decision making after receiving information from unknown people. *Social Cognitive and Affective Neuroscience, 4,* 23–34. doi:10.1093/scan/nsn034

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., & McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: Social neuroscience and the complementing nature of social and biological approaches. *Psychological Bulletin, 126,* 829–843. doi:10.1037/0033-2909.126.6.829

Camerer, C. F., & Weigelt, K. (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica, 56,* 1–36. doi:10.2307/1911840

Carter, C. S. (2007). Sex differences in oxytocin and vasopressin: Implications for autism spectrum disorders? *Behavioural Brain Research, 176,* 170–186. doi:10.1016/j.bbr.2006.08.025

Casolini, P., Kabbaj, M., Leprat, E., Piazza, P. V., Rouge-Pont, E., Angelucci, L., . . . Maccari, S. (1993). Basal and stress-induced corticosterone secretion is decreased by lesion of mesenchephalic dopaminergic neurons. *Brain Research, 622,* 311–314. doi:10.1016/0006-8993(93)90836-C

Cesarini, D., Dawes, C. T., Fowler, J. H., Johannesson, M., Lichtenstein, P., & Wallace, B. (2008). Heritability of cooperative behavior in the trust game. *Proceedings of the National Academy of Sciences, 105,* 3721–3726. doi:10.1073/pnas.0710069105

Coleman, J. S. (1990). *Foundations of social theory.* Cambridge, MA: Harvard University Press.

Cowen, P. J. (2002). Cortisol, serotonin and depression: All stressed out? *The British Journal of Psychiatry, 180,* 99–100. doi:10.1192/bjp.180.2.99

Critchley, H. D., Mathias, C. J., & Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron, 29,* 537–545. doi:10.1016/S0896-6273(01)00225-2

Crockett, M. J., Clark, L., Tabibnia, G., Lieberman, M. D., & Robbins, T. W. (2008). Serotonin modulates behavioral reactions to unfairness. *Science, 320,* 1739. doi:10.1126/science.1155577

Damasio, A. (2005). Brain trust. *Nature, 435,* 571–572. doi:10.1038/435571a

Debiec, J. (2005). Peptides of love and fear: Vasopressin and oxytocin modulate the integration of information in the amygdala. *BioEssays, 27,* 869–873. doi:10.1002/bies.20301

De Dreu, C. K. W., Greer, L. L., Handgraaf, M. J. J., Shalvi, S., Van Kleef, G. A., Baas, M., . . . Feith, S. W. W. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science, 328,* 1408–1411. doi:10.1126/science.1189047

de Kloet, E. R., Oitzl, M. S., & Joels, M. (1999). Stress and cognition: Are corticosteroids good or bad guys? *Trends in Neuroscience, 22,* 422–426. doi:10.1016/S0166-2236(99)01438-1

Delgado, M. R., Frank, R. H., & Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience, 8,* 1611–1618. doi:10.1038/nn1575

de Winter, R. F., Van Hemert, A. M., DeRijk, R. H., Zwinderman, K. H., Frankhuijzen-Sierevogel, A. C., Wiegant, V. M., & Goekopp, J. G. (2003). Anxious-retarded depression: Relation with plasma vasopres-

sin and cortisol. *Neuropsychopharmacology, 28,* 140–147. doi:10.1038/sj.npp.1300002

Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Quarterly, 34,* 373–396.

Ditzen, B., Schaer, M., Gabriel, B., Bodenmann, G., Ehlert, U., & Heinrichs, M. (2009). Intranasal oxytocin increases positive communication and reduces cortisol levels during couple conflict. *Biological Psychiatry, 65,* 728–731. doi:10.1016/j.biopsych.2008.10.011

Dohmen, T., Falk, A., Huffman, D., & Sunde, U. (2006). The intergenerational transmission of risk and trust attitudes. IZA Discussion Paper, No. 2380, The Institute for the Study of Labor, Bonn, Germany.

Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. (2007). Oxytocin improves "mind-reading" in humans. *Biological Psychiatry, 61,* 731–733. doi:10.1016/j.biopsych.2006.07.015

Donaldson, Z. R., & Young, L. J. (2008). Oxytocin vasopressin and the neurogenetics of sociality. *Science, 322,* 900–904. doi:10.1126/science.1158668

Dubal, S., Foucher, A., Jouvent, R., & Nadel, J. (2011). Human brain spots emotion in non humanoid robots. *Social Cognitive and Affective Neuroscience, 6,* 90–97. doi:10.1093/scan/nsq019

Ebstein, R. P., Israel, S., Lerer, E., Uzefovsky, F., Shalev, I., Gritsenko, I., . . . Yirmiya, N. (2009). Arginine vasopressin and oxytocin modulate human social behavior. *Annals of the New York Academy of Sciences, 1167,* 87–102. doi:10.1111/j.1749-6632.2009.04541.x

Eisenegger, C., Naef, M., Snozzi, R., Heinrichs, M., & Fehr, E. (2010). Prejudice and truth about the effect of testosterone on human bargaining behaviour. *Nature, 463,* 356–359. doi:10.1038/nature08711

Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review, 96,* 1611–1630. doi:10.1257/aer.96.5.1611

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: The neural circuitry of social preferences. *Trends in Cognitive Sciences, 11,* 419–427. doi:10.1016/j.tics.2007.09.002

Fehr, E. (2009a). On the economics and biology of trust. *Journal of the European Economic Association, 7,* 235–266. doi:10.1162/JEEA.2009.7.2-3.235

Fehr, E. (2009b). Social preferences and the brain. In P. W. Glimcher, C. F. Camerer, E. Fehr, and R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain,* (pp. 215–232). Amsterdam, The Netherlands: Elsevier.

Fehr, E. Fischbacher, U., & Kosfeld, M. (2005). Neuroeconomic foundations of trust and social preferences: Initial evidence. *American Economic Review, 95,* 346–351. doi:10.1257/000282805774669736

Ferris, C. F., & Delville, Y. (1994). Vasopressin and serotonin interactions in the control of agonistic behavior. *Psychoneuroendocrinology, 19,* 593–601. doi:10.1016/0306-4530(94)90043-4

Ferris, C. F., Melloni, Jr., R. H., Koppel, G., Perry, K. W., Fuller, R. W., & Delville, Y. (1997). Vasopressin/serotonin interactions in the anterior hypothalamus control aggressive behavior in golden hamsters. *Journal of Neuroscience, 17,* 4331–4340.

Field, T., Grizzle, N., Scafidi, F., & Schanberg, S. (1996). Massage and relaxation therapies' effects on depressed adolescent mothers. *Adolescence, 31,* 903–911.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research.* Reading, MA: Addison-Wesley.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50,* 531–534. doi:10.1016/j.neuron.2006.05.001

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London–B, 358,* 459–473. doi:10.1098/rstb.2002.1218

Frith, U., & Frith, C. D. (2010). The social brain: Allowing humans to boldly go where no other species has been. *Philosophical Transactions of the Royal Society of London–B, 365,* 165–176. doi:10.1098/rstb.2009.0160

Füllbrunn, S., Richwien, K., & Sadrieh, A. (2011). Trust and trustworthiness in anonymous virtual worlds. *Journal of Media Economics, 24,* 48–63. doi:10.1080/08997764.2011.549432

Furlan, P. M., DeMartinis, N., Schweizer, E., Rickels, K., & Lucki, I. (2001). Abnormal salivary cortisol levels in social phobic patients in response to acute psychological but not physical stress. *Biological Psychiatry, 50,* 254–259. doi:10.1016/S0006-3223(00)01126-4

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game. *Neuroimage, 16,* 814–821. doi:10.1006/nimg.2002.1117

Garau, M., Slater, M., Pertaub, D.-P., & Razzaque, S. (2005). The responses of people to virtual humans in an immersive virtual environment. *Presence, 14,* 104–116. doi:10.1162/1054746053890242

Gazzaniga, M. S., Ivry, R., & Mangun, G. R. (2009). *Cognitive neuroscience: The biology of the mind* (3rd edition), New York, NY: W. W. Norton.

Gefen, D., Benbasat, I., & Pavlou, P. A. (2008). A research agenda for trust in online environments. *Journal of Management Information Systems, 24,* 275–286. doi:10.2753/MIS0742–1222240411

Guastella, A. J., Kenyon, A. R., Alvares, G. A., Carson, D. S., & Hickie, I. B. (2010). Intranasal arginine vasopressin enhances the encoding of happy and angry faces in humans. *Biological Psychiatry, 67,* 1220–1222. doi:10.1016/j.biopsych.2010.03.014

Hanley, N. R., & Van de Kar, L. D. (2003). Serotonin and the neuroendocrine regulation of the hypotha-

lamic-pituitary-adrenal axis in health and disease. *Vitamins and Hormones, 66,* 189–255. doi: 10.1016/S0083-6729(03)01006-9

Harlan, R. E. (1988). Regulation of neuropeptide gene expression by steroid hormones. *Molecular Neurobiology, 2,* 183–200. doi:10.1007/BF02935345

Harris, G. W. (1948). Neural control of the pituitary gland. *Physiological Reviews, 28,* 139–179.

Harris, J. A., Vernon, P. A., & Boomsma, D. I. (1998). The heritability of testosterone: A study of Dutch adolescent twins and their parents. *Behavior Genetics, 28,* 165–171. doi:10.1023/A:1021466929053

Hart, A. J., Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., & Rauch, S. L. (2000). Differential response in the human amygdala to racial outgroup vs ingroup face stimuli. *Neuroreport, 11,* 2351–2355. doi:10.1097/00001756-200008030-00004

Heinrichs, M., Baumgartner, T., Kirschbaum, C., & Ehlert, U. (2003). Social support and oxytocin interact to suppress cortisol and subjective responses to psychosocial stress. *Biological Psychiatry, 54,* 1389–1398. doi:10.1016/S0006-3223(03)00465-7

Heinrichs, M., & Domes, G. (2008). Neuropeptides and social behavior: Effects of oxytocin and vasopressin in humans. *Progress in Brain Research, 170,* 337–350. doi:10.1016/S0079-6123(08)00428-7

Heinrichs, M., Meinlschmidt, G., Wippich, W., Ehlert, U., & Hellhammer, D. H. (2004). Selective amnesic effects of oxytocin on human memory. *Physiology & Behavior, 83,* 31–38.

Heinrichs, M., von Dawans, B., & Domes, G. (2009). Oxytocin, vasopressin, and human social behavior. *Frontiers in Neuroendocrinology, 30,* 548–557. doi:10.1016/j.yfrne.2009.05.005

Huber, D., Pierre, V., & Ron, S. (2005). Vasopressin and oxytocin excite distinct neuronal populations in the central amygdala. *Science, 308,* 245–248. doi:10.1126/science.1105636

Huettel, S. A., Stowe, C. J., Gordon, E. M., Warner, B. T., & Platt, M. L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron, 49,* 765–775. doi:10.1016/j.neuron.2006.01.024

Ikemoto, S., & Panksepp, J. (1999). The role of nucleus accumbens dopamine in motivated behavior: A unifying interpretation with special reference to reward-seeking. *Brain Research Reviews, 31,* 6–41. doi: 10.1016/S0165-0173(99)00023-5

Inoue, T., Kimura, T., Azuma, C., Inazawa, J., Takemura, M., Kikuchi, T., . . . Saji, F. (1994). Structural organization of the human oxytocin receptor gene. *Journal of Biological Chemistry, 269,* 32451–32456.

Insel, T. R. (1997). The neurobiology of social attachment. *American Journal of Psychiatry, 154,* 726–735.

Insel, T. R., & Young, L. J. (2001). The neurobiology of attachment. *Nature Reviews Neuroscience, 2,* 129–136. doi:10.1038/35053579

Internet World Stats (2010). Retrieved February 21, 2011 from http://www.internetworldstats.com

Israel, S., Lerer, E., Shalev, I., Uzefovsky, F., Riebold, M., Laiba, E., . . . Ebstein, R. P. (2009). The oxytocin receptor (OXTR) contributes to prosocial fund allocations in the dictator game and the social value orientations task. *PLoS ONE, 4,* e5535. doi:10.1371/journal.pone.0005535

Johnson, R. T., & Breedlove, S. M. (2010). Human trust: Testosterone raises suspicion. *Proceedings of the National Academy of Sciences, 107,* 11149–11150. doi:10.1073/pnas.1005376107

Johnson, W. (2007). Genetic and environmental influences on behavior: Capturing all the interplay. *Psychological Review, 114,* 423–440. doi: 10.1037/0033-295X.114.2.423

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science, 308,* 78–83. doi:10.1126/science.1108062

Kirsch, P., Esslinger, C., Chen, Q., Mier, D., Lis, S., Siddhanti, S., . . . Meyer-Lindenberg, A. (2005). Oxytocin modulates neural circuitry for social cognition and fear in humans. *Journal of Neuroscience, 25,* 11489–11493. doi:10.1523/JNEUROSCI.3984-05.2005

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology, 19,* 313–333. doi:10.1016/0306-4530(94)90013-2

Kleinsmith, L. J., & Kaplan, S. (1963). Paired-associate learning as a function of arousal and interpolated interval. *Journal of Experimental Psychology, 65,* 190–193. doi:10.1037/h0040288

Knafo, A., Israel, S., Darvasi, A., Bachner-Melman, R., Uzefovsky, F., Cohen, L., . . . Ebstein, R. P. (2008). Individual differences in allocation of funds in the dictator game associated with length of the arginine vasopressin 1a receptor RS3 promoter region and correlation between RS3 length and hippocampal mRNA. *Genes Brain & Behavior, 7,* 266–275. doi: 10.1111/j.1601-183X.2007.00341.x

Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior, 20,* 165–182. doi:10.1007/BF02281954

Komura, Y., Tamura, R., Uwano, T., Nishijo, H., Kaga, K., & Ono, T. (2001). Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature, 412,* 546–549. doi: 10.1038/35087595

Koscik, T. R., & Tranel, D. (2011). The human amygdala is necessary for developing and express-

ing normal interpersonal trust. *Neuropsychologia, 49,* 602–611. doi:10.1016/j.neuropsychologia.2010.09.023

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature, 435,* 673–676. doi:10.1038/nature03701

Krain, A. L., Wilson, A. M., Arbuckle, R., Castellanos, F. X., & Milham, M. P. (2006). Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. *Neuroimage, 32,* 477–484. doi:10.1016/j.neuroimage.2006.02.047

Krueger, F., Grafman, J., & McCabe, K. (2008). Neural correlates of economic game playing. *Philosophical Transactions of the Royal Society of London–B, 363,* 3859–3874. doi:10.1098/rstb.2008.0165

Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Zahn, R., Strenziok, M., . . . Grafman, J. (2007). Neural correlates of trust. *Proceedings of the National Academy of Sciences, 104,* 20084–20089. doi:10.1073/pnas.0710103104

Kugler, T., Connolly, T., & Kausel, E. E. (2009). The effect of consequential thinking on trust game behavior. *Journal of Behavioral Decision Making, 22,* 101–119. doi:10.1002/bdm.614

Kuhnen, C. M., & Knutson, B. (2005). The neural basis of financial risk-taking. *Neuron, 47,* 763–770. doi:10.1016/j.neuron.2005.08.008

Landgraf, R., & Neumann, I. D. (2004). Vasopressin and oxytocin release within the brain: A dynamic concept of multiple and variable modes of neuropeptide communication. *Frontiers in Neuroendocrinology, 25,* 150–176. doi:10.1016/j.yfrne.2004.05.001

Lee, Y. Y., & Lin, J. L. (2009). The effects of trust in physician on self-efficacy, adherence and diabetes outcomes. *Social Science and Medicine, 68,* 1060–1068. doi:10.1016/j.socscimed.2008.12.033

Lieberman, M. D., & Eisenberger, N. I. (2009). Pains and pleasures of social life. *Science, 323,* 890–891. doi:10.1126/science.1170008

Loup, F., Tribollet, E., Dubois-Dauphin, M., & Dreifuss, J. J. (1991). Localization of high affinity binding sites for oxytocin and vasopressin in the human brain: An autoradiographic study, *Brain Research, 555,* 220–232. doi:10.1016/0006-8993(91)90345-V

Luhmann, N. (1979). *Trust and power*. New York, NY: John Wiley & Sons.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review, 20,* 709–734.

McCabe, K., Houser, D., Ryan, L., Smith, V., & Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. *Proceedings of the National Academy of Sciences, 98,* 11832–11835. doi:10.1073/pnas.211415698

McCabe, K., Rigdon, M., & Smith, V. (2003). Positive reciprocity and intentions in trust games. *Journal of Economic Behavior & Organization, 52,* 267–275. doi:10.1016/S0167-2681(03)00003-9

McCabe, K. A., & Smith, V. (2000). A two person trust game played by naive and sophisticated subjects. *Proceedings of the National Academy of Sciences, 97,* 3777–3781. doi:10.1073/pnas.040577397

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience, 27,* 1–28. doi:10.1146/annurev.neuro.27.070203.144157

McKittrick, C. R., Blanchard, D. C., Hardy, M. P., & Blanchard, R. J. (2009). Social stress effects on hormones, brain, and behaviour. In D. W. Pfaff, A. P. Arnold, A. M. Etgen, S. E. Fahrbach, and R. T. Rubin (Eds.), *Hormones, brain and behavior*. (pp. 341−347). San Diego, CA: Academic. doi:10.1016/B978-008088783-8.00009-7

McKnight, H. D., & Chervany, N. L. (2001). What trust means in e-commerce customer relationships: An interdisciplinary conceptual typology. *International Journal of Electronic Commerce, 6,* 35–59.

McKnight, H. D., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research, 13,* 334–359. doi:10.1287/isre.13.3.334.81

Mikolajczak, M., Gross, J. J., Lane, A., Corneille, O., de Timary, P., & Luminet, O. (2010). Oxytocin makes people trusting, not gullible. *Psychological Science, 21,* 1072–1074. doi:10.1177/0956797610377343

Miller, G. (2005). Economic game shows how the brain builds trust. *Science, 308,* 36. doi:10.1126/science.308.5718.36a

Morris, J., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature, 383,* 812–815. doi:10.1038/383812a0

Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R. C., & Grammer, K. (2007). Amygdala activation at 3T in response to human and avatar facial expressions of emotions. *Journal of Neuroscience Methods, 161,* 126–133. doi:10.1016/j.jneumeth.2006.10.016

Murgatroyd, C., Wigger, A., Frank, E., Singewald, N., Bunck, M., Holsboer, F., . . . Spengle, D. (2004). Impaired repression at a vasopressin promoter polymorphism underlies overexpression of vasopressin in a rat model of trait anxiety. *Journal of Neuroscience, 24,* 7762–7770. doi:10.1523/JNEUROSCI.1614-04.2004

Nguyen, G. C., Laveist, T. A., Harris, M. L., Datta, L. W., Bayless, T. M., & Brant, S. R. (2009). Patient trust-in-physician and race are predictors of

adherence to medical management in inflammatory bowel disease. *Inflammatory Bowel Diseases, 15,* 1233–1239. doi:10.1002/ibd.20883

Numan, R. (2000). *The behavioral neuroscience of the septal region.* New York, NY: Springer. doi:10.1007/978-1-4612-1302-4

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304,* 452–454. doi:10.1126/science.1094285

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105,* 11087–11092. doi:10.1073/pnas.0805664105

Patchev, V. K., & Almeida, O. F. X. (1995). Corticosteroid regulation of gene expression and binding characteristics of vasopressin receptors in the rat brain. *European Journal of Neuroscience, 7,* 1579–1583. doi:10.1111/j.1460-9568.1995.tb01153.x

Pavlou, P. A., Liang, H., & Xue, Y. (2007). Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal Agent Perspective. *MIS Quarterly, 31,* 1–32.

Petrovic, P., Kalisch, R., Singer, T., & Dolan, R. J. (2008). Oxytocin attenuates affective evaluations of conditioned faces and amygdala activity. *Journal of Neuroscience, 28,* 6607–6615. doi:10.1523/JNEUROSCI.4572-07.2008

Phelps, E. A., & LaBar, K. S. (2006). Functional neuroimaging of emotion and social cognition. In R. Cabeza & A. Kingstone (Eds.), *Handbook of functional neuroimaging of cognition* (pp. 421–453). Cambridge, MA: MIT Press.

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behaviour. *Neuron, 48,* 175–187. doi:10.1016/j.neuron.2005.09.025

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *Journal of Cognitive Neuroscience, 12,* 729–738. doi:10.1162/089892900562552

Phillips, M., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., . . . Gray, J. A. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society of London – B, 265,* 1809–1817. doi:10.1098/rspb.1998.0506

Phillips, M., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., . . . David, A. S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature, 389,* 495–498. doi:10.1038/39051

Powell, E. W., & Rorie, D. (1967). Septal projections to nuclei functioning in oxytocin release. *American Journal of Anatomy, 120,* 605–610. doi:10.1002/aja.1001200310

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a "theory of mind"? *Behavioral and Brain Sciences, 1,* 515–526. doi:10.1017/S0140525X00076512

Prewitt, C. M., & Herman, J. P. (1994). Lesion of the central nucleus of the amygdala decreases CRH mRNA expression and stress-induced ACTH release. *Annals of the New York Academy of Sciences, 746,* 438–440. doi:10.1111/j.1749-6632.1994.tb39279.x

Price, C. J., & Friston, K. J. (2005). Functional ontologies for cognition: The systematic definition of structure and function, *Cognitive Neuropsychology, 22,* 262–275. doi:10.1080/02643290442000095

Reuter, M., Montag, C., Altmann, S., Bendlow, F., Elger, C., Kirsch, P., . . . Falk, A. (2009). Genetically determined differences in human trust behavior: The role of the oxytocin receptor gene. *NeuroPsychoEconomics Conference Proceedings,* 21.

Riedl, R., Hubert, M., & Kenning, P. (2010). Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. *MIS Quarterly, 34,* 397–428.

Righetti, F., & Finkenauer, C. (2011). If you are able to control yourself, I will trust you: The role of perceived self-control in interpersonal trust. *Journal of Personality and Social Psychology, 100,* 874–886. doi:10.1037/a0021827

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., & Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron, 35,* 395–405. doi:10.1016/S0896-6273(02)00755-9

Rilling, J. K., & Sanfey, A. G. (2011). The neuroscience of social decision-making. *Annual Review of Psychology, 62,* 23–48. doi:10.1146/annurev.psych.121208.131647

Rimmele, U., Spillmann, M., Bärtschi, C., Wolf, O. T., Weber, C. S., Ehlert, U., & Wirtz, P. H. (2009). Melatonin improves memory acquisition under stress independent of stress hormone release. *Psychopharmacology, 202,* 663–672. doi:10.1007/s00213-008-1344-z

Rodrigues, S. M., Saslow, L. R., Garcia, N., John, O. P., & Keltner, D. (2009). Oxytocin receptor genetic variation relates to empathy and stress reactivity in humans. *Proceedings of the National Academy of Sciences, 106,* 21437–21441. doi:10.1073/pnas.0909579106

Rostene, W., Sarrieau, A., Nicot, A., Scarceriaux, V., Betancur, C., Gully, D., . . . Berod, A. (1995). Steroid effects on brain functions: An example of the action of glucocorticoids on central dopaminergic and neurotensinergic systems. *Journal of Psychiatry & Neuroscience, 20,* 349–356.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. F. (1998). Not so different after all: A

cross-discipline view of trust. *Academy of Management Review, 23,* 393–404. doi:10.5465/AMR.1998.926617

Rushton, J. P., & Ankney, C. D. (2007). The evolution of brain size and intelligence. In S. M. Platek, J. P. Keenan, & T. K. Shackelford (Eds.), *Evolutionary cognitive neuroscience,* (pp. 121–161). Cambridge, MA: MIT Press.

Rushton, J. P., & Jensen, A. R. (2008). James Watson's most inconvenient truth: Race realism and the moralistic fallacy. *Medical Hypotheses, 71,* 629–640. doi:10.1016/j.mehy.2008.05.031

Sanchez-Pages, S., & Turiegano, E. (2010). Testosterone, facial symmetry and cooperation in the prisoners' dilemma. *Physiology & Behavior, 99,* 355–361. doi:10.1016/j.physbeh.2009.11.013

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300,* 1755–1758. doi:10.1126/science.1082976

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593–1599. doi:10.1126/science.275.5306.1593

Schultz, W. (2000). Multiple reward signals in the brain. *Nature Reviews Neuroscience, 1,* 199–207. doi:10.1038/35044563

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36,* 241–263. doi:10.1016/S0896-6273(02)00967-4

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology, 57,* 87–115. doi:10.1146/annurev-.psych.56.091103.070229

Seppänen, R., Blomqvist, K., & Sundqvist, S. (2007). Measuring inter-organizational trust: A critical review of the empirical research in 1990–2003. *Industrial Marketing Management, 36,* 249–265. doi:10.1016/j.indmarman.2005.09.003

Shahrokh, D. K., Zhang, T. Y., Diorio, J., Gratton, A., & Meaney, M. J. (2010). Oxytocin-dopamine interactions mediate variations in maternal behavior in the rat. *Endocrinology, 151,* 2276–2286. doi:10.1210/en.2009-1271

Siegal, M., & Varley, R. (2002). Neural systems involved in "theory of mind." *Nature Reviews Neuroscience, 3,* 463–471.

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., & Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron, 41,* 653–662. doi:10.1016/S0896-6273(04)00014-5

Singer, T. (2009). Understanding others: Brain mechanisms of theory of mind and empathy. In P. W. Glimcher, C. F. Camerer, E. Fehr, & R. A. Poldrack (Eds.), *Neuroeconomics: Decision making and the brain* (pp. 251–268). Amsterdam, The Netherlands: Elsevier.

Skuse, D. H., & Gallagher, L. (2011). Genetic influences on social cognition. *Pediatric Research, 69,* 85R–91R. doi:10.1203/PDR.0b013e318212f562

Smith, J. H. (2010). Trusting the avatar. *Games and Culture, 5,* 298–313. doi:10.1177/1555412009359764

Sturgis, P., Read, S., Hatemi, P. K., Zuh, G., Trull, T., Wright, M. J., & Martin, M. G. (2010). A genetic basis for social trust. *Political Behavior, 32,* 205–230. doi:10.1007/s11109-009-9101-5

Stuss, D. T., Gallup G. G. Jr., & Alexander, M. P. (2001). The frontal lobes are necessary for "theory of mind." *Brain, 124,* 279–286. doi:10.1093/brain/124.2.279

Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior, 59,* 364–382.

Swan, J. E., Bowers, M. R., & Richardson, L. D. (1999). Customer trust in the salesperson: An integrative review and meta-analysis of the empirical literature. *Journal of Business Research, 44,* 93–107. doi:10.1016/S0148-2963(97)00244-0

Takahashi, T., Ikeda, K., Ishikawa, M., Kitamura, N., Tsukasaki, T., Nakama, D., & Kameda, T. (2005). Interpersonal trust and social stress-induced cortisol elevation. *Neuroreport, 16,* 197–199. doi:10.1097/00001756-200502080-00027

Takahashi, T., Ikeda, K., Ishikawa, M., Tsukasaki, T., Nakama, D., Tanida, S., & Kameda, T. (2004). Social stress-induced cortisol elevation acutely impairs social memory in humans. *Neuroscience Letters, 363,* 125–130. doi:10.1016/j.neulet.2004.03.062

Takahashi, T. (2005). Social memory, social stress, and economic behaviors. *Brain Research Bulletin, 67,* 398–402. doi:10.1016/j.brainresbull.2005.06.006

Taylor, S. E., Burklund, L. J., Eisenberger, N. I., Lehman, B. J., Hilmert, C. J., & Lieberman, M. D. (2008). Neural bases of moderation of cortisol stress responses by psychosocial resources. *Journal of Personality and Social Psychology, 95,* 197–211. doi:10.1037/0022-3514.95.1.197

Theodoridou, A., Rowe, A. C., Penton-Voak, I. S., & Rogers, P. J. (2009). Oxytocin and social perception: Oxytocin increases perceived facial trustworthiness and attractiveness. *Hormones and Behavior, 56,* 128–132. doi:10.1016/j.yhbeh.2009.03.019

Thompson, R. R., George, K., Walton, J. C., Orr, S. P., & Benson, J. (2006). Sex-specific influences of vasopressin on human social communication. *Proceedings of the National Academy of Sciences, 103,* 7889–7894. doi:10.1073/pnas.0600406103

Thompson, R., Gupta, S., Miller, K., Mills, S., & Orr, S. (2004). The effects of vasopressin on human facial responses related to social communication. *Psychoneuroendocrinology, 29,* 35–48. doi:10.1016/S0306-4530(02)00133-6

Tillfors, M., Furmark, T., Marteinsdottir, I., Fischer, H., Pissiota, A., Langstrom, B., & Fredrikson, M. (2001). Cerebral blood flow in subjects with social phobia during stressful speaking tasks: A PET

study. *American Journal of Psychiatry, 158,* 1220–1226. doi:10.1176/appi.ajp.158.8.1220

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social Cognitive and Affective Neuroscience, 3,* 119–127. doi:10.1093/scan/nsn009

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/ avoidance behaviors. *Annals of the New York Academy of Sciences, 1124,* 208–224. doi:10.1196/annals.1440.012

Tost, H., Kolachana, B., Hakimi, S., Lemaitre, H., Verchinski, B. A., Mattay, V. S., . . . Meyer-Lindenberg, A. (2010). A common allele in the oxytocin receptor gene (OXTR) impacts prosocial temperament and human hypothalamic-limbic structure and function. *Proceedings of the National Academy of Sciences, 107,* 13936–13941. doi:10.1073/pnas.1003296107

Toulmin, S. E. (1958). *The use of arguments.* Cambridge, MA: Cambridge University Press.

Tricomi, E. M., Delgado, M. R., & Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron, 41,* 281–292. doi:10.1016/S0896-6273(03)00848-1

Turkheimer, E. (1998). Heritability and biological explanation, *Psychological Review, 105,* 782–791. doi:10.1037/0033-295X.105.4.782-791

Verbalis, J. (1999). The brain oxytocin receptor(s?). *Frontiers in Neuroendocrinology, 20,* 146–156. doi:10.1006/frne.1999.0178

Vilares, I., Dam, G., & Kording, K. (2011). Trust and reciprocity: Are effort and money equivalent? *PLoS ONE, 6,* e17113. doi:10.1371/journal.pone.0017113

Welch, M. R., Rivera, R., Conway, B. P., Yonkoski, J., Lupton, P. M., & Giancola, R. (2005). Determinants and consequences of social trust. *Sociological Inquiry, 75,* 453–473. doi:10.1111/j.1475-682X.2005.00132.x

Wermter, A. K., Kamp-Becker, I., Hesse, P., Schulte-KÖrne, G., Strauch, K., & Remschmidt, H. (2010). Evidence for the involvement of genetic variation in the oxytocin receptor gene (OXTR) in the etiology of autistic disorders on high functioning level. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics, 153B,* 629–639.

Weyers, P., Mühlberger, A., Hefele, C., & Pauli, P. (2006). Electromyographic responses to static and dynamic avatar emotional facial expressions. *Psychophysiology, 43,* 450–453. doi:10.1111/j.1469-8986.2006.00451.x

Winston, J. S., Strange, B. A., O'Doherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5,* 277–283. doi:10.1038/nn816

Wood, R. M., Rilling, J. K., Sanfey, A. G., Bhagwagar, Z., & Rogers, R. D. (2006). Effects of tryptophan depletion on the performance of an iterated prisoner's dilemma game in healthy adults. *Neuropsychopharmacology, 31,* 1075–1084. doi:10.1038/sj.npp.1300932

Yanagisawa, K., Masui, K., Furutani, K., Nomura, M., Ura, M., & Yoshida, H. (2011). Does higher general trust serve as a psychosocial buffer against social pain? An NIRS study of social exclusion. *Social Neuroscience, 6,* 190–197. doi:10.1080/17470919.2010.506139

Zak, P. J., Borja, K., Matzner, W. T., & Kurzban, R. (2005a). The neuroeconomics of distrust: Sex differences in behavior and physiology. *American Economic Review, 95,* 360–363. doi:10.1257/000282805774669709

Zak, P. J., & Fakhar, A. (2006). Neuroactive hormones and interpersonal trust: International evidence. *Economics and Human Biology, 4,* 412–429. doi:10.1016/j.ehb.2006.06.004

Zak, P. J., & Knack, S. (2001). Trust and growth. *The Economic Journal, 111,* 295–321. doi:10.1111/1468-0297.00609

Zak, P. J., Kurzban, R., Ahmadi, S., Swerdloff, R. S., Park, J., Efremidze, L., . . . Matzner, W. T. (2009). Testosterone administration decreases generosity in the ultimatum game. *PLoS ONE, 4,* e8330. doi:10.1371/journal.pone.0008330

Zak, P. J., Kurzban, R., & Matzner, W. T. (2004). The neurobiology of trust. *Annals of the New York Academy of Sciences, 1032,* 224–227. doi:10.1196/annals.1314.025

Zak, P. J., Kurzban, R., & Matzner, W. T. (2005b). Oxytocin is associated with human trustworthiness. *Hormones and Behavior, 48,* 522–527. doi:10.1016/j.yhbeh.2005.07.009

Zak, P. J., Stanton, A. A., & Ahmadi, S. (2007). Oxytocin increases generosity in humans. *PLoS ONE, 2,* e1128. doi:10.1371/journal.pone.0001128

Zak, P. J. (2003). Trust. *Capco Institute Journal of Financial Transformation, 7,* 13–21.

Zak, P. J. (2011). The physiology of moral sentiments. *Journal of Economic Behavior & Organization, 77,* 53–65. doi:10.1016/j.jebo.2009.11.009

Zethraeus, N., Kocoska-Maras, L., Ellingsen, T., von Schoultz, B., Hirschberg, A., & Johannesson, M. (2009). A randomized trial of the effect of estrogen and testosterone on economic behaviour. *Proceedings of the National Academy of Sciences, 106,* 6535–6538. doi:10.1073/pnas.0812757106

Received June 10, 2011
Revision received September 19, 2011
Accepted October 13, 2011 ■