



JOINTACTION4HRI PROJECT  
ANR-16-CE33-0017

# Joint Actions, Commitments and Need to Belong

Víctor Fernández Castro

Institut | Nicod



Human-Robot Joint Action, September 19, 2018 - Paris, France

# Introduction

- Among the many devices involved in joint action, **commitments** appear to play a key role.
- In particular, commitments can dramatically facilitate joint actions by stabilizing expectations, reducing the uncertainty of the interaction and improving coordination among agents.
- However, a commitment can only play these roles if it is credible in the first place.
- Yet, the credibility of commitments is not a straightforward matter, especially when one notes that credibility depends upon the motivation of the committed agent to honor her commitment in many situations where alternative options that maximize her interests are available.
- **The credibility problem**: what is our most fundamental motivation for complying with our commitments?

Michael & Pacherie (2015)

# Overview

- Commitments and their role in joint action
- The credibility problem
- Two normative approaches to the credibility problem
- Reputation and credibility
- Social emotions and credibility
- The need to belong as a fundamental motivation

# What are commitments

Commitments are characterized by Gilbert (2007) as "creatures of the will":

- A commitment is **by the will** in the sense that it is an exercise of a subject's will.
- A commitment is **of the will** in the sense that it binds the will in a certain way.
- Commitments have **content**.
- Commitments have **authors** and **recipients**.

# How are commitments established?

- Traditionally, philosophers have connected the establishment of commitments to explicit verbal agreements, e.g. by intentionally communicating that one intends to  $x$  through a promise (Austin, 1975; Gilbert, 2009).
- However, commitments are not necessarily established through explicit verbal agreements. For instance, one might indicate through gestures or facial expressions that one will perform the action (Ledyard, 1995; Sally, 1995; Scalon, 1998) appropriate

# How are commitments established?

- In addition, a sense of commitment can arise under certain conditions, even in the absence of any verbal or non-verbal agreement (Michael et al. 2016; Michael and Salice 2016; Lo Presti 2013)
- **Situational affordances**: we may feel committed to push the open button of the elevator when we see someone trying to get in and the doors are closing.
- More generally, one may feel committed to contribute to another agent's goal simply by identifying this goal and realizing that the contribution of another agent is crucial to their achieving this goal (Michael, Sebanz & Knoblich 2016, Michael and Székely 2018)

# How are commitments established?

- **Repetition**: repeated patterns of interaction can create a sense of commitment (Gilbert 2006, Michael et al. 2015).
- **Effort**: the perception that one's partner is investing effort may generate a sense of commitment
  - For instance, Székely & Michael (2018) found that the perception of a partner's investment of effort in a joint activity led participants to remain engaged longer despite increasing boredom.

# Joint action, coordination and predictability

- Successful joint action depends on the efficient coordination of participant agents' goals, intentions, plans, and actions.
- Efficient coordination requires mutual predictability. (e.g. through the alignment of representations)
- Various forms of uncertainty can undermine predictability, the alignment of representations and hence coordination.

# Uncertainties

These various forms of uncertainty fall under three broad categories (Michael & Pacherie 2014).

- Motivational uncertainty:
  - How convergent are our respective interests?
  - What goals do we share and can we promote together?
- Instrumental uncertainty
  - What should we do to achieve our joint goal?
  - How are roles to be distributed?
  - When and where should we act?
- Common ground uncertainty:
  - How much of what is relevant to our deciding on a joint goal, planning for that goal and executing our plan is common ground or mutually manifest to us?

# Commitments and Predictability

- It is plausible that the main function of commitments in joint action is to reduce uncertainty and make agents more predictable to one another.
- Commitments may enable agents to have more reliable expectations about each others' actions than would otherwise be possible, thus facilitating cooperation and coordination.

# The credibility problem

- However, these benefits only accrue if commitments are credible in the first place, that is, if the authors of commitments do more often than not act in accordance with their commitments and if their recipients trust the authors to act as committed.
- Yet, the credibility of commitments is not a straightforward matter, however. The problem is that it may appear irrational to engage in commitments, to the extent that they foreclose options which may be more attractive than the action to which one is committed .
- Thus, if an agent makes a commitment to perform a particular action, and her interests, desires or opportunities for action subsequently change, it is not clear why she should remain motivated to honor the commitment.

# Normative approaches to the credibility problem

- In the philosophical literature on commitments, two broad **normative** approaches to the credibility problem can be discerned:
- One approach appeals to considerations of **practical rationality** (e.g., Bratman)
- The other approach appeals to the **deontic nature of commitments** (e.g. Gilbert).

# Bratman on the credibility of commitments (I)



- On Bratman's theory (Bratman 2014), intentions are construed as commitments to act and, as such, as distinctive elements of human planning agency that go beyond the ordinary desires and beliefs characteristic of simple purposive agency.
- Shared intentions, in turn, are construed as complexes of interlocking and interdependent intentions of individual participants.
- Intentions thus allow us to become planning agents, freeing us from the confines of the present, as well as social agents, freeing us from the confines of our own self and allowing us to coordinate with others.

# Bratman on the credibility of commitments (2)



- In order to accrue these benefits, intentions must be subject to norms of practical rationality (means-end coherence, consistency, agglomeration)
- In particular intentions are rationally required to resist reconsideration and be stable, as their instability would defeat the very purpose of planning agency and shared agency.
- Thus, on Bratman's view the credibility of commitments would be premised on the practical rationality of agents: to the extent that an agent is practically rational, we should expect her intentions to be stable and hence her to act in accordance with her intentions.

# Bratman on the credibility of commitments (3)



- This practical rationality approach to credibility lacks sufficient robustness.
- Practical rationality doesn't require us to never reconsider our intentions, it demands that we reconsider only if solid reasons to do so come to light.
- Since it can be rational for an agent to be open to re-considering her intentions, fluctuations in her interests threaten the credibility of her commitments
- Given that shared intentions are construed as a network of interlocking individual intentions, an agent's personal reasons for reconsidering her intentions may lead to the unraveling of the whole network of commitments.

# Gilbert on the credibility of commitments (I)



- On Gilbert's account of shared intentions as involving joint commitments, commitments are credible for a different reason.
- In forming a joint commitment, the parties to the commitment together impose obligations on each other to act in conformity with the commitment, and concomitant rights to demand of one another that they so act.
- In addition, since a joint commitment can only be rescinded with the consent of all the parties involved, absent such consent, agents remain obligated to act in conformity with the shared intention even if their interests have fluctuated.

# Gilbert on the credibility of commitments (2)



- On this deontic approach, the credibility of commitments is premised on the social normative assumption that people act as their obligations dictate.
- An appeal to the normative force of obligation provides, according to Gilbert, for a more stable framework than simply an appeal to the practical rationality of individual agents.
- However, one might justifiably demand an explanation of what motivates people to act as they are obligated to, and Gilbert's account merely asserts that people are so motivated without providing an explanation of why this is so.
- Absent such an explanation, Gilbert's account of the credibility of commitments must be regarded as incomplete.

# Interim summary (I)

- Normative approaches to the credibility problem for commitments suffer from important limitations.
- The practical rationality approach appears to lack sufficient stability
- The deontic approach that appeals to obligations begs the question unless it also offers an explanation for why people are motivated to act as obligated.
- Both normative approaches fail to offer clear suggestions regarding the specific psychological mechanisms leading humans to be motivated to fulfill their commitments.

# Credibility problem: the brief

- We need an account of the proximal psychological mechanisms that motivate us to commit to others and abide by our commitments.
- We should aim at identifying the most fundamental such mechanism, fundamental in the sense both that:
  - i. it operates early in ontogeny and can be part of an explanation for the sense of commitment that emerges in early childhood
  - ii. it can provide a plausible basis for more sophisticated motivations.

# Reputation (I)

- How can the practical rationality approach to credibility be reinforced?
- **Agents care for their reputation** and that their concern for their reputation may counterbalance their inclination to renege on their commitments.
- Several evolutionary theories suggest that people cooperate to maintain a good reputation in their social environment where this reputation in turn attracts valuable partners , thus positively affecting their future benefit.
- This approach then offers an answer to the credibility problem in terms of reputation management. Even if an agent's personal reasons to engage in a particular joint action have changed, he may still care about his reputation as a cooperator and this concern for his reputation may motivate him to act as committed.

# Reputation (3)

- Although the importance of reputation in the stabilization and replication of cooperative and joint action is undeniable (e.g. Trivers 1971; Alexander 1987; Haley & Fessler 2005; Nowak & Sigmund 2005), it is less clear that it constitutes our primary motivation for honoring our commitments.
- In fact, one may object, children acquire a sense of commitment and at least some understanding of how implicit commitments work before they start caring for their reputation and develop a capacity for reputation management.

# Understanding of commitment

- In a set of studies, Gräfenhain et al. (2009) investigated children's emerging understanding of commitments.
- In their first study, they tested children's reactions when an experimenter abandoned a joint action, depending on whether or not the experimenter had made an explicit commitment to the joint action.
- They found that 3 year-olds, but not 2 year-olds, protested significantly more when a commitment had been violated than when there had been no commitment.
- In their second study with 3- and 4-year-old children, they tested the children's understanding of their own obligation to a committed joint activity.
- They found that when they were enticed away from a joint activity with an adult, children in both age groups acknowledged their leaving significantly more often when they had made a joint commitment to act together than when they had not.

# Reputation management

- In contrast, developmental studies on reputation management suggest that children only start exhibiting actions aimed at promoting their own reputation by the age of five (see Silver and Shawn, 2018; Engelmann & Rapp, 2018 for reviews).
- For example:
  - 5-year-old children are more generous when their behavior is perceived by their partner (Leimgruber, Shaw, Santos & Olson, 2012).
  - 5- to 6-year-olds avoid cheating when they believe they are observed by another person (Piazza, Bering, & Ingram, 2011), even when this person is an imaginary character (e.g. Princess Alice).
  - 5-years-old (but not 3-years-olds) communicate evaluative information to partners (gossip) about a third party's proclivity to cooperate (Engelman et al. 2016) .
- In addition, there is no evidence of explicit reputation management before age 8.
- For example:
  - children around age 8 start to explicitly refer to self representational concerns when explaining behavior in front of an audience (Banerjee et al. 2012).
  - 8-year old, but not 6-year old, children use appropriate verbal self-representational strategies to maximize their chances to be selected as a partner for a game (Aloise-Young 1993).

# Limits of the reputation approach

- Although reputation management may sometimes explain or contribute to explaining why we honor our commitments, it does not seem to constitute a plausible general solution to the credibility problem.
- On the one hand, developmental findings suggest that human understanding of implicit commitments appears before the capacity for reputation management.
- On the other hand, equating our motivation to fulfill commitments with our motivation to maintain a good reputation is uninformative, since it simply replaces the question of why we care about honoring our commitments with the question of why we care about our reputation.
- Thus, it seems we have to look elsewhere for the most basic motivation at work in sustaining commitments.

# The social emotion approach

- Why would an agent feel compelled to act as her obligations dictate, to contribute her part to a joint action or to comply with their commitments?
- A plausible move to complement the deontic account could appeal to social emotions (e.g., guilt, shame, pride) as such a driving force.
- On this approach, our tendency to fulfill our obligations would be a result of our to avoid negative social emotions and seek positive ones.

# What are social emotions?

- Although all emotions may be said to be social in the sense that they can be modulated by social factors, many theorists also distinguish a subset of emotions, usually including shame, embarrassment, guilt, pride, jealousy, admiration and so on, as “social emotions” in a more specific sense.
- According to Hareli & Parkinson (2008), social emotions are social in a different way to other emotions, in that:
  - they necessarily depend on other people’s thoughts, feelings or actions, as experienced, recalled, anticipated or imagined at first hand, or instantiated in more generalized consideration of social norms or conventions (131)

# Social emotions and the credibility of commitments

- How can social emotions help solve the credibility problem?
- The general idea is that the anticipation of emotional outcomes of actions changes the payoff structure for an agent's action options.
  - For example, if one agent could get away with renegeing on her commitments, she may nevertheless refrain from doing so because she wants to avoid the negative emotional outcome that she expects to ensue from renegeing on her commitments (e.g., guilt, shame).
- In the context of development, this would imply that children's understanding of commitments should depend upon the development of their ability to anticipate social emotions.

# Social emotions: developmental trajectory (1)

- The predominant view in developmental psychology is that children begin to exhibit pride and embarrassment around their second birthdays, showing public elation when performing well at difficult tasks, and blushing and hiding their faces when they do not do well at some task or other.
- It is also around the same time that they first pass the mirror test (Bischof-Köhler 1991).
- As Rochat (2008) notes:

Placed in front of a mirror with a mark on the face, the child often will not simply self-refer and remove the mark, but also show embarrassment, even blushing... In such secondary or self-conscious emotions, children demonstrate unambiguously that what they hold as representation of themselves (i.e., self-knowledge) factors the view of others' (249).

# Social emotions: developmental trajectory (2)

- Vaish et al. (2016) have recently suggested that guilt is an early form of social emotion that emerges by three years of age as a way of repairing social bonds when harm is inflicted on others.
- In their studies, the experimenters tested the children in four conditions, varying whether or not a mishap caused harm to someone and whether children themselves caused that mishap or not.
- They found that 3-year-olds exhibited more reparative behavior when they were the causal agent of the harm, suggesting that guilt as a social emotion motivating pro-social behavior emerges around age 3.
- Thus social emotions appear to start motivating behavior around age 3.

# Social emotions: developmental trajectory

## (3)

- In contrast to this evidence,, the social emotions we need to appeal to in order to explain adherence to commitments typically depend on a capacity to anticipate or imagine what the thoughts, feelings and actions of our partners would be, were we to renege on them.
- However, developmental evidence suggests that the ability to imagine situations in which social emotions might be experienced does not appear until 7 or 8 years of age
  - Children under this age rarely refer to such complex emotions in their speech (Ridgeway et al. 1985).
  - When presented with vignettes where an agent either succeeds or fails at some action with a moral significance according to their effort, their luck, or outside intervention, children younger than 7 or 8 are not proficient at inferring the resultant moral emotions, such as shame, guilt, pride and anger (Thompson and Paris 1981; Weiner et al. 1982; Thompson 1987a,b).
  - When asked to describe situations likely to provoke a given social emotion, children aged 7 or more, but not children aged 5, are able to do so (Harris et al., 1987).

# Limits of the social emotions approach

- Like reputation concerns, the anticipation of social emotions may sometimes explain or contribute to explaining why we honor our commitments.
- Yet, their developmental timeline suggest that they cannot constitute the most basic motivational force at work in sustaining commitments.
- Once, around age 3, children experience social emotions, they begin to protest when others fail to act in accordance to their commitments and to initiate reparative behavior when they breach their own commitments.
- But it is only later, starting around age 7, that they become able to reliably anticipate social emotions and to use these anticipated emotions to regulate commitment related behavior.

# Interim summary (2)

- The basic rationale behind the reputation view and the social emotion view is that the key to solving the credibility problem is to offer an account of the fundamental motivation that drives us to act as committed, as without such a motivational story, normative approaches to commitments remain largely incomplete.
- However, the reputation view and the social emotions view are both problematic as accounts of the fundamental motivation behind commitment compliance.
- While reputation management and social emotions certainly play a role in motivating us to act as committed, the general timetable for their development is inconsistent with the idea that they can be our primary motivations for complying with our commitments.
- In addition, both are sophisticated motivational states, that seem to depend on more basic motivation: Why we care for our reputation in the first place? Why should we care for how others feel about us?

# The social motivation hypothesis

- In a series of recent papers, Godman, Nagatsu and Salmela (Godman 2013; Godman et al. 2014; Salmela & Nagatsu 2016) have proposed what they call the social motivation hypothesis, according to which:
  - "There is a particular psychological disposition whose role is to orient us toward affiliative stimuli, which yields social reward (affect) and enables the formation of social bonds." (Godman et al., 2014: 13)
- In particular, they argue that many joint actions are driven, at least in part, by agents' social motivations rather than merely by their shared intentions (i.e, agents find acting together with others rewarding in its own right).
- They also argue, more generally, that the social motivation hypothesis represents a basic explanation of the appeal of pro-social behavior (in terms of anticipated social rewards), that provides a plausible scaffold for other more sophisticated motivations.

# The need to belong

- The belongingness hypothesis can be seen as a more constrained version of the social motivation hypothesis.
- According to Baumeister and Leary (1995):
  - Belongingness should be conceptualized as a **need** and a **fundamental human motivation** rather than a mere desire and such that failure to satisfy it ought to be marked by serious distress and long-term negative consequences.
  - Fulfilling the need to belong involves satisfying two criteria: (i) individuals must have relatively frequent, positively valenced (or at least non-aversive) interactions with at least a few other people, and (ii) individuals need to perceive that there is an interpersonal bond or relationship marked by stability, affective concern, and continuation into the foreseeable future.

# Belongingness vs social motivation hypothesis

- Thus, the belongingness hypothesis, but not the more generic social motivation hypothesis per se, entails that:
  - people should strive to achieve a certain minimum quantity and quality of social bonds but , once this level is surpassed, their motivation should diminish.
  - interactions with a constantly changing sequence of partners will be less satisfactory than repeated interactions with the same persons.
  - people should be willing to devote more energy into preserving and consolidating existing bonds than interacting with strangers and interactions with strangers should be appealing mainly as potential first-steps towards long-term contact.

# Need to belong in development

- One important virtue of the view that the need to belong is the fundamental motivation behind commitment and commitment compliance is it is evidenced very early in ontogeny.
- From early in development, children take pleasure in social interactions and engage in behaviors that serve to prolong positive engagement.
- For example, by eight weeks of age, infants smile in response to their social partners and by 12 weeks of age, they rarely smile outside of positive face-to-face exchanges with others (Rochat et al. 1999).
- As early as 8 weeks-old, infants also engage in proto-conversations with caregivers and other agents (Rocha, et al. 1999; Trevarthen & Aitken 2001).
- 6 month-old infants strongly prefer to interact with people who engage in contingent interactions with them (Hay et al. 1983, 2004; Jacobson 1981).

# Need to belong in development

- Children exhibit a selective preference for imitating, engaging, attending or helping those agents who look warmer and friendlier or prosocial (Hamlin and Wynn, 2011; Hamlin et al, 2007, 2010; Lakin and Chartrand, 2003; Nielsen, 2006; 2009; Over and Carpenter, 2009).
- 6-month-old children prefer to interact with a character they observed helping another than one than one they observed hindering another (Hamlin et al. 2007)
- 18-month-old children differ in their copying skills depending on whether the models demonstrating the actions act socially or are aloof: they focus on copying the outcome of the demonstrated action when the model is aloof, but are as likely to focus on copying actions as outcomes when the model behaves socially (Nielsen 2006) .

# Need to belong in development

- These findings show that children do not only interact with caregivers, which could suggest a preference for those who improve their survival, or people in general, which could suggest a general empathic preference, but with those who seem ready or more apt to maintain interactions with them.
- This evidence regarding selective preference over interaction partners speak, Over (2016) suggests, in favor of the idea that at a very young age children seek to affiliate with people, preferring those who exhibit behaviors and features that make them more appropriate for maintaining systematic and long-term relations.

# Need to belong, reputation and social emotions

- Unlike social emotions and reputation management, the need to belong does not presuppose sophisticated cognitive abilities (self-recognition, self-representation, capacities for mindreading and for anticipating the mental states of others).
- Rather, it involves a basic "set of psychological dispositions and biological mechanisms biasing the individual to preferentially orient to the social world (social orienting), to seek and take pleasure in social interactions (social reward), and to work to foster and maintain social bonds (social maintaining). (Chevallier 2012:231).
- As pointed out by Godman et al. (2014) the experience of a reward attached to social affiliative stimuli can lead to social motivation through a simple process of associative learning:
  - " because rewards are typically attached to certain affiliative stimuli, they become increasingly associated in such a way that they yield (social) motivations in expectation of distinctively social rewards. (12-13)"

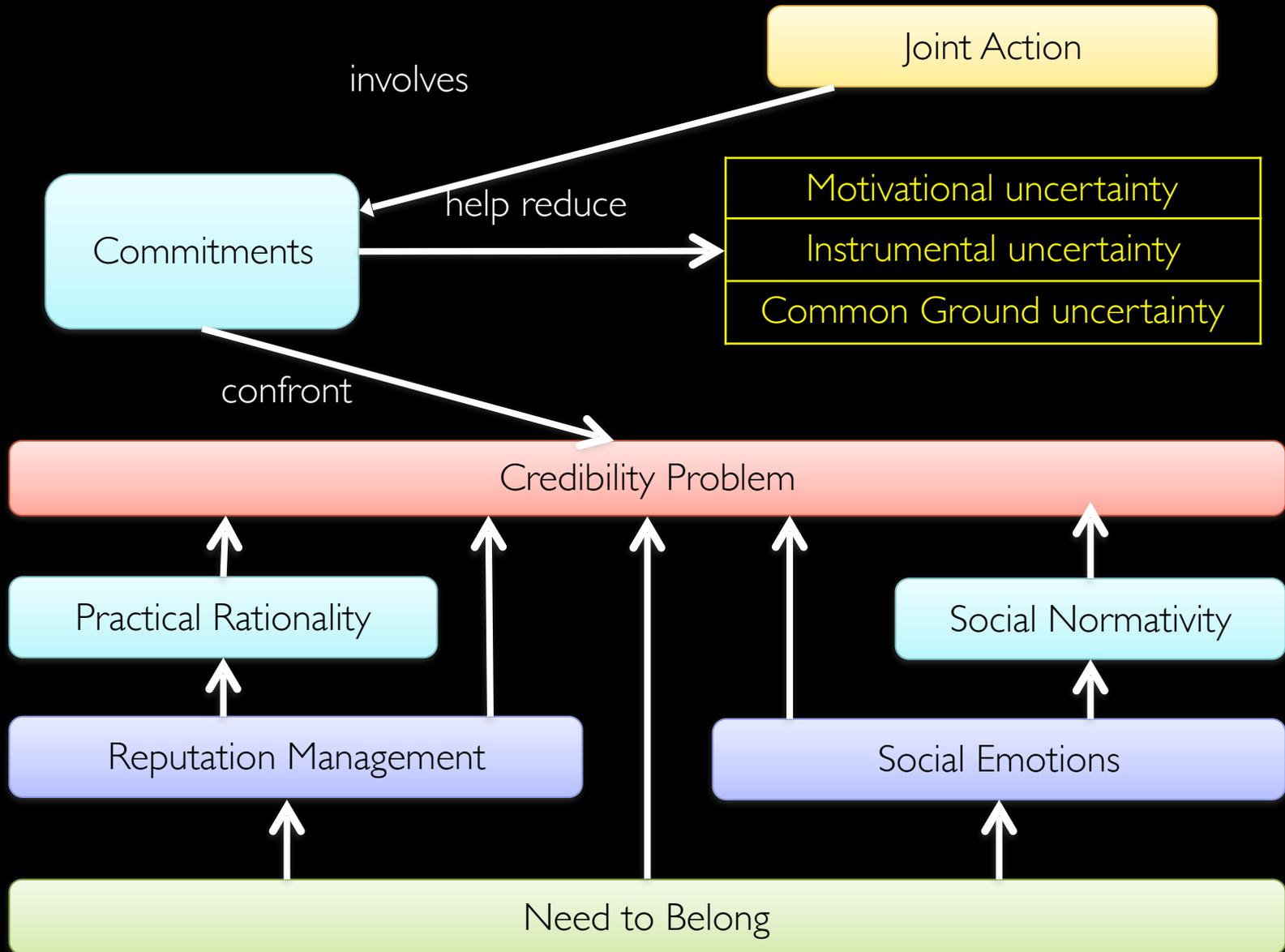
# Need to belong, reputation and social emotions

- The need to belong can also provide an answer to the questions left pending by the reputation and the social emotions accounts:
  - Why do we care about others' assessments of us and our behavior?
  - Why do we care about our reputation?
- For instance, Vaish et al (2018) propose that social emotions such as guilt serve to repair social bonds when harm is inflicted on others, but this presupposes that children already value social bonds and care about their preservation.
- Similarly, positing a need to belong can help explain why we care about our reputation, insofar as the extent to which others are willing to form and maintain social bonds with us depends on our reputation.
- Thus, while they are more complex motivations that depend on more sophisticated cognitive abilities, reputation and social emotions are scaffolded at least in part by the need to belong and the behavior they motivate contributes to the satisfaction of this need.

# Need to belong and sense of commitment

- The need to belong is conceptualized as the need to form and maintain long-term, strong, stable interpersonal relationships, and not simply a drive for social contact.
- If the need to belong is the fundamental motivation behind commitments compliance, then from these characteristics, one can predict that people would be more motivated to conform to commitments in contexts where the others are participants in long-standing relations or in the presence of cues about the aptness of the other as a social partner
- This prediction receives confirmation from the everyday observation, as well as experimental evidence, that repetition, as well as perception of effort on the part of a partner can give rise to a stronger sense of commitment .

# Conclusion: layers of social motivation



Thanks!