Understanding the psychological mechanisms that constrain the transfer of dialogue effects

Katherine O’Lone

Abstract: An intrinsic component of evaluating the success of a dialogue programme is the concept of transfer. This refers to the means by which the effects of dialogue extend beyond the immediate participants to their wider social groups and go on to influence broader societal and policy change. In this paper I am interested in exploring, from a psychological perspective, why positive dialogue effects at an individual level (i.e., the micro level) sometimes fail to motivate future positive behaviour in the local social milieu (i.e., the meso level). This failure of dialogue effects to permeate beyond the immediate group can mean that transfer is limited, thus raising questions about the effectiveness of dialogue interventions. Drawing on psychological research that looks at mechanisms of indirect contact (whereby positive effects of contact with an out-group member spread beyond the immediate setting) I propose that a factor which might hinder the process of transfer is a cognitive bias called vicarious moral licensing. I propose that by understanding the psychological mechanisms that stymie the transfer of individual attitudinal change in to group behavioural outcomes, one can better address the crucial question of how to maximise the degree of transfer following from dialogue.

Keywords: Transfer, Intergroup dialogue, Indirect contact effects, Vicarious moral licensing, Dialogue, Psychological bias

Introduction

Transfer includes the processes by which a project’s effect on participants is spread or transmitted beyond that group to influence other groups, practices or policies, and make broader changes in society (United States Institute for Peace 2016, 11).

Katherine O’Lone holds a BA (1st Class Hons) in Linguistics from University College, London, an MA (Distinction) in Cognition and Culture from Queen’s University, Belfast and an MSc and PhD in Applied Social Psychology, both from Royal Holloway, University London. Her PhD research, supervised by Professor Ryan McKay, focused on the cognitive science of religion and morality, in particular moral transference. This refers to episodes where past (im)moral behaviour is displaced into the present, where past in-group behaviour is displaced onto us, and where moral imperatives are displaced onto other (e.g., supernatural) agents. Her main research interests are: moral cognition and extremism; religion and pro-sociality; and religiosity and morality. Katherine’s most recent work has been published in The Journal for the Cognitive Science of Religion and investigates whether primes of divine forgiveness increase support for state-sanctioned mechanisms of punishment. As part of her work at the Woolf Institute, Cambridge, she works on a project that seeks to develop a reliable set of indicators to measure the effectiveness of interfaith dialogue initiatives.
A wealth of research in social psychology has found that direct contact between members of different outgroups, under certain conditions, is an effective predictor of improved intergroup attitudes (Allport 1954; Brown and Hewstone 2005; Pettigrew and Tropp 2006; Wright 2009). More recently researchers have found that attitudes about out-group members can be improved through mechanisms of indirect contact (see Dovidio et al. 2011; MacInnis and Hodson 2019; Mazziotta, et al. 2011). For example, learning that a close in-group member is friends with an out-group member (Wright et al. 1997); imagining positive interaction with an out-group member (Crisp and Turner 2009) and even observing an in-group member interact with an out-group member (Mazziotta et al. 2011) have all been shown to enhance positive intergroup attitudes.

It is plausible to assume that these mechanisms of indirect contact play a crucial role in transferring the positive effects of dialogue out into the broader community, thus leading to improved intergroup relations at the group level. Yet there is evidence from social psychology and from dialogue practitioners themselves to suggest that all too often the positive effects of dialogue on intergroup attitudes do not permeate beyond the immediate setting (Abu-Nimer 1999), nor do they lead to subsequent behavioural change and policy support in the wider group (Dovidio and Gaertner 2000, 2004; Quillan et al. 2017).

So there is a contradiction. On the one hand there is a large amount of reliable evidence to suggest that both direct and indirect contact with an out-group leads to positive attitudinal changes and improvements in intergroup relations. Yet on the other hand, there is compelling evidence to suggest that contact effects (both direct and indirect) may not have any meaningful impact on people's actual behaviour. Moreover, in certain circumstances they may actually lead to the avoidance of the out-group (Shelton and Richeson, 2005) and a reluctance to engage in collective action. Researchers have recently started to address this discrepancy (MacInnis and Hodson 2018) and in this paper I aim to contribute to this area of investigation by proposing that a psychological bias called vicarious moral licensing (see Kouchaki 2011) might also play a role in the failure to translate positive attitudes gained via contact effects, into subsequent behavioural outcomes at the wider group level (e.g., collective action and policy support).

**Direct intergroup contact effects**

Research in social psychology suggests that intergroup dialogue is a promising method for ameliorating problematic intergroup relations (Dessel 2010; Dessel and Rogge 2008; Dessel et al. 2006; Nagda et al. 2009; Spencer et al. 2008). For example, reported positive outcomes of participation in intergroup dialogue at the individual level range from increased perspective taking and empathy; appreciation
of group differences and structural inequalities; an enhanced sense of similarity with out-group members and a motivation to build bridges and alliances (Nagda 2006; Nagda et al. 2009; Werkmeister-Rozas 2007). At the community level too, the positive outcomes of intergroup dialogue are numerous. They include: decreased stereotyping, increased trust and reduction of threat, the facilitation of relationships, improved communication, establishing common ground, increased perspective taking, generation of grassroots collaboration, and increased commitment to social justice action (DeTurk 2006; Dessel 2010; United Nations Development Program 2004; LeBaron and Carstarphen 1997; Southwest Educational Development Lab 2000; Rodenborg and Huynh 2006; Spencer et al. 2008).

The theoretical rationale for much of the work on intergroup dialogue is based on Allport's landmark 'contact hypothesis' (Allport 1954). The basic premise is that contact between members of different groups is an effective way to reduce prejudice, partly through the reduction in intergroup anxiety (Brown and Hewstone 2005; Pettigrew and Tropp 2008), although for contact to be successful at reducing intergroup conflict Allport stipulated some prerequisite features

1

(Dovidio et al. 2015; Pettigrew 2016). In the decades since it was first published, extensive empirical evidence in support of contact theory has emerged. In 2006 for example the results of meta-analytic studies revealed a small but positive effect of intergroup contact on out-group attitudes (Pettigrew and Tropp 2006), a finding that has since been replicated (e.g., Beelman and Heinemann 2014; Lemmer and Wagner 2015). A larger effect size has been reported for intergroup friendship on improved intergroup attitudes (Davies et al. 2011).

**Indirect intergroup contact effects**

Recent research suggests that these positive effects of intergroup contact are not bound to the immediate dialogue situation (i.e., they are not restricted to direct, face to face interaction with out-group members) and can be transmitted through indirect modes of contact. For example, studies suggest that *extended* contact, that is, merely knowing that an in-group member has a close positive relationship with an out-group member can lead to reduced negative out-group attitudes and intergroup bias (Wright et al. 1997). These findings converge with a body of research that demonstrates people infer their own attributes by observing the behaviour of other group members, as if they had observed themselves performing the action (Galinsky et al. 2008; Goldstein and Cialdini 2007). Even perceiving the actions and emotions of similar others activates the same neural mechanisms we employ when we produce those actions and emotions ourselves (Conson et al. 2017; Decety and Grezes 2006). For example, studies that

---

1 (1) Equal status within the contact (i.e., dialogue) situation. (2) Contact that allows people to get to know outgroup members as individuals. (3) Cooperation in the pursuit of common goals. (4) The support of authority figures (see Pettigrew 2016).
expose participants to shared multi-sensory experiences have been found to alter their mental representations of their identity and to increase the perception of similarity with others (Paladino et al. 2010; Tajadura-Jiménez et al. 2012).

Similar positive effects have been noted through the use of *vicarious* contact, that is, observing the positive interactions of an in-group member with an out-group member influences ideas about how to behave and perceived norms for intergroup behaviour (Gomez et al. 2011; Schiappa et al. 2005; Turner et al. 2008). Even experimentally inducing a sense of merged identity with an out-group member has been shown to lead participants to perceive themselves as possessing the attributes of the target which they then incorporate into their own self-concept (Goldstein and Cialdini 2007). Under certain conditions simply mentally simulating an instance of positive contact with an out-group member, or *imagined contact*, can significantly reduce intergroup bias (Crisp and Turner 2009), as indeed can instances of virtual contact (White et al. 2014), albeit to a lesser degree (Lemmer and Wagner 2015).

The above research provides strong rationale for assuming that any positive effects of dialogue would extend beyond the original setting. It also suggests that dialogue initiatives need not be limited to direct engagement with out-group members, similar positive effects may be found from online participation or in a virtual environment, for example (see Schwab et al. 2017). It also suggests that by virtue of individual members of an in-group engaging in dialogue with an out-group member, any positive effects will transfer to the wider community via mechanisms of indirect contact (i.e., vicarious or extended contact).

**Intergroup contact and a lack of transfer**

Despite the findings outlined above, there is research in social psychology which suggests that, in some circumstances, the positive attitudinal effects of intergroup contact do not permeate beyond the immediate circle of participants. Nor do these effects generalise to other members of the out-group. Liked members of an out-group are perceived as exceptions to a negative rule, rather than stereotypical of the out-group at large (Amir 1976; Cook 1978), for example. In fact, more recent psychological research suggests that increased contact with minority groups can, particularly in situations where there is a degree of competition for resources, increase prejudice and inter-group hostility (Nagel 1995; Saguy 2018). Indeed studies have shown that as minority groups grow, so too does local opposition to race-targeted policies and programmes (e.g., Quillian 1996; Taylor 1998).

---

2 And of course, the opposite is true; that viewing negative behaviour and interactions with out-group members can influence behaviour and reinforce old stereotypes. A reason why continued negative portrayals and narratives about out-groups in the press can heavily influence prejudice (Shaver et al. 2018; Weisbuch et al. 2009).
Dialogue practitioners provide anecdotal evidence to suggest that positive effects of dialogue often disappear after participants return to their normal lives and social groups (Abu-Nimer 1999) often due to fear or resistance from within the community. Abu-Nimer (1999) noted that across six Arab-Jewish dialogue initiatives in Israel, the programmes did manage to alter participants’ perceptions about and interactions with each other but rarely did they affect behaviour outside of the dialogue contexts. He argued that for transfer to occur the attitudinal changes in participants must be translated into behavioural changes that in turn influence others. Only in this way will any successful effects of dialogue filter out and make larger societal impacts. Without this shift from attitudinal to behavioural change, dialogue efforts are but a ‘drop in the ocean’ and any observed effects can easily be reversed once back within the in-group context (Abu-Nimer 1999).

**Divergent effects: intergroup contact effects on attitudes vs. behaviour**

Despite a large amount of evidence to suggest that contact (both direct and indirect) is a reliable predictor of positive intergroup attitudes, there is not a similar degree of evidence examining the role of contact on positive intergroup *behaviour* (McInnis and Hodson 2019). Of the research that has been conducted, it has been shown that often improved intergroup attitudes do not translate into behavioural outcomes. For example, studies have found that despite attitudes towards out-group members becoming more positive over a period of time, this has little impact on in-group members’ actual behaviour (e.g., Dovidio and Gaertner 2000, 2004; Koen and Durrheim 2010; Quillan et al. 2017). Improvements in attitudes do not necessarily correlate with improvements in accompanying behaviours. Nor does the evidence suggest that intergroup contact predicts other relevant behavioural outcomes, such as support for policy (Dixon et al. 2012) or collective action. As MacInnis and Hobson state, ‘it appears that relative to negative intergroup attitudes, negative intergroup behaviours (e.g., avoidance; discrimination) are more persistent’ (MacInnis and Hodson 2018, 13).

Despite people’s willingness to express improved intergroup attitudes (Fasbender and Wang 2017), their subsequent behaviours have not been sufficiently measured. The few studies that have examined behavioural outcomes tentatively suggest some degree of association between increased contact and positive behavioural outcomes. For example, it was found that among youths in Northern Ireland (McKeown et al. 2012) and among children in racially diverse schools (McKeown et al. 2017) intergroup contact was associated with positive behavioural outcomes (i.e., less segregated seating arrangements). Yet, despite the findings that intergroup contact did initially predict more willingness to sit next to out-group members, this effect decayed after time and did not last.
Given these findings, it is critical that we better understand the factors that hinder positive attitudinal change being translated into relevant behavioural outcomes. A clearer understanding of the variables that constrain transfer will allow us to make headway in (a) ensuring that transfer does take place, and (b) that it is successful and sustainable. In the next section I discuss some explanations for the divergent effects of intergroup contact (i.e., why intergroup contact predicts positive attitude change but not necessarily relevant behavioural outcomes).

**Explaining the divergent effects of intergroup effects**

MacInnis and Hodson (2018) propose that issues of external validity and measurement limitations may go some way to explaining the divergent effects of intergroup contact. As they point out, the findings of the controlled laboratory studies that make up the bulk of the psychological literature may simply not be reflective of intergroup contact as it happens in the “real world” (see also Dixon et al. 2007). Similarly, the measures that researchers used to assess contact effects may simply be too crude to capture subtle changes as they occur in real world settings:

> It may not be possible for intergroup contact measures to fully capture the complexities of human relationships. Perhaps it is only the contact experiences tapped by common contact measures or contact that takes place under well-controlled laboratory conditions that reliably predicts reduced prejudice. (MacInnis and Hodson 2018, 15)

This over-application of laboratory findings to real world settings has characterised many previous attempts to assess intergroup initiatives. For example, in the field of prejudice reduction, research conducted in the laboratory has far out-paced research conducted in the field (Paluck and Green 2009). This may mean that those seeking to conduct programmes aimed at societal or policy change are relying on evidence that is not applicable to real world settings (Scottish Centre for Crime and Justice Research 2015). Arguably until this gap is closed, the conclusions that one can draw from the academic literature are, perhaps, limited (Paluck and Green 2009).

This explanation is very plausible, and certainly, as a culture of dialogue evaluation emerges one must face uncomfortable questions about the extent to which one can extrapolate findings from a laboratory to the real world. As compelling as this explanation is, however, there is a long tradition in psychology that has examined the attitude-behaviour gap, that is, that attitudes do not reliably and consistently predict behaviours (Bhattacherjee and Sanford 2006; Venkatesh et al. 2003). This implies that the divergent effects of intergroup contact may have a cognitive explanation. In the next section I propose such an explanation and suggest a mechanism that might underpin these divergent effects; a cognitive bias called *vicarious moral licensing*. 
Moral licensing

Research has found that recalling past moral behaviours renders us more likely to engage in immoral activities without worrying about feeling or appearing immoral (Khan and Dhar 2006; Kouchaki and Jami 2018; Merritt et al. 2010). This phenomenon, **moral licensing**, has been documented in a wide range of contexts including recruitment decision-making (Cascio and Plant 2015; Monin and Miller 2001; Effron, et al. 2009); racial attitudes (Effron et al. 2009; Effron et al. 2012; Mann and Kawakami 2012); consumer choice (Khan and Dhar 2006; Mazar and Zhong 2010); environmental judgements and behaviours (Meijers et al. 2015; Nilsson et al. 2016; Noblet and McCoy 2017; Tiefenbeck et al. 2013); food consumption (Eskine 2013; Garvey and Bolton 2017) and pro-social behaviour (Conway and Peetz 2012; Jordan et al. 2011).

For example, male participants who had the opportunity to establish non-sexist credentials by disagreeing with overtly sexist statements were subsequently more likely to indicate that a particular job was better suited for a man than a woman (Monin and Miller 2001). In the racial domain, non-black participants who had been given feedback stating they were progressing on egalitarian goals were subsequently more likely to choose to sit further away from black individuals and closer to whites as well as indicate greater implicit racial prejudice (Mann and Kawakami 2012). Similarly, participants who had the opportunity to display non-racist credentials by endorsing Barack Obama’s presidential candidacy (Effron et al. 2009) or by selecting a black candidate for a category-neutral job (Monin and Miller 2001) were more likely to exhibit preferential judgements of white candidates in a later job hiring scenario. In fact, simply indicating they would do a future good deed reduced charitable donations (Khan and Dhar 2006).

Vicarious moral licensing

Research has shown that people not only use the past moral behaviour of similar others to inform their own self-concept but they use these credentials to license subsequent immoral behaviour, a type of **vicarious** moral licensing (Kouchaki 2011; Meijers et al. 2018; Newman and Brucks 2018). For example, Kouchaki (2011) asked white participants to indicate whether they thought an available job position was better suited for a particular race. Those who were told of the moral superiority of their group in a prior task were more likely to indicate that a job position was better suited for a white candidate than an African-American candidate (Kouchaki 2011, Study 1). This suggests the past moral behaviour of the in-group enhances one’s moral self-concept, which then licenses discriminatory attitudes on a subsequent task. Similarly, when given information about a group member’s previous non-discriminatory behaviour (i.e., having chosen a Hispanic applicant in a prior task), participants went on to make
discriminatory ratings against the Hispanic for a job position that was stereotypically associated with whites (Kouchaki 2011).

It might be that the knowledge an in-group member has had a positive interaction with an out-group member bolsters an individual's moral identity. The individual then feels licensed to express prejudiced intergroup attitudes without compromising their self of moral self-worth. The positive attitudinal effects garnered from intergroup contact never translates into broader behavioural outcomes because, via a process of vicarious moral licensing, the initial positive effects are used to justify and license entrenched negative intergroup attitudes.

The moderating role of group identification

The degree to which a person is influenced by the behaviour of other group members depends on the strength of their identification with the group (Goldstein and Cialdini 2007; Tajfel 1982; Tajfel and Turner 1979). Therefore, past in-group behaviour can trigger vicarious moral licensing only if identification with the group is strong enough (Kouchaki 2011; Newman and Brucks 2018). For example, Kouchaki found that participants who were told of their group’s past non-discriminatory decisions expressed more prejudiced attitudes towards a Hispanic applicant on a subsequent job hiring task when they identified more with the in-group (Kouchaki 2011, Study 3).

Similarly, participants who strongly identified with a corporate brand (i.e., Nike) used information about that brand’s socially responsible behaviour to vicariously license subsequent selfish behaviour in the context of an economic game. Participants indicating lower levels of ‘self-brand overlap’ however, did not exhibit licensing behaviours (Newman and Brucks 2017, Study 1). Meijers et al. (2018) conducted a series of studies to assess the vicarious moral licensing effect in dyadic relationships. They found that participants who both read about and imagined a close other (e.g., friend or partner) performing an environmentally friendly action subsequently went on to behave in a less environmentally friendly way (compared to participants who read about or imagined their close other performing a neutral action) (Meijers et al., 2018). In line with both Kouchaki (2011) and Newman and Brucks (2018), Meijers et al. (2018) found that the degree of identification with the target moderated any vicarious licensing effects. Less environmentally friendly behaviours were only performed after participants read or thought about a close other’s ethical behaviour, rather than a non-close other. Meijers et al. (2018) also performed a meta-analysis on two of their studies, which revealed a significant vicarious moral licensing effect and yielded a small to moderate mean effect size.
Vicarious moral licensing and intergroup contact

Unlike research on contact theory, research on moral licensing and vicarious moral licensing is relatively new and therefore lacks multiple attempts to objectively measure effect sizes. To address this, there has been a recent focus on conducting meta-analyses to provide an objective measure of the strength of the moral licensing effect\(^3\) (Blanken et al. 2015) and the *vicarious* moral licensing effect (Meijers et al. 2018). These preliminary meta-analyses provide compelling evidence for the robustness of these effects and therefore a degree of confidence in this area of research. Given these promising findings, this fledgling research agenda could provide useful insights to help reduce the lack of transfer of dialogue effects that are not explained by contact theory alone.

For example, the degree of group identification is a factor that warrants more investigation in the context of the transfer of dialogue effects. Multiple studies have shown that the strength of one’s identification with the group moderates vicarious moral licensing effects (Kouchaki 2011; Newman and Brucks 2018). That is, the more someone identifies with the group, the more they will use the past moral behaviours of in-group members to license prejudiced attitudes and judgements. This finding has immediate implications for the transfer of positive effects of dialogue. The prediction being that an individual who observes an in-group member having a positive interaction with an out-group member will be more likely to use this information to license subsequent problematic attitudes the more they identify with the group.

What this means for transfer is that regardless of promising intergroup attitudinal changes within the dialogue setting, these will not permeate into the wider community if group identification is too strong. Research in social psychology has long recognised the role of group identification in fostering problematic intergroup relations (Branscombe et al. 2002; Doosje et al. 1998; Tajfel 1982; Roccas et al. 2006). For example, a study with Israeli participants looking at the effect of collective guilt on the propensity to engage in collaborative behaviour (i.e., with Palestinians) found that the more participants felt collective guilt, the greater their readiness to cooperate. However, the stronger their identification with the target in-group the less their readiness to collaborate with the out-group (Solomatina and Austers 2014). As Kelly points out,

individuals who identify strongly with a social group (and whose self-esteem is therefore highly dependent on it) will be more likely than weak identifiers

---

\(3\) Blanken et al. (2015) for example conducted a meta-analysis of 91 studies ($N = 7,397$) comparing a moral licensing condition with a control condition. The results yielded an average effect size of $d = .31$, 95% CI [0.23, 0.38]; a small to medium effect on Cohen’s (1988) classification.
to differentiate between fellow in-group members and members of alternative social groups. By this means, the social world becomes clearly divided into “us” and “them” and the more favourable perceptions of the in-group promote a positive social identity and self-esteem. (Kelly 1993, 60)

For my purpose it means that, theoretically at least, transfer will not happen if the degree of group identification in the wider social milieu passes a certain threshold. The implication following from this is that research at the wider group level needs to be done to address the ways that people construct their group identities and the types and degrees of group identification this fosters (see Sassenberg and Wieber 2005). It might be that group-level initiatives aimed at addressing degree of group identification would go some way to countering vicarious moral licensing effects, thus meaning that positive intergroup effects from indirect contact stand a better chance of transfer.

**Conclusion**

I have taken a theoretical stance in this paper, and in some places the ideas I have laid out have been speculative. However, what I have hoped to convey is the importance of future research into vicarious moral licensing effects and how they operate at a group level, particularly the moderating role of group identification. I am calling for the construction of a theoretical hybrid as a framework for understanding the issues of the transfer of dialogue effects; one that combines insights from the well-established contact theory and the more fledgling research on moral licensing. However, in the early conceptual stages, this theoretical hybrid is not intended as a framework for understanding the processes of transfer at multiple levels. I am suggesting this will be useful for gaining insights into the mechanisms of transfer from the *micro* (individual level) to the *meso* (local group) level. In order to understand meso to macro level transfer it is likely that a different theoretical hybrid will need to be constructed; the social and psychological and social mechanisms involved in transfer at this superordinate level are no doubt different and beyond the scope of this paper.

In short, just as important as understanding the positive effects of dialogue on individuals, is understanding the biases in the minds of the communities to which they return. In doing so, one can perhaps go some way to maximising the effectiveness of dialogue by ensuring a higher degree of transfer. Ultimately this would mean that the hard-earned individual positive effects garnered from dialogue would go on to influence the broader community, thus greatly increasing the possibility of social cohesion and peaceful relations between previous hostile communities.
Bibliography


other’s behavior’, Brain Research, 1079 (1): 4-14.
and their Implications for Agency and Alliance Building’, Communication
A reality check for the contact hypothesis’, American Psychologist, 60 (7): 697-711.
negative evaluations the problem and is getting us to like one another more the
association: When one’s group has a negative history’, Journal of Personality and
intergroup bias through intergroup contact: Twenty years of progress and future
through direct, extended and other forms of indirect contact’, Group Processes and
Psychology, 36, 1-52.
The licensing effect of immoral counterfactual behaviors’, Journal of Personality
and Social Psychology, 103, 916-932.
Eskine, K. J. (2013) ‘Wholesome foods and wholesome morals?: Organic foods
reduce prosocial behavior and harshen moral judgments’, Social Psychological and
Personality Science, 4, 251–254.
and hiring decisions: Testing the moderating role of decision makers’ core self-


Quillian, L., Pager, D., Hexel, O., and Midtbøen, A. (2017) ‘Meta-analysis of field experiments shows no change in racial discrimination in hiring over
Understanding the psychological mechanisms that constrain the transfer of dialogue effects


