Trust in responsive regulation theory: a critical appraisal from a trust perspective

Frédérique E. Six
Department of Governance Studies, Faculty of Social Sciences
VU University Amsterdam
De Boelelaan 1081, 1081 HV AMSTERDAM, the Netherlands
E f.e.six@vu.nl; T + 3120 598 6908; M +316 4821 1892

Paper presented at ECPR Standing Group on Regulation & Governance
Exeter, June 27-29, 2012

Abstract
Trust plays an important role in Responsive Regulation Theory (RRT). This study critically appraises RRT from a trust perspective, using insights from recent trust research. Three points of critique are made. First, in RRT cooperation is equated with trust, which is outdated. Furthermore, trust based on the benign big gun is deterrence-based trust, which nowadays is not considered trust, but low distrust. Second, within the pyramid, trust and control are treated as substitutes, while recent research shows that trust and control are better conceptualized as complements. Third, processes of trust building and repair are underplayed. Too little attention is paid to interpersonal and reflexive competencies and RRT assumes no influence of regulator generalized (dis)trust in regulatees. The conclusion is that the central role for trust in responsive regulation is as valid in 2012 as it was in 1992. Introducing insights from recent trust research helps to improve RRT.

Key words: trust, control, responsive regulation theory, relational signalling, regulation
Trust in responsive regulation theory: a critical appraisal from a trust perspective

Trust has been found to be important in regulatory relations (Braithwaite & Makkai 1994; Murphy 2004; Murphy et al. 2009). Responsive Regulation Theory (RRT) places trust at its heart. RRT’s key advice to regulators is to adopt “a dynamic regulatory strategy of dialogue and trust as a first choice followed by escalation to more punitive regulation when trust is abused” (Braithwaite & Makkai 1994, p. 1). Braithwaite (2002, p. 36) showed that “when lawbreakers believe they are treated as trustworthy by those who regulate them, they are more likely to comply with the law in future”. The empirical support for these claims comes from several studies. For example, Braithwaite and Makkai (1994) empirically investigated the impact of trust on compliance in the inspection of Australian aged care homes. They found that regulatees’ perception of regulator trust in them increased regulatee compliance at the next inspection (18-20 months later). Murphy (2004) investigated the effect of the trust Australian tax payers have in the Tax Office. She found that regulatee trust in the regulator leads to higher voluntary compliance. Procedural justice is an important antecedent to this regulatee trust.

According to RRT, successful cooperative regulation may be achieved when (1) regulators use a tit-for-tat strategy; (2) have access to a hierarchical range of sanctions (the enforcement pyramid); and (3) the punitiveness of the most severe sanctions are sufficiently high (Ayres & Braithwaite 1992, p. 40). The tit-for-tat strategy is based on game theory. In this strategy actors start with cooperation and if the other actor reciprocates, a cooperative exchange may develop. If the other actor responds with defection, taking advantage of the initial trust, then the first actor retaliates with defection. The actor copies the other actor’s action. Ayres and Braithwaite (1992) called this cooperation trust. The general principle of responsiveness is to tailor the regulator intervention style to the propensity and capacity of regulatees to comply. Regulatees are trustworthy when they have both the competence and positive intentions to do what the regulator expects from them, i.e. comply (cf Nooteboom 2002). Authors critically reviewing RRT struggle with how to go about making this judgment about the regulatees’ trustworthiness based on RRT, or in regulatory terminology, “how to decide when to punish and when to persuade” (Braithwaite 2002, p. 29). Regulators in RRT are assumed to have the skill and will to act responsibly towards regulatees, to be tough on the bad guys and soft on the good guys.

This study critically appraises RRT from a trust perspective, using current insights from trust research. Since the publication of Ayres and Braithwaite’s book on RRT two decades ago, trust has received burgeoning scholarly attention, in particular in the field of organization and management research (see for example special issues in Academy of Management Review (1998, 2009), International Sociology (2005) or Group & Organization Management (2007)). The research question is: how can current insights from trust research enrich and improve responsive regulation theory?

Three points of critique are made. First, the way RRT conceptualizes trust needs improvement. In the early 1990s these conceptualizations of trust were more commonplace, but twenty years later they are no longer accepted as appropriate (Rousseau et al. 1997; Möllering 2006a). Second, RRT implicitly assumes trust (persuasion) and control (deterrence) are substitutes, in other words when one trusts, one does not control and when one controls
there is no trust. More and more empirical evidence is emerging in support of the perspective that trust and control are not substitutes but complements. Trust and control may be applied simultaneously and may reinforce each other (Das & Teng 1998; Möllering 2005; Weibel, 2007). Six (forthcoming) has argued that this perspective also applies to regulation. The third point of critique is the observation that insufficient attention is paid to processes of trust building and repair. The implication is that important concepts are currently absent or underplayed in RRT, in particular regulator generalized trust or distrust in regulatees and the interpersonal and reflexive competencies of both regulator and regulatee.

RRT, as formulated in Ayres and Braithwaite (1992) and subsequent elaborations by Braithwaite and colleagues (e.g., Braithwaite et al. 2007), is based on extensive empirical research performed by Braithwaite and colleagues over many years. RRT, itself, however, does not come with testable hypotheses. Several authors have derived hypotheses from RRT and tested them empirically. Operationalization of many key concepts in RRT proved problematic and the results of the tests provided no, or at best weak, support for RRT. This study argues that incorporation of current insights from trust research into RRT, as proposed here, will help solve (some of) these theoretical puzzles.

The conclusion from this critical appraisal is that the central role for trust in responsive regulation is as valid in 2012 as it was in 1992. Introducing insights from current trust research helps to enrich and improve RRT.

I first address these three points of critique in more detail in the next three sections. Then I show how these points may help to explain some current problems with empirical tests of RRT. Finally conclusions are drawn with implications for theory and practice.

**Conceptualization of trust**

Trust has been studied in many different academic disciplines and this has resulted in many different definitions. A definition that is widely accepted is that of Rousseau et al. (1998 p. 395), “trust is a psychological state comprising the intention to accept vulnerability based upon the positive expectations of the intentions or behavior of another”. Trust is a judgment based on knowledge about another party’s trustworthiness or ‘untrustworthiness’, respectively, but this knowledge is incomplete. Trust implies that there is uncertainty about the trustee’s future behavior. Möllering (2006a, p. 356) argued that trust inevitably involves a leap of faith in which the “irreducible social vulnerability and uncertainty [are suspended] as if they were favorably resolved.” The crux of trust is that it is not simply the more trust the better it is. Trust is not always warranted. But at the same time, generic distrust is also not warranted (Hardin 1993). The key question is: how well do I know the other so that I can trust the other to do what I need her to do? (Gabarro 1978). Trust and distrust are both selective: How well can the trustworthy be distinguished from the not trustworthy? How much tolerance for errors in the judgment is acceptable?

RRT conceptualizes trust using game theory. Regulatees are trusted as long as they comply. Regulators are to use the tit-for-tat strategy and start a new regulatory relation with trust. In the days that RRT was first published (early 1990s) it was fairly common to use game theory and the ways in which cooperation may be stimulated through games, in trust research. Cooperation in such games was seen as the cooperating actor trusting the other actor. Nowadays, this is considered to be incorrect (Möllering 2006a). Furthermore, the enforcement pyramid that is at the centre of RRT and through which regulators navigate up or down using
a tit-for-tat strategy is based on “the benign big gun”. “Regulators will be more able to speak softly when they carry big sticks” (Ayres & Braithwaite 1992, p. 19). Again, at the time of writing, trust based on deterrence, which is what is described here, was accepted as a type of trust. Nowadays, deterrence-based trust is no longer considered trust, but is considered low distrust (Rousseau et al. 1998). Using sanctions as a deterrent to achieve cooperation is not the same as trusting. Cooperation is not the same as trust. “Not all cooperation requires that the actors trust each other and not all actors who trust each others cooperate necessarily” (Möllering 2006a, p. 41). Cooperation tends to lead to trust.

Deterrence-based trust is about taking away negative intentions in the trustee by putting in place sanctions for when the trust is violated. Trust, however, according to the definition, is about beliefs in the presence of positive intentions, which is not the same (Rousseau et al. 1998). Lindenberg (2000) argued that one first needs to take away distrust before trust can be built. He introduced the notion of legitimate distrust situations, “situations where any explicit or implicit promise […] is blatantly against the self-interest of the promising party” (Lindenberg 2000, p. 12). Because the distrust is seen as legitimate, that is “reasonable observers would say that any other reasonable person put into this situation” would judge similarly, remedies can be relationally neutral, meaning that the distrusting individual can “claim the necessity of remedies, pinpoint to a menu of solutions and show good faith at the same time” (Lindenberg 2000, p. 12). In other words, sanctions may need to be present to take away distrust before trust can be built. This issue is related to the relation between trust and control, which is discussed below. Lewicki et al. (1998) argued that distrust and trust are two separate concepts and not two ends of one continuum. Low trust is not the same as high distrust. Just as actors require knowledge about the other actor before they can have high trust, they also require knowledge to actively distrust. If this knowledge is missing, they will most likely be low on trust and distrust.

**Relation between trust and control**

The second point of critique concerns how the relation between trust and control is conceptualized. Much of the debate about how best to achieve regulatee compliance to regulator objectives is in essence about this relation. One of the reasons there is a continuing debate is because researchers argue based on different perspectives on the relation (Das & Teng 1998; Möllering 2005). If one interprets the cooperative enforcement style as trust, as Braithwaite and Makkai (1994) do, and the punitive or repressive style as control, then RRT assumes both trust and control are needed for effective regulation. It furthermore assumes that trust and control are substitutes; when one applies a cooperative enforcement style, one does not at the same time apply a punitive style, and vice versa. At the lower levels of the enforcement pyramid, cooperative, trusting styles are applied, while at the top of the pyramid punitive, controlling styles are applied (Ayres & Braithwaite 1992).

More and more empirical evidence is emerging in support of the perspective that trust and control are not substitutes but complements. Trust and control may be applied simultaneously and may reinforce each other (Das & Teng 1998; Möllering 2005; Weibel 2007; Weibel et al. 2009). Six (forthcoming) has argued that this perspective also applies to regulation. Trust and control are complements as long as the control measures are experienced by regulatees as supporting their self-determination. Controls may thus strengthen voluntary compliance. To strengthen regulatee voluntary compliance, regulators should always tell regulatees they have a choice, give positive feedback where possible and treat regulatees with procedural justice and respect; this can be combined with implementing controls.
In general, control can be defined as a process of regulating other people’s behaviour to make it more predictable and achieve the outputs that the controller wants to achieve (e.g., Das & Teng 1998). It addresses the question: how do controllers generate and retain the commitment of other people to tasks that call for heavy investments of time and effort and yet are difficult to prescribe, program and monitor? Two basic types of control exist in the literature: external measure-based control and internal value-based control (Eisenhardt 1985). The former has been called formal, objective coercive, restrictive or bureaucratic control (Adler & Borys 1996; Das & Teng 1998; Elias 2009). Bureaucratic control focuses on the enforced obedience to the rules (Adler & Borys 1996). Internalized control has been called normative, clan, enabling, promotive or social control (Adler & Borys 1996; Das & Teng 1998; Elias 2009) and is more directed to an internalization of the rules and an identification with the organization or public interest. Promotive control refers to a controlling actor attempting “to influence [a controlled actor] while listening to his or her opinions and including him or her in the decision-making processes” (Elias 2009, p. 370). Enabling control is aimed at enabling controlled actors to “master their tasks” (Adler & Borys 1996, p. 62).

External measure-based controls tend to diminish self-determination, while internal value-based controls tend to enhance self-determination (Deci & Ryan 2000; Ryan & Deci 2000). This is in line with research showing that governance on the basis of extrinsic motivation (i.e. external control) may crowd out governance on the basis of internalized motivation, which is crucial for trust (e.g., Frey 1997; Frey & Jegen 2001; Ryan & Deci 2000). Tenbrunsel and Messick’s (1999) research into the effect of decision frames on compliance provides explanations for the underlying mechanism. The presence of regulation without sanctions leads to fairly high levels of compliance, because many regulatees decide to comply from a moral decision frame. When sanctions are introduced, most regulatees approach the decision to comply or not from a calculative business frame. The decision to comply is then strongly influenced by the severity and likelihood of the sanction (i.e. external control).

Whether control crowds out trust depends on whether certain conditions are met (Weibel 2007). Social determination theory (SDT) may be used to derive these conditions (Deci & Ryan 2000; Ryan & Deci 2000). Six (forthcoming) applied SDT to regulation theory to identify the conditions under which regulation may enhance regulatee self-determination and as a result lead to higher voluntary compliance. The conceptual model is derived from Weibel’s (2007) model for the relation between trust and control in the hierarchical relationship between managers and subordinates. Below only a summary is given, see Six (forthcoming) for more details.

**Self Determination Theory (SDT) and regulation research**

Self-determined activities are “activities that people do naturally and spontaneously when they feel free to follow their inner interests” (Deci & Ryan 2000, p. 234). When performing self-determined activities, people have an internal perceived locus of causality. When, however, external rewards are introduced for these activities, “people tend to feel controlled by the rewards, prompting a shift in the perceived locus of causality for the behavior from internal to external” (Deci & Ryan 2000, p. 234). SDT posits that the more people internalize and integrate socially sanctioned values and norms, the more self-determined their actions and the more likely it is that their locus of causality is internal. SDT is based on supporting the three basic psychological needs that people have: autonomy, competence and relatedness. According to SDT, it is part of the way human organisms are designed and work that they “want to engage [in] interesting activities, to exercise capacities, to pursue connectedness in
social groups, and to integrate intrapsychic and interpersonal experiences into a relative unity” (Deci & Ryan 2000, p. 229).

Weibel’s (2007) premise was that the more control measures support people’s basic needs for competence, autonomy and relatedness, the more likely it is that people internalize and integrate the socially sanctioned values and norms. And the more subordinates internalize these organizational values, the more trustworthy they are to managers and other organizational members. People want to feel part of and accepted by the social groups they live in. Relatedness is enhanced when regulatees are able to internalize the values and regulations of their social groups (Deci & Ryan 2000, p. 238). Relatedness is also enhanced if the authority within the group, such as a leader within an organization, imposes the formal controls with positive intentions, in other words if he trusts rather than distrusts those controlled (Weibel, 2007 Figure 1). Competence is enhanced the more regulatees are supported in their ability “to understand or grasp the meaning or rationale behind the regulation and an ability to enact it” (Deci & Ryan 2000, p. 238). This process is enhanced the more regulators give constructive feedback (Weibel 2007, p. 509). The need for autonomy is supported the more regulatees are given the opportunity “to freely process and endorse transmitted values and regulations” and the opportunity to give “personal meaning and valence to acquired regulations” (Deci & Ryan 2000, p. 238). In work situations, this implies that the need for autonomy is supported the more regulatees participate in the development of formal controls and the execution of those controls (Weibel 2007, p. 508). Autonomy is furthermore enhanced, the more rewards and sanctions “are administered on the basis of a broad and more subjective evaluation”, rather than “strictly contingent on the evaluation of standard fulfilment” (Weibel 2007, p. 509).

Six (forthcoming) argued that these conditions, identified by Deci and Ryan (2000) and Weibel (2007), are also applicable to the hierarchical relationship between regulator and regulatee, the focus of RRT. When the actors subjected to the formal controls (the ‘regulatees’) internalize and integrate the values of the actors imposing the formal controls (the ‘regulators’), then the regulatees are more likely to comply with the regulation in a self-determined way, on their own volition. The quality of the compliance will be higher. SDT shows that value internalization is stimulated when the regulatees feel their needs for autonomy, competence and relatedness are supported by the regulators (Deci & Ryan 2000; Ryan & Deci 2000). The conditions that make formal controls more or less in support of value internalization and self-determination have been discussed in the previous section and are presented in the first column of Table 1. In the second column, concepts from regulation theory that relate to each of the self-determination enhancing conditions are listed.

PROCESSES OF BUILDING AND REPAIRING TRUST

A third point of critique is the observation that RRT pays insufficient attention to processes of trust building and repair. The implication is that important concepts are currently absent or underplayed in RRT. For trust to be built effectively and trust repair to have a chance - after (perceived) trust violations - interpersonal and reflexive competencies for both regulator and regulatee are important (Six & Sorge, 2008). RRT appears to assume that the regulator has the interpersonal and reflexive competencies to choose the appropriate strategies within the pyramid and communicate clearly with regulatees. Mascini and van Wijk (2009) showed that
these assumptions do not always hold empirically. In terms of regulatee competencies, attention is paid to the compliance related competencies (Winter & May 2001), but not to interpersonal and reflexive competencies. Again, Mascini and van Wijks’s research showed the relevance of these competencies. Furthermore, motivational postures and general propensity to trust the other actor are important in trust building. For regulatees this is acknowledged in Valerie Braithwaite’s motivational postures (2009) with disengagement resembling generalized distrust of government. For regulators, however, this is not problematized. RRT assumes no influence of regulator generalized trust or distrust towards regulatees.

RRT proposes that trust is important in regulatory relations as it fosters voluntary compliance. This study supports that proposition. But trust in regulatory relations does not necessarily occur naturally. It needs to be built and maintained. And since trouble in work relations is inevitable (Six & Skinner, 2010), trust may also need to be repaired at some point in time. Trust is a three-part relation: A trusts B to do C (Hardin 1993). The regulator wants to trust the regulatee that s/he will comply. The regulatee wants to trust the regulator that it will treat him/her fairly, with procedural justice (Haines & Gurney 2003; Murphy et al. 2009) and without unnecessary hassle. In a regulatory relation, the regulator is both trustor and trustee, as is the regulatee. A’s trust in B is to a large extent based on A’s perception of B’s trustworthiness. Does B have the competence (skill) and intentions (will) to do what A expects?

**Trust building processes**

The way in which regulator and regulatee interact influences the degree to which trust is built and maintained, both regulator’s trust in the regulatee as well as regulatee’s trust in the regulator. Trust is ultimately based on the perceptions of the trustor, not the intentions of the trustee, and multiple interpretations of the trustee’s actions and intentions may be possible. The interactive nature of the trust building process means that the trustor’s predisposing beliefs, such as generalized trust or distrust, are crucial as they determine his or her initial attitude. The self-fulfilling prophesy of beliefs (Zand 1972) means that a trustor’s beliefs are most likely to be confirmed through the impact of his or her actions – driven by these beliefs – on the trustee: in upward spiralling processes of positive trust experiences, A’s trust in B is confirmed, that is, B acts according to A’s pattern of expectations and trust is built. If A concludes that B is sufficiently trustworthy, A will act to make himself vulnerable to the actions of B; B in turn will perceive A’s action as indications of A’s trustworthiness and will likely act in line with A’s expectations, which will be perceived as confirmation of A’s initial trust (Figure 1).

---

**FIGURE 1 ABOUT HERE**

Trustors look for two things in the behaviour of trustees. First, they check if the behavior shows the competence to perform according to expectations (ability dimension of trustworthiness). Second, trustors look for signs in the behaviour of trustees for indications whether the trustee is interested in maintaining the relationship in the future (intentions dimension). These signs are called relational signals (Lindenberg 2000). Relational signals are ‘behavioural clues that allow us to make inferences about other people’s interest in maintaining a mutually rewarding social relationship with us’ (Wittek 1999, p. 8). A positive relational signal is any behaviour that reassures another actor and is perceived by the second
actor as an indication that the first actor wants to maintain a mutually rewarding relationship. A negative relational signal is any behaviour that makes another actor uncomfortable, perceiving this behaviour as an indication that the first actor does not want to maintain a mutually rewarding relationship. The notion of a relational signal reflects the insight that much of human behaviour, and particularly interpersonal interaction and communication, is not only about the exchange of information, but also about defining the nature of the relationship between the actors involved. What relational signals mean depends on shared understandings, built up through interaction (Wittek 1999). When signals are meaningful in the interaction between two or more actors, they include ‘expressions given off’, seemingly involuntary aspects like blushing or other body language. These are only partly open to manipulation. This is important, since actors who have no relational interest can and probably will exploit relational signalling (Deutsch 1973). It is difficult for most people to pretend consistently that they have good intentions when in fact they do not, as they will nearly always give off signals to the contrary. To be credible, the trustee needs to regularly signal interest in the wellbeing of the trustor (Lindenberg 2000). Relational signalling theory (e.g., Lindenberg 2000; Wittek 1999) also stresses the dual function of relational signals: positive relational signals show that the sending actor wishes to act appropriately and is concerned about the wellbeing of the receiving actor, and they strengthen a corresponding mind-set in the receiver. In regulatory relations, regulators represent the public interest and societal needs, such as protection against major hazards, ensuring minimum quality standards in health, care and education. Regulatees need to signal interest in these societal needs, rather than regulator wellbeing.

The dynamic described above has been empirically tested in interpersonal work relations and interorganizational business relations (e.g., Zand 1972; Zaheer et al. 1998), but not yet in regulatory relations. John Braithwaite and colleagues have repeatedly stressed the importance of the actual interaction between the inspectors and the members of the regulatee organization they are in contact with (e.g., Braithwaite et al. 2007). Implicitly they also talk about relational signals. Mascini and van Wijk (2009) do so explicitly without elaborating the concept.

Relational signals are in the eye of the beholder. In other words, whether an action carries positive or negative (or ambiguous) relational signals is determined by the perception of the receiver rather than the intention of the sender. Mascini and van Wijk’s (2009, p. 40) research provides several examples, “entrepreneurs react based on how they perceived an inspection rather than based on the inspector’s intention … and … entrepreneurs certainly do not always perceive the enforcement style in the same manner as was intended by the inspector.” Braithwaite and Makkai’s (1994) finding that regulatees’ perception of regulator trust in them increased future compliance by regulatees also illustrates this point.

**Generalized trust and distrust**

An important driver of trust building is the general propensity to trust others (Mayer et al. 1995). In regulatory relations regulator’s generalized trust, or distrust, in regulatees is also relevant, as is regulatee’s generalized trust, or distrust, in regulators and government. In political science and public management research, much attention is paid to citizen’s trust in government, although a lot of criticism is possible regarding the way empirical research is generally conducted (e.g. Van de Walle and Six 2011). In regulation research attention is paid to regulatee attitudes or motivational postures. The disengagement posture resembles high generalized distrust of regulators, and possibly government more generally (V. Braithwaite 2009).
In regulation research, as is also the case in political science or public management research, hardly any attention is paid to regulator generalized trust, or distrust in regulatees. If measured at all, it is not called trust. Mascini and van Wijk (2009, p. 33) found that 47% of inspectors believed “there are more bad companies than good companies”, i.e. showed generalized distrust towards regulatees, while only 31% believed the opposite, showing generalized trust. Inspectors with generalized distrust of regulatees will most likely signal their distrust to regulatees, probably without being aware of it. Even if they go through the motions of the prescriptions of RRT, regulatees will most likely pick up the distrust and not perceive the rest. Equally, if regulatees have the motivational posture of disengagement as dominant and have a generalized distrust of government and thus also of inspectors visiting their premises, they are unlikely to perceive any positive relational signal the inspector sends, however hard the inspector tries (cf Six 2005). As Mascini and van Wijk (2009, p. 40) showed, “A negative relational signal frequently dominated an entrepreneur’s overall perception of the inspector’s conduct and resulted precisely in the negative effects the inspectors were trying to prevent with their predominantly persuasive approach.”

Attention to relationship
Trust research also shows that to build trust attention must not only be paid to the task at hand, but also to the relationship itself. Giving positive feedback and praise is among the most important relational behaviours that help build trust (e.g., Six & Sorge 2008). As shown below, the crux in repairing trust after a (perceived) trust violation is making clear that the person is OK, but the problem is with the behaviour (e.g., Six & Skinner 2010). Makkai and Braithwaite (1994) showed the importance of distinguishing between showing disapproval of the behaviour and the person. Reintegrative shaming disapproves of the behaviour while accepting the person, and disintegrative shaming disapproves of both behaviour and person. The former led to much better compliance in the future.

Dealing with (perceived) trust violations
If all goes well and regulators find regulatees in compliance, trust is likely to build between regulator and regulatee. But what happens to the relationship when regulators find something that appears to be not in compliance with the law? Equally, what happens in the relationship when regulatees feel they are not treated fairly? What happens when such trouble occurs is of central concern in the context of relationship maintenance and repair and (unnecessary) damage to the relationship can be avoided if such trouble is dealt with effectively. Trouble is the disruption of the flow of expectations which, at least initially, is experienced as unpleasant (Six & Skinner 2010). Trouble events disrupt the trustor’s expectations of trustworthiness; such an event ‘produces a sense of disruption of trust, or profound confusion’ (Zucker 1986, p. 59). However, episodes of trouble occur for a number of reasons and are not necessarily the intentional violation or abuse of trust (described as betrayal by Elangovan and Shapiro 1998) or violation of trust such as occurs ‘when the person’s actions are freely chosen’ (Lewicki and Bunker 1996, p. 127), although this may be the conclusion drawn by the trustor. Events experienced as trouble by the trustor may be due to the presence of radical uncertainty (Knight 1921), random lapses in efficiency (Hirschman 1970), short-term temptations (Lindenberg 2000) or mishap situations. The trustor who experiences trouble, apparently caused by a trustee, needs to decide what this means for the trust he or she can place in the trustee in the future and how to respond to the event. What happens in the critical moments when trouble occurs is crucial in determining the nature of the trust in the relationship in the future. Fundamental to the trouble event and its aftermath are the messages conveyed by
verbal and non-verbal actions of those involved. Importantly, it is solely the recipient’s perception that determines whether these messages are interpreted as positive or negative and what their significance is for the recipient’s wellbeing (Wittek 1999). Again, relational signals are crucial when dealing with trouble events and (perceived) trust violations.

The regulator needs to decide after an incident of noncompliance whether the lapse is due to an unfortunate mishap or misunderstanding, an honest mistake, or due to wilful calculation. When not dealt with properly, chances are that the regulator makes an incorrect assessment. If he concludes that trust has been violated while this is not the case (e.g. it was an honest mistake), he will from now on act based on this unwarranted distrust. And given the precarious nature of the self-fulfilling prophecy in the trust process, it is unlikely that trust is restored. If on the other hand, the regulator concludes that the regulatee is still trustworthy, while this is unwarranted, this misjudgement will be corrected in the future when the noncompliance is repeated. Eventually the appropriate conclusion, warranted distrust, will be drawn. Unwarranted distrust is thus unlikely to be corrected while unwarranted trust is (cf Hardin 1993). The same dynamic applies to the regulatee, and even more strongly so, given citizens’ tendency to view government as a homogenous actor, which is incorrect. Distrusting regulatees will disengage from constructive interaction with government agencies and thus make responsive regulation impossible. And in most instances inspectors do not have the time available to spend with disengaged regulatees to take away the distrust and (re)build trust. Thus the dynamics of trust building and repair in regulatory relations makes it *a priori* more likely that relations end up in low trust, possibly even distrust, rather than achieve sustainable high levels of trust. High levels of trust are more likely when regulatees are large corporations because the likelihood that the competencies are present is higher and the frequency of interaction is likely to be higher than with small corporations or individual citizens (e.g., Mascini and van Wijk 2009).

**Positive feedback and praise**

Above, Makkai and Braithwaite’s research on reintegrative shaming was used to illustrate the importance of saying yes to the person, while saying no to the behaviour. Giving sincere positive feedback whenever possible is a strong way to signal that the actor is respected and accepted and makes it easier to achieve that disapproval of a certain behaviour is seen as just that and not rejection of the person. Makkai and Braithwaite (1993) showed how informal praise from an inspector to regulatees improved compliance. (Braithwaite et al. 2007, p. 138) though, found that inspectors often shy away from positive feedback.

**Procedural justice**

Several regulatory studies showed the importance of procedural justice (e.g., Haines & Gurney 2003; Murphy *et al.* 2009). Procedurally just behaviour by the regulator helps build trust because it sends positive relational signals. It signals support for relatedness and acceptance of the regulatee. This stimulates self-determination and voluntary compliance. Braithwaite *et al.* (2007, p. 211) add the importance of treating regulatees respectfully while firmly confronting noncompliance.

**Interpersonal and reflexive competencies**

For processes of trust building and repair to be successful, both regulatee and regulator need to have sufficient interpersonal and reflexive competencies. Regulator reflexive competencies are about the degree to which they reflect on their own generalized trust or distrust of regulatees and also their trust or distrust in the particular regulatee in question. How to interpret the mix of relational signals perceived? What level of trust or distrust is warranted?
Is some doubt about trust based on lack of competence or bad intentions? Reflexive competencies of regulatees address similar questions.

Interpersonal competencies are important to improve the relational signaling process. The more aware and competent each actor, the higher the likelihood that the actor sending the signals sends them clearly and with little ambiguity and noise. And also the higher the likelihood that the receiver of the relational signals perceives it as intended by the sender. Awareness of how trust building and repair processes work will also increase the likelihood that, in particular, regulators send positive relational signals whenever possible, through positive feedback and praise. Interpersonal competencies of regulators are also important to make good judgments of the degree to which regulatees are trustworthy (Heimer 2011). Braithwaite et al. (2007) refer in several places to the importance of such competencies, without explicitly integrating it into RRT. For example, “open communication from both sides helps build trust and fosters compliance” (Braithwaite et al. 2007, p. 124); “US inspectors receive training … which is unsuitable for preparing them for a dialogic role” (p. 138).

RRT and other regulation theories pay no attention to the importance of the interpersonal and reflexive competencies of regulatees. Attention is paid to the compliance related competencies of regulatees, do they know what to do in order to comply with the law? (Winter & May 2001). Mascini and van Wijk’s (2009) research among small entrepreneurs in food stores and restaurants showed that many did not speak the (Dutch) language. If one cannot communicate at all, a responsive regulatory approach aimed at building trust is not possible. In other instances, their research raises the question whether the inability of inspectors to get their nuanced message across was purely due to lack of communication skills on their part. It is also possible that the small entrepreneurs lacked the communication and reflexive competencies for a constructive dialogue with authorities.

The requirements with regard to interpersonal and reflexive competencies are high (Heimer 2011), but this is not unique to regulatory relations. All human interactions with high uncertainty, ambiguity and insufficient information require difficult assessments of the others’ trustworthiness.

**Potential of incorporation of current trust insights into RRT**

RRT, as formulated in Ayres and Braithwaite (1992) and subsequent elaborations by Braithwaite and colleagues (e.g., Braithwaite et al. 2007), is based on extensive empirical research performed by Braithwaite and colleagues over many years. RRT, itself, however, does not come with testable hypotheses. Several authors have derived testable hypotheses from RRT and tested them empirically. Operationalization of many key concepts in RRT proved problematic and the results of the tests provided no, or at best weak, support for RRT. Nielsen (2006, p. 395) tested whether inspectors were responsive in the way in which RRT prescribes. She investigated some 2,500 breaches in four different regulatory domains in Denmark. “The empirical analyses show that regulatory inspectors manage to act responsively, but only to a small degree and not necessarily in the way that the theories of responsive and tit-for-tat regulation recommend.” Mascini and van Wijk (2009, p. 27) investigated RRT’s prescription that “enforcers should not shift to coercing before it has become clear that persuading does not work”. They developed three presuppositions and found that all three are problematic when empirically tested in inspection relations of the Dutch Consumer Food and Product Safety Authority. “Enforcement agents apply different styles in comparable cases; they are impeded in applying the most appropriate style; and they
do not control the perverse consequences of their conduct because regulatees tend to perceive it as more coercive than intended by inspectors” (Mascini & Van Wijk 2009, p. 27). Nielsen and Parker (2009) empirically investigated to what degree regulators show tit-for-tat responsiveness or restorative justice responsiveness, and with what effect on regulatee compliance. They investigated 141 interactions between large Australian businesses and the Australian Competition and Consumer Commission, specifically investigations after accusations of breaching the Trade Practice Act. They found “little evidence of tit for tat responsiveness actually occurring in practice. To the extent that tit for tat responsiveness does exist, [they found] a small amount of evidence that is has the hypothesized effect on behaviour but not attitude. [They found] clearer evidence of restorative justice responsiveness having the hypothesized effect on attitudes, but not on behaviour” (Nielsen & Parker 2009, p. 376).

This study argues that some of these findings may be better explained when RRT is improved using insights from current trust research. For example, in RRT, regulator’s generalized trust or distrust is ignored, which could explain some of the findings in all three studies. The finding that tit for tat responsiveness improves regulatee behaviour but not attitude may be explained by the fact that this type of responsiveness is deterrence-based trust which is really low distrust rather than trust. The regulatee probably perceives regulator actions as purely controlling and without any attempt to build trust. Nielsen and Parker’s (2009) finding that restorative justice improves regulatee attitude towards inspectors but not regulatee behaviour, i.e. compliance, may be explained by showing that singularly positive, i.e. trusting, behaviour of the inspector is likely to create positive attitudes, but does nothing to improve behaviour. The inspector does not control. Finally, in the operationalization of the concepts to be able to measure them empirically, the distinction between an actor’s intention behind an action and how it is perceived by the other is usually ignored, while this is crucial for trust-building.

Short and Toffel (2010) studied the effect of the regulatory environment on the likely impact of self-regulation on compliance levels. They investigated how different enforcement strategies enhance or undermine normative motivations to self-regulate and comply. Organizations were more likely to “follow through on their commitments to self-regulate when they (and their competitors) are subject to heavy regulatory surveillance” and when explicit threats of sanctions are absent (p. 361). Explicit threats of sanctions carry negative relational signals and are deterrence-based controls. This will crowd out intrinsic motivations to comply and reduce trust. Surveillance on the other hand, especially routine regulatory inspections, “lack the ‘cuing’ or ‘framing’ effect of actual sanctions or direct threats” (Short & Toffel 2010, p. 371). Such routine surveillance irrespective of the trustworthiness of the regulatee is likely to be perceived as relationally neutral as long as inspectors behave respectfully and fair (cf Lindenberg’s legitimate distrust situations).

Conclusions
This study critically appraises RRT from a trust perspective, using insights from current trust research. Three points of critique are made. First, in RRT cooperation is equated with trust, which is outdated. Furthermore, trust based on the benign big gun is deterrence-based trust, which nowadays is not considered trust, but low distrust. Second, within the pyramid, trust and control are treated as substitutes, while empirical evidence shows that trust and control are better conceptualized as complements. Third, processes of trust building and repair are underplayed. Too little attention is paid to interpersonal and reflexive competencies and RRT
assumes no influence of regulator generalized trust or distrust. The conclusion is that the central role for trust in responsive regulation is as valid in 2012 as it was in 1992.

An important general critique of RRT has been that it came without propositions. And when testable hypotheses were derived and tested, the results were disappointing (Mascini & Van Wijk 2009; Nielsen 2006; Nielsen & Parker 2009). This study suggests that part of the explanation is that the way in which insights from trust research were integrated into RRT was incomplete. If the three points of critique presented above are integrated into RRT, theoretical models may be formulated with propositions. (Six forthcoming) has made a start for a model with propositions in which SDT is applied to the relation between trust and control in regulatory relations. This model needs to be empirically tested. A theoretical model needs to be developed for the interactive processes within the regulatory relations with relational signals as the central concept. What is the influence of relational signals within the regulatory relation on processes of trust building and repair?

This study assumed that the regulatee is a homogenous actor. If the regulatee is an individual (e.g., a tax payer or a recipient of state benefits) or a small business (e.g., a small food store owner) this is appropriate. Many regulatees, however, are larger organizations (e.g., care homes, schools, chemical facilities) and the concept of organizational trust becomes relevant. What makes such organizations trustworthy in the eyes of the regulator? Both Möllering (2006b) and Zaheer et al. (1998) point to the fact that organizational trust is based on the interpersonal trust between the individual inspectors and the individuals within the regulated organization that they are in direct contact with (the boundary spanners), and on trust in regulatee organizational systems. Organizational trust is seen to hinge on the “collective characteristics of an administrative organization and top management group which are not reducible to features of individual actors and which ensure some continuity of activities and direction when those actors change” (Whitley 1987, p. 133). Further research is needed to develop the elements that make larger organizations trustworthy to regulators. Management-based regulation (Coglianese & Lazer 2003; Gunningham & Sinclair 2009) and corporate compliance systems (Parker 2003; Parker & Nielsen 2009) are likely to play an important role, as well as the “tone at the top”. Regulators are also not homogenous actors. Individual inspectors represent organizations and different parts or individuals from the agency may respond differently. More importantly, many regulatees do not distinguish between different government agencies. So when different regulatory agencies interact with one regulatee and send different relational signals, the regulatee is likely to get confused and become suspicious, lowering trust and possibly strengthening distrust.

Effective responsive regulation is hard work because trust building and maintenance is hard work, in whatever relationship. The conditions surrounding many regulatory relations may make it even harder to effectively build and maintain trust where warranted. For example, the frequency of direct interaction is often low; many different government agencies may interact with the regulatee, sending different, possibly conflicting, relational signals; there are high barriers to exit most regulatory relations; and budget cuts and political decisions may make it more difficult for regulators to act responsively in accordance with RRT. This does not make the theory wrong nor should it lead to practitioners giving up and not even trying to implement RRT. However imperfect daily regulatory practice may be, there is (currently) no better alternative theory.

The conclusion from this critical appraisal of RRT using a trust perspective has two sides. On the one hand, it reconfirms the importance of trust in regulatory relations and RRT. It also
reconfirms the importance of responsiveness in regulation. On the other hand, it shows that actually putting RRT into practice is very complicated and challenging. Both the regulator and the regulatee must not only have the positive intentions to enter into a constructive, responsive relation, they must also have all the required competencies. These include those required for the task at hand as well as the interpersonal and reflexive competencies for trust building and repair. These practical challenges, however, do not diminish the theoretical strength of RRT. When testing (elements of) RRT one should be careful to distinguish between lack of proof due to poor implementation or due to theoretical shortcomings.

References


Table 1: Self-determination enhancing conditions and regulation theory

<table>
<thead>
<tr>
<th>SELF-DETERMINATION ENHANCING CONDITIONS</th>
<th>RELATED CONCEPTS IN REGULATION RESEARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relatedness</strong></td>
<td></td>
</tr>
<tr>
<td>- Internalize values and regulations of one’s social groups</td>
<td>- Social motivations, social norms (Winter &amp; May, 2001)</td>
</tr>
<tr>
<td>- Positive intentions regulator</td>
<td>- Regulator trust in regulatee (Murphy, 2004)</td>
</tr>
<tr>
<td></td>
<td>- Procedural justice (Murphy et al., 2009)</td>
</tr>
<tr>
<td><strong>Competence</strong></td>
<td></td>
</tr>
<tr>
<td>- Ability to understand or grasp meaning or rationale behind regulation</td>
<td>- Knowledge of rules (Winter &amp; May, 2001)</td>
</tr>
<tr>
<td></td>
<td>- Educating regulatees (Ayres &amp; Braithwaite, 1992)</td>
</tr>
<tr>
<td></td>
<td>- Speak and understand regulator language (Mascini &amp; van Wijk, 2009)</td>
</tr>
<tr>
<td>- Ability to enact regulation</td>
<td>- Capacity to comply (Winter &amp; May, 2001)</td>
</tr>
<tr>
<td>- Constructive feedback</td>
<td>- Constructive feedback (Braithwaite et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>- Praise (Makkai &amp; Braithwaite, 1993)</td>
</tr>
<tr>
<td><strong>Autonomy</strong></td>
<td></td>
</tr>
<tr>
<td>- Opportunity for regulatee to freely process and endorse transmitted values and regulation</td>
<td>- Self-regulation, enforced self-regulation (Ayres &amp; Braithwaite, 1992)</td>
</tr>
<tr>
<td>- Opportunity to give ‘personal’ meaning and valence to acquired regulation</td>
<td>- Self-regulation, enforced self-regulation (Ayres &amp; Braithwaite, 1992)</td>
</tr>
<tr>
<td>- Internal locus of causality</td>
<td>- Normative motivations (Winter and May, 2001)</td>
</tr>
<tr>
<td>- Degree of participation</td>
<td>- Enforced self-regulation (Ayres &amp; Braithwaite, 1992)</td>
</tr>
<tr>
<td></td>
<td>- Management-based regulation (Coglianese &amp; Lazer, 2003; Gunningham &amp; Sinclair, 2009; Parker, 2003)</td>
</tr>
<tr>
<td>- Combination of rewards/evaluation</td>
<td>- Restorative justice (Braithwaite, 2002; Macrory, 2006; Nielsen &amp; Parker, 2009)</td>
</tr>
<tr>
<td></td>
<td>- Procedural justice and control (Makkai &amp; Braithwaite, 1996)</td>
</tr>
</tbody>
</table>

Source: adapted from Six (forthcoming)
Figure 1: Trust building process