CHAPTER

23

Measures of Intergroup Contact

Simon Lolliot[1], *Benjamin Fell*[1], *Katharina Schmid*[1], *Ralf Wölfer*[1], *Hermann Swart*[2], *Alberto Voci*[3], *Oliver Christ*[4], *Rachel New*[1] *and Miles Hewstone*[1]

[1]University of Oxford, Oxford, UK; [2]Stellenbosch University, Stellenbosch, South Africa; [3]University of Padova, Padova, Italy; [4]Philipps-University Marburg, Germany

Intergroup contact has long been heralded as a means by which intergroup relations can be improved (Baker, 1934; Williams, 1947). The ideas put forward by these early theorists were integrated by Allport (1954), who proposed that under the 'optimal conditions' of equal status, collaborative work towards a common goal, and support from relevant authorities, intergroup contact will result in improved intergroup relations. This became known as the 'contact hypothesis' and, more recently, as 'contact theory' (e.g., Hewstone, 2009). Three impressive meta-analyses have confirmed (a) the basic contention that positive intergroup contact is negatively associated with intergroup prejudice (Pettigrew & Tropp, 2006); (b) that this relationship is stronger the more intimate the contact becomes, with cross-group friendships being more strongly associated with lower prejudice than is general contact (Davies, Tropp, Aron, Pettigrew, & Wright, 2011; Pettigrew & Tropp, 2006); and (c) that the processes of anxiety reduction and empathy enhancement, especially, help translate the beneficial effects of contact into less prejudiced attitudes (Pettigrew & Tropp, 2008).

Numerous present-day societies are characterized by social segregation (for an authoritative historical overview, see Nightingale, 2012), despite becoming increasingly diverse (e.g., Hooghe, Reeskens, Stolle, & Trappers, 2009). Given that intergroup contact between groups competing for limited resources — such as jobs — exacerbates prejudice (Esses, Jackson, Dovidio, & Hodson, 2005), tensions, often born in segregated areas (see Gallagher, 1995), can boil over into intergroup conflict (e.g., xenophobic attacks in South Africa [Steenkamp, 2009]). As societies become more diverse and competition for scarce resources increases, the role of positive intergroup contact as a social intervention tool remains as, if not more, important today as when it was first introduced almost a century ago (e.g., Bogardus, 1928).

MEASURES REVIEWED HERE

We review the psychometric properties of some of the key scales used to measure the most pertinent constructs to intergroup contact research. To achieve this, the chapter is divided into four over-arching sections covering seven constructs germane to research on intergroup contact. These are as follows:

Measuring the Dimensions of Intergroup Contact

1. General Intergroup Contact Quantity and Contact Quality (Islam & Hewstone, 1993)
2. Cross-Group Friendships (Turner, Hewstone, & Voci, 2007a)
3. Extended Contact via Work Colleagues, Neighbors, Friends, and Family Members (Tausch, Hewstone, Schmid, Hughes, & Cairns, 2011)
4. Positive and Negative Intergroup Contact

Measures of Personality and Social Psychological Constructs.
DOI: http://dx.doi.org/10.1016/B978-0-12-386915-9.00023-1

(i) Negative Experiences Inventory (Stephan, Stephan, Demitrakis, Yamada, & Clason, 2000)
(ii) Valenced Contact (Barlow, Paolini, & Pedersen, 2012; Pettigrew, 2008)

Mediators of the Effects of Intergroup Contact on Outgroup Attitudes

5. Intergroup Anxiety (Stephan & Stephan, 1985)

Moderators of the Relationship Between Intergroup Contact and Outgroup Attitudes

6. Salience of Group Memberships
(i) Measure 1 (Voci & Hewstone, 2003)
(ii) Measure 2 (Harwood, Hewstone, Paolini, & Voci, 2005)

Measuring Attitudes Towards Outgroups

7. Outgroup Attitudes
(i) Feeling Thermometers (Converse, Dotson, Hoag, & McGee, 1980)
(ii) The General Evaluation Scale (Wright, Aron, McLaughlin-Volpe, & Ropp, 1997)

OVERVIEW OF THE MEASURES

We aim to provide the researcher with a 'toolkit': a set of psychometrically sound measures that can be used by both the novice and experienced intergroup contact researcher. We also hope that this chapter will help shape the future of intergroup contact research. For example, while we are aware of the beneficial effects of positive contact and the mechanisms through which it improves outgroup attitudes, we echo the call for more research into the effects of negative contact experiences (e.g., Pettigrew, 2008). Understanding both the positive and negative influences of intergroup contact, and, especially, how they interact, can only help researchers, policy makers, and governments better understand and tackle the opportunities, and sometimes the challenges, faced by individuals living in diverse societies. However, in order to not overstate our claims, it is important to first highlight some of the methodological considerations that need to be kept in mind while reading this chapter.

With the exception of one article (Stephan & Stephan, 1985), no research has been published with the sole purpose of testing the psychometric properties of the measures presented below, either by using Campbell and Fiske's (1959) multitrait–multimethod matrix (MTMM), or confirmatory factor analytic (CFA) techniques (e.g., Marsh & Hocevar, 1988). We have had to rely on a rather 'post-hoc' analysis of the available correlation tables, exploratory factor analyses (EFA), and CFAs (often used as part of structural equation modeling protocols, see Anderson & Gerbing, 1988). We have tried to identify correlation tables that make use of some form of MTMM. With regard to factorial validity, we have sought evidence from reported EFAs and CFAs. Such evidence also helps demonstrate *convergent* validity (if the individual items are loaded significantly by their expected factor; Anderson & Gerbing, 1988), and *divergent* validity (if the items are loaded only by their hypothesized factor). Lastly, we have tried to find this evidence (replicating where possible) from samples drawn from different age groups and different countries to provide confidence in the psychometric soundness of the scale under investigation. Where possible, we have also included standardized regression weights when discussing predictive validity to give an idea of the strength of the relationship between two variables (controlling for the other variables present in the respective models). If no regression weights are reported, it is either because unstandardized regression weights were originally reported, or it was unclear whether standardized or unstandardized weights were being reported.

Our measures assess constructs that we expect to change over time. For example, given an intervention to promote positive intergroup contact, we would expect contact to increase and become more positive, anxiety to decrease, and attitudes to become more positive (or at least less negative). While social psychologists do think of (outgroup) attitudes, especially, as fairly stable dispositions, it is strong attitudes that are more temporally stable over time (Visser & Krosnick, 1998). We also expect attitudes to change depending on situational circumstances. In fact, Wilson, Lindsey, and Schooler (2000) proposed that explicit attitudes change relatively easily, while Dovidio, Kawakami, Johnson, Johnson, and Howard (1997) noted that attitudes vary as a function of whether they are expressed privately or in public. These considerations notwithstanding, it is useful to have information about the psychometric properties of the measures that we use, which we have endeavored to provide.

Substantial empirical research supports Allport's (1954) hypothesis that intergroup contact can, and typically does, reduce prejudice towards outgroup members (e.g., Brown & Hewstone, 2005; Pettigrew & Tropp, 2006). Central to the study of the effects of intergroup contact on outgroup prejudice is the operationalization or conceptualization of intergroup contact. An important distinction must be made between the *opportunity for contact* – the potential for intergroup contact to occur as a function of the degree of diversity within a given context (Wagner, Hewstone, & Machliet, 1989) – and actual *direct intergroup encounters* (Pettigrew & Tropp, 2006). While opportunities for contact are sometimes considered as an approximation of face-to-face intergroup contact (e.g., Prestwich, Kenworthy, Wilson, & Kwan-Tat, 2008), it cannot be assumed that because ingroup and outgroup members are in close proximity to one another that they are interacting with each other (e.g., Khmelkov & Hallinan, 1999). Any investigation into the relationship between intergroup contact and outgroup prejudice needs to include a measure of direct interactions between ingroup and outgroup members. Here we discuss three measures of direct contact that are central to the study of intergroup contact and its effects on outgroup prejudice, namely contact quantity, contact quality, and cross-group friendships.

General Intergroup Contact Quantity and Contact Quality (CQCQ)

(Islam & Hewstone, 1993).

Variable

Contact *quantity* relates to the frequency with which someone has direct intergroup encounters whereas contact *quality* reflects the extent to which face-to-face intergroup encounters are experienced positively or negatively.

Description

Islam and Hewstone (1993) undertook one of the earliest studies to explore the differential impact of self-reported CQCQ on outgroup prejudice (see sample details below). Islam and Hewstone measured contact quantity and contact quality using five items each. While the five contact quantity items ask respondents how much contact they have had with the chosen outgroup in a variety of settings, the five contact quality items assess the overall valence of the contact experience. A number of studies have used adapted three-item and four-item versions of Islam and Hewstone's (1993) contact quantity measure, including student and adult samples in Italy (e.g., Voci & Hewstone, 2003) and student samples in Northern Ireland (e.g., Tausch, Tam., Hewstone, Kenworthy, and Cairns, 2007). Similarly, Islam and Hewstone's (1993) contact quality measure has been adapted and successfully applied as a two-item (e.g., Tausch, Hewstone, Kenworthy, Cairns, & Christ, 2007), three-item (Voci & Hewstone, 2003, Study 1), and four-item (e.g., Voci & Hewstone, 2003, Study 2) measure. Sometimes, adapted versions of both the contact quantity and quality measures have been used (e.g., Blascovich et al., 2001).

Islam and Hewstone originally chose seven-point anchors for their items. Other researchers have, however, successfully used five-point rating scales (e.g., Voci & Hewstone, 2003). The contact quality items have also been rated using a five or seven-point bipolar adjective scale (e.g., Tam, Hewstone, Kenworthy, & Cairns, 2009). Higher scores on the contact quantity measure indicate having had more contact experiences with the chosen outgroup. Higher scores on the contact quality scale indicate having had more pleasant contact experiences with the chosen outgroup.

Sample

The Islam and Hewstone (1993) study utilized a sample of 65 Hindu ($M_{age} = 23.06$ years) and 66 Muslim ($M_{age} = 22.14$ years) students attending a Bangladeshi university. Hindu respondents were asked about their intergroup encounters with Muslims ($M_{Contact\ quantity} = 4.99$, $M_{Contact\ quality} = 4.54$), and Muslim respondents were asked about their intergroup encounters with Hindus ($M_{Contact\ quantity} = 4.14$, $M_{Contact\ quality} = 5.36$).

Reliability

Internal Consistency

Contact quantity formed an internally consistent measure for both of Islam and Hewstone's (1993) sample groups ($\alpha = .90$ and .82 for Hindus and Muslims, respectively; Islam, 1992). Neither Islam and Hewstone (1993) nor Islam (1992) reported Cronbach alpha coefficients for the contact quality measure.

Cronbach alpha coefficients for four-item measures of contact quantity range from .72 (Voci & Hewstone, 2003, Study 1) to .84 (Tausch, Tam et al., 2007, Study 2), while for three-item measures, alpha coefficients range

from .71 (Tausch, Tam et al., 2007, Study 1) to .73 (Tausch, Hewstone et al., 2007). For the contact quality scale, while it appears that four-item adaptations exhibit alpha coefficients ranging from .82 (Voci & Hewstone, 2003, Study 2) to .86 (Tausch, Tam, et al., 2007, Study 2), a two-item short version of this measure ($\alpha = .79$; Tausch, Hewstone et al., 2007) may be adequate in surveys where space for survey items is limited (e.g., representative samples).

In a two-wave longitudinal study, the Binder, Zagefka, and Brown (2009) three-item adaptation of Islam and Hewstone's (1993) contact quality measure exhibited Cronbach alpha coefficients at the first wave ($\alpha = .71$) and six months later at wave two ($\alpha = .73$).

Test–Retest

Binder et al. (2009) reported a six-month test–retest correlation of $r = .52$ for their three-item adaptation of Islam and Hewstone's (1993) contact quality scale.

Validity

Convergent/Concurrent

Numerous studies (that used adapted versions of these measures) have reported a significant positive correlation between contact quantity and quality ranging from $r = .36$ (Tausch, Hewstone et al., 2007; Voci & Hewstone, 2003, Study 1) to $r = .49$ (Tausch, Tam et al., 2007, Study 2). Binder et al. (2009) reported that their three-item measure of contact quality correlated significantly ($r_{\text{time 1}} = .51$, $r_{\text{time 2}} = .47$) with their two-item measure of cross-group friendships (see next section). Islam and Hewstone (1993) reported significant positive correlations between their contact quantity measure and perceived outgroup variability ($r = .55$) as well as between their contact quality measure and outgroup attitudes ($r = .57$; see also Tausch, Tam et al., 2007).

Divergent/Discriminant

Measures of contact quantity can generally be distinguished (via low, and often non-significant correlations) from ingroup identification (e.g., $r = -.14$; Tausch, Hewstone et al., 2007), perceived symbolic threat (e.g., $r = -.09$; Tausch, Tam et al., 2007, Study 1), and intergroup anxiety (e.g., $r = -.09$; Tausch, Tam et al., 2007, Study 2). Similarly, measures of contact quality can be distinguished from ingroup identification (e.g., $r = -.11$; Tausch, Tam et al., 2007, Study 1), and perceived relative group status (e.g., $r = .10$; Tausch, Hewstone et al., 2007).

Construct/Factor Analytic

Islam and Hewstone (1993) subjected their CQCQ items to a principal components analysis with varimax rotation ($N = 131$) and reported that each individual item was loaded significantly by its respective factor (loadings for contact quantity factor $\geq .76$; loadings for contact quality factor $\geq .52$). While there was evidence for some item cross-loadings (largest cross-loading = .56), the general pattern of factor loadings supported the hypothesized factor structure. This pattern of relationships was evidenced for both the Hindu and Muslim respondents.

Criterion/Predictive

Islam and Hewstone (1993) reported that contact quantity was significantly associated with more positive explicit outgroup attitudes ($\beta = .12$), more perceived outgroup variability ($\beta = .46$), and less intergroup anxiety ($\beta = -.23$). Islam and Hewstone's (1993) measure has also been shown to have an impact on physiological and neurological reactions to outgroup members (e.g., Blascovich et al., 2003, Study 3). Islam and Hewstone (1993) found that contact quality predicted lower intergroup anxiety ($\beta = -.52$) and more positive explicit outgroup attitudes ($\beta = .48$). Other studies have found contact quality to be predictive of lower intergroup anxiety ($\beta = -.68$ and $-.45$), and perceived realistic ($\beta = -.28$ and $-.30$) and symbolic ($\beta = -.40$ and $-.31$; e.g., Tausch, Tam et al., 2007, Study 1 and Study 2, respectively) threats, as well as predictive of blatant ($\beta = .20$) and subtle ($\beta = .12$; e.g., Mähönen, Jasinskaja-Lahti, & Liebkind, 2011) outgroup attitudes.

Location

Islam, M.R., & Hewstone, M. (1993). Dimensions of contact as predictors of intergroup anxiety, perceived outgroup variability, and outgroup attitude: An integrative model. *Personality and Social Psychology Bulletin, 19*, 700–710.

Results and Comments

CQCQ are important predictors of reduced outgroup prejudice, yet their effects on outcome measures are distinct; contact quality is often considered a stronger negative predictor of intergroup anxiety and prejudice than is contact quantity (e.g., Islam & Hewstone, 1993; Tausch, Hewstone et al., 2007). Given these differential effects, exploring the outcomes of distinct dimensions of direct contact offers a nuanced view of the relationship between contact and prejudice and greater insights into how prejudice may be reduced via intergroup contact.

GENERAL INTERGROUP CONTACT QUANTITY AND CONTACT QUALITY SCALE

Contact quantity

How much contact do you have with [outgroup]...

1. ...at college
2. ...as neighbors
3. ...as close friends

Scaled from 1 = *none at all* to 7 = *a great deal*

How often have you...

1. ...engaged in informal conversations with outgroup members
2. ...visited the homes of outgroup members

Scaled from 1 = *not at all* to 7 = *very often*

Contact quality

To what extent did you experience the contact with [outgroup] as...

1. ...equal (1 = *definitely not* to 7 = *definitely yes*)
2. ...involuntary or voluntary (1 = *definitely involuntary* to 7 = *definitely voluntary*)
3. ...superficial or intimate (1 = *very superficial* to 7 = *very intimate*)
4. ...pleasant (1 = *not at all* to 7 = *very*)
5. ...competitive or cooperative (1 = *very competitive* to 7 = *very cooperative*)

Cross-Group Friendships (CGF)

(Turner et al., 2007a).

Variable

Cross-group friendships are considered to be a particularly beneficial form of intergroup contact because they are characterized by frequent interactions (contact quantity) that are more likely to be characterized by Allport's (1954) 'optimal' conditions (contact quality; Pettigrew, 1998). A number of cross-sectional studies, spanning a variety of contexts, participants, and outgroups have reported a reliable negative relationship between cross-group friendships and a range of prejudice measures (for reviews see Turner, Hewstone, Voci, Paolini, & Christ, 2007b; Vonofakou et al., 2008). The meta-analytic findings of Pettigrew and Tropp (2006; see also Davies et al., 2011) provide clear support for the effects of cross-group friendships on prejudice reduction. They reported that the 154 tests that included cross-group friendship as a measure of contact showed a significantly stronger negative relationship with prejudice (mean $r = -.25$) than did the remaining 1211 tests that did not use cross-group friendships as a measure of contact (mean $r = -.21$; the difference between the two correlation coefficients was significant, $p < .05$).

Self-reported cross-group friendships are generally measured on an aggregate level in terms of the number of friendships a respondent has from one or more outgroups, the frequency of the interactions with these cross-group friends, and the quality of these interactions. As such, these measures are similar to those of contact quantity and quality developed by Islam and Hewstone (1993); however, they refer to the quantity and quality of contact with outgroup *friends* in particular as opposed to with members of the outgroup in general (e.g., Binder et al., 2009). In this section, we will review one of the more popular scales used to measure cross-group friendships, Turner et al.'s (2007a) cross-group friendship scale.

Description

Turner et al. (2007a) offered a four-item cross-group friendships inventory (Study 3) as well as closely related abbreviated two-item (Studies 1 and 4) and three-item (Study 2) version of the scale. Here, we will focus on the two-item (Study 4) and four-item (Study 3) measures. Turner et al. (2007a) used a five-point rating scale for both

the four-item and two-item measures of CGF where higher scores indicated more and higher quality cross-group friendships. The two-item cross-group friendships scale has been successfully used in both cross sectional (e.g., Swart, Hewstone, Christ, & Voci, 2010) and longitudinal (e.g., Binder et al., 2009; Swart, Hewstone, Christ, & Voci, 2011) research.[1]

Sample

Turner et al. (2007a) measured cross-group friendships with Asian amongst White British high school children (Study 3, $N = 164$, $M_{age} = 13.60$ years), and White British university students (Study 4, $N = 142$, $M_{age} = 19.90$ years). The mean CGF scores for these two studies were reported as $M = 1.82$ (Study 3) and $M = 2.77$ (Study 4).

Reliability

Internal Consistency

The two items used to measure cross-group friendships yielded a Cronbach alpha coefficient of $\alpha = .84$ (Swart et al., 2010) and have been shown to correlate positively with each other ($r = .63$, Turner et al., 2007a, Study 4). The two items have also been shown to correlate positively with each other over a six-month period ($r_{time\ 1} = .45$, $r_{time\ 2} = .40$; Binder et al., 2009). Likewise, the Turner et al. (2007a, Study 3) four-item measure of cross-group friendships yielded an alpha coefficient of $\alpha = .82$.

Test–Retest

Binder et al. (2009) reported a six-month test–retest reliability coefficient of $r = .73$ for the two-item CGF measure. The Swart et al. (2011) three-wave longitudinal study (with six months separating measurement occasions), using the two-item CGF measure, reported strong six-month (time 1 and time 2, $r = .56$; time 2 and time 3, $r = .65$) and one year (time 1 and time 3, $r = .57$) correlations.

Validity

Convergent/Concurrent

For the two-item measure of cross-group friendships, Turner et al. (2007a, Study 4) reported significant positive correlations with self-disclosure ($r = .64$), intergroup empathy ($r = .21$), intergroup attitudes ($r = .22$), and intergroup trust ($r = .43$). Swart et al. (2011) reported positive correlations between their two-item measure of cross-group friendships and affective empathy ($.11 \leq r \leq .28$). Turner et al. (2007a, Study 3) reported significant positive correlations between their four-item cross-group friendship measure and opportunity for contact ($r = .30$), extended friendships ($r = .62$, see next section), self-disclosure ($r = .53$), and outgroup attitudes ($r = .25$).

Divergent/Discriminant

Binder et al. (2009) reported significant negative correlations between the two-item cross-group friendship measure and intergroup anxiety ($-.23 \leq r \leq -.25$) as did Swart et al. (2011; $-.18 \leq$ r $\leq -.23$). Furthermore, Binder et al. (2009) reported significant negative correlations between cross-group friendships and social distance ($-.35 \leq r \leq -.39$) and negative emotions ($-.28 \leq r \leq -.33$). Turner et al. (2007a, study 3) reported significant negative correlations between their four-item measure of cross-group friendships and intergroup anxiety ($r = -.29$).

Construct/Factor Analytic

Turner et al. (2007a), using similar items to those in Study 4, subjected their two-item CGF measure together with their measure of outgroup attitudes (Study 1, $N = 60$) to an exploratory principal axis factor analysis with oblimin rotation. Two factors were extracted, with the CGF and intergroup attitude items being loaded significantly by their respective factors (all factor loadings for the CGF items were $\geq .66$; see Swart et al., 2010; Swart et al. 2011, for similar findings). Binder et al. (2009) entered the two-item CGF measure into a principal components analysis with social distance, intergroup anxiety, quality of contact, positive intergroup emotions, negative intergroup emotions, and typicality (see section on contact moderators, below). They found that all items loaded strongly onto their respective factors with minimal cross-loadings. Thus, while Binder and colleagues warned against possible conceptual overlap between their social distance scale and cross-group friendships (they argued

[1]Swart et al. (2010, 2011) used the same anchors as Turner et al. (2007a); Binder et al. (2009), while using cross-group friendship items analogous to the three aforementioned studies, used slightly different anchors.

that intergroup contact is an actual behavior whereas the social distance scale measures an intention towards contact), their results indicated that these two constructs are, in fact, distinct.

Lastly, in a CFA, Swart et al. (2011) found that the two CGF items loaded strongly and significantly onto their designated latent variable at each time point (all factor loadings were equal to or larger than .72). Swart et al. (2010, Studies 1 and 2) found that the two-item CGF measure demonstrated metric invariance between two population groups (mixed-race and White South African high school samples).

Criterion/Predictive

Turner et al. (2007a, Study 3) reported that respondents who had greater opportunities for engaging in contact with outgroup members also reported having more cross-group friendships (e.g., $\beta = .30$). Cross-group friendships have been shown to predict reduced intergroup anxiety (Swart et al., 2010; Swart et al., 2011), greater affective empathy (e.g., Swart et al., 2010; Swart et al., 2011), greater self-disclosure (e.g., . $\beta = .34$ and .63, Turner et al., 2007a, Studies 3 and 4, respectively), more favorable outgroup attitudes (Swart et al., 2010; Swart et al., 2011), less social distance ($\beta = -.05$, Binder et al., 2009) and fewer negative emotions ($\beta = -.10$, Binder et al., 2009).

Location

Turner, R.N., Hewstone, M., & Voci, A. (2007a). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. *Journal of Personality and Social Psychology, 93*, 369–388.

Results and Comments

Cross-group friendships are considered to be the strongest dimension of direct contact, impacting on a wide range of outcomes (Davies et al., 2011; Pettigrew & Tropp, 2006). They embody both contact quantity and quality (Pettigrew, 1998), and, as such, form a useful measure as an alternative to assessing contact quantity and quality separately. Cross-group friendships have been successfully used as a two-item measure, measuring the number of outgroup friends and how much time is spent interacting with these outgroup friends, or as a more comprehensive four-item measure, among younger children, adolescents, and adults, making it an indispensable part of a contact researcher's toolkit.

An obvious limitation of all measures of contact to date is their reliance on self-reports. However, Hewstone, Judd, and Sharp (2011) demonstrated that respondents' ratings of their own intergroup contact were highly correlated with observer ratings. Their research included two items measuring cross-group friendships that were almost identical to those used in Turner et al. (2007a, Study 3), as well as contact quantity and quality items based on those found in Islam and Hewstone (1993). Therefore, the Hewstone et al. (2011) study lends crucial support to the convergent validity of such self-report measures of intergroup contact using the items we have highlighted here (see also Dhont, Van Hiel, De Bolle, & Roets, 2012, for additional evidence).

CROSS-GROUP FRIENDSHIPS SCALE

1. How many close friends do you have at school who are [outgroup]?
2. How many close friends do you have outside school who are [outgroup]?
3. How often do you spend time with [outgroup] friends when you are at school?
4. How often do you spend time with [outgroup] friends outside school?

Notes:

Scale for items 1 and 2: 1 = *none*, 2 = *one*, 3 = *between two and five*, 4 = *between five and ten*, 5 = *more than ten*.

Scale for items 3 and 4: 1 = *never*, 2 = *occasionally*, 3 = *sometimes*, 4 = *quite a lot*, 5 = *all the time*.

The two item measure used in Study 3 included the items:

1. How many [outgroup] friends do you have at University?
2. How often do you spend time with [outgroup] friends when you are at University?

Notes:

Scale for item 1: 1 = *none*, 2 = *one*, 3 = *between two and five*, 4 = *between five and ten*, 5 = *more than ten*.

Scale for item 2: 1 = *never*, 2 = *occasionally*, 3 = *sometimes*, 4 = *quite a lot*, 5 = *all the time*.

Extended Contact via Work Colleagues, Neighbors, Friends, and Family Members

(Tausch et al., 2011).

Variable

While the benefits of direct intergroup contact are well established (Davies et al., 2011; Pettigrew & Tropp, 2006), research also shows that people can benefit from intergroup contact without having it themselves. To this end, Wright and colleagues (1997) introduced the *extended contact* hypothesis: prejudice can be reduced by knowing that an ingroup member has an outgroup friend. In four studies − two correlational and two experimental − Wright and colleagues showed that individuals who reported knowing an ingroup member with an outgroup friend also reported lower prejudice scores. Since the publication of the Wright et al. paradigmatic article, much research has supported this basic tenet (even after controlling for direct contact; for a review, see Turner et al., 2007b). Recent research has begun to test Wright's original ideas in full (Turner, Hewstone, Voci, & Vonofakou, 2008) and expand the theory of extended contact. For example, extended contact is especially effective in segregated areas, where opportunities for direct intergroup contact with the outgroup are low, and among respondents with lower levels of direct contact (Christ et al., 2010).

Although measures of extended and direct cross-group friendships are typically highly correlated, extended contact and direct contact achieve reduced prejudice through different mediating mechanisms. For example, extended contact is thought to reduce prejudice by changing one's perception of the social norms surrounding the outgroup (Turner et al., 2008; Wright et al., 1997) whereas direct contact works especially by reducing intergroup anxiety (see below; Islam & Hewstone, 1993) and by promoting self-disclosure (e.g., Turner & Feddes, 2011) and empathy (Brown & Hewstone, 2005; Pettigrew & Tropp, 2008). Given the increasingly important role that is being afforded to extended contact in intergroup research, we include the construct as part of our toolkit for intergroup contact research.

Measures used to tap a respondent's level of extended contact typically ask respondents to rate the number of ingroup members they know to have friends from a given outgroup (Wright et al., 1997). There are, however, different types of 'ingroup members' to which the question could refer. Traditionally, extended contact effects have been studied by asking if the participant has a close ingroup friend (e.g., Paolini, Hewstone, Cairns, & Voci, 2004) or family member (e.g., Turner et al., 2007a, Studies 2 and 3) who has an outgroup friend. Research often aggregates across the different ingroup exemplars (i.e., friends and family members) to yield an index of extended contact. For example, Turner et al. (2008) formed composite measures of overall extended contact using extended contact via general ingroup members, friends, best friends, and family members. It should be noted, however, that extended contact via different ingroup exemplars has differential effects on outgroup attitudes. Tausch et al. (2011) found that the closer the relationship between the respondents and the ingroup exemplar was (e.g., ingroup neighbor versus family member), the stronger the positive association between extended contact and outgroup trust. Given these results, we will focus on the items used to measure extended contact in the Tausch et al. (2011) study as it represents the most systematic investigation into the differential effects of different types of extended contact and provides the most complete set of items to do so.

Description

The Tausch et al. extended contact measures use four items each to tap extended contact experiences with an outgroup through ingroup friends and family. Extended contact via neighbors and work colleagues was assessed using three items. All items were rated on a six-point rating scale. Higher extended contact scores indicate more extended contact with the outgroup.

Sample

Tausch and colleagues (2011) used an adult sample ($N = 424$, $M_{age} = 45.81$ years, $SD = 15.76$) of Northern Irish respondents to investigate the effects of the different types of extended contact on outgroup trust. The composite means were as follows: extended contact via neighbors ($M = 2.61$, $SD = 1.04$), work colleagues ($M = 2.48$, $SD = 0.90$), ingroup friends ($M = 2.70$, $SD = 1.00$), and family members ($M = 2.73$, $SD = 1.08$).

Reliability

Internal Consistency

Cronbach alpha coefficients were reported for extended contact via neighbors (.73), colleagues (.75), ingroup friends (.84), and family members (.81; Tausch et al., 2011).

Test–Retest

No information on test–retest reliability is currently available.

Validity

Convergent/Concurrent

Tausch et al. (2011) reported correlation coefficients between the different forms of extended contact that ranged from medium (extended contact via work and family $r = .53$) to strong (extended contact via friends and family $r = .72$). Tausch et al. also reported significant positive correlations between the various types of extended contact and direct contact ($.22 \leq r \leq .41$). Furthermore, extended contact via work colleagues ($r = .38$), neighbors ($r = .37$), friends ($r = .49$), and family members ($r = .47$) all correlated significantly and positively with outgroup trust.

Divergent/Discriminant

Closeness of neighbors, work colleagues, friends, and family members (rated on a five-point scale) correlated weakly with extended contact via neighbors ($0 \leq r \leq .05$), work colleagues ($.06 \leq r \leq .14$), friends ($.03 \leq r \leq .11$), and family members ($-.04 \leq r \leq .03$).

Criterion/Predictive

Tausch et al. (2011) reported that extended contact was positively associated with outgroup trust for extended contact via friends ($\beta = .25$) and family members ($\beta = .23$) but not via neighbors or work colleagues (both βs $\leq .07$). Moderated by perceived closeness, extended contact was significantly and more strongly associated with outgroup trust if the respondent perceived their relationship to be close with the work colleague ($\beta = .16$), neighbor ($\beta = .19$), friend ($\beta = .38$), and family member ($\beta = .31$). All beta weights were non-significant for low closeness.

Location

Tausch, N., Hewstone, M., Schmid, K., Hughes, J., & Cairns, E. (2011). Extended contact effects as a function of closeness of relationship with ingroup contacts. *Group Processes & Intergroup Relations*, *14*, 239–254.

Results and Comments

The Tausch et al. (2011) items represent the most comprehensive approach to measuring self-reported extended contact. Covering extended contact experiences from friends, family members, neighbors, and work colleagues, these items provide a reliable means by which a number of hypotheses relating to the extended contact effect can be assessed. The correlations provided by Tausch et al. (2011) help demonstrate both convergent and divergent validity. Furthermore, their multi-item nature permits researchers to take advantage of latent variable modeling techniques allowing for more accurate parameter estimation. Future research, however, needs to subject these items to a stricter analysis of validity with longitudinal analyses needed to confirm their test–retest reliability. Despite these shortcomings, these items represent the best set of questions available for estimating self-reported levels of extended contact and thus form an integral part of the intergroup contact researcher's toolkit.

EXTENDED CONTACT VIA WORK COLLEAGUES, NEIGHBORS, FRIENDS, AND FAMILY MEMBERS

How many of your ingroup [*neighbors* (excluding item 1 below) / *colleagues* (excluding item 2 below) / *friends* / *family members*] have...

1. ...neighbors from [the outgroup]
2. ...work colleagues from [the outgroup]
3. ...close friends from [the outgroup]
4. ...were married to someone from [the outgroup]

Notes:
Scored from: 1 = *none* to 6 = *all*.
Reproduced with permission.

Positive and Negative Intergroup Contact

Variable

Since its inception, research into intergroup contact has experienced a marked positivity bias (see Pettigrew & Tropp, 2006; Pettigrew, 2008), with the majority of work focusing on its potential beneficial outcomes. This is,

perhaps, unsurprising given that the genesis of contact theory lies in Allport's (1954) conditions for *optimal* contact. However, by their very nature, these conditions imply the possibility of suboptimal contact, which one might infer as resulting in either a lack of change in intergroup relations, or, more troublingly, a change for the worse. Studies of naturalistic contact (i.e., contact not specifically intended to improve intergroup relations) suggest that without deliberate implementation of Allport's optimal conditions, the consequences of contact can be positive or negative (e.g., Dijker, 1987; Paolini, Harwood, & Rubin, 2010), while scenarios that explicitly violate the optimal conditions have produced unequivocally negative results (e.g., Pettigrew, 1998).

Increasing group diversity promotes opportunities for both positive *and* negative contact. Although some studies show a net reduction in prejudice in ethnically diverse areas (e.g., Wagner, Christ, Pettigrew, Stellmacher, & Wolf, 2006), there is evidence that group diversity can engender higher levels of intergroup enmity (e.g., Cernat, 2010; Quillian, 1995). The potential prevalence and toxic influence of negative contact has led several authors (e.g., Barlow et al., 2012; Pettigrew, 2008) to call for greater emphasis on this relatively understudied construct. Given the important implications that negative outgroup experiences can have on intergroup relations (e.g., Barlow et al., 2012), we believe that this type of contact and its consequences needs to be moved up the future contact research agenda. To this end, we now evaluate the measures of negative contact that exist and make recommendations as to which items should form part of the toolkit.

Negative Experiences Inventory (NEI)

(Stephan et al., 2000).

Description

To our knowledge, the first attempt to investigate the effects of negative contact was a 17-item measure by Stephan et al. (2000) who posited it to be an antecedent to intergroup threat. Negative contact was measured by asking respondents to report how often they had experienced various examples of negative contact with outgroup members. A 9-point Likert-type rating scale, ranging from *never* to *very frequently*, was used with higher scores representing more negative contact. The approach of Stephan et al. to measuring negative contact based on specific experiences has been adopted in a number of subsequent studies. For example, Stephan et al. (2002) and Aberson and Gaffney (2008) utilized a shorter, 14-item version to assess negative intergroup contact, the psychometric properties of which we discuss below.

Sample

For their full 17-item negative contact scale, Stephan et al. (2000) asked female undergraduates from New Mexico (Study 1: $N = 125$; Study 3: $N = 194$) and Hawaii (Study 2: $N = 161$) about their negative contact experiences with men. The authors reported the means for each study as follows: Study 1, $M = 44.13$; Study 2, $M = 36.00$; Study 3, $M = 38.60$ (no standard deviations were reported).

The shorter 14-item NEI version (Stephan et al., 2002) was first used to assess negative contact experiences between Black and White students from across the US ($N = 1011$, $M_{age} = 18.9$ years). The same 14-item scale was used by Aberson and Gaffney (2008) with a sample of White undergraduates ($N = 386$, $M_{age} = 20.9$ years, $SD = 5.1$).

Reliability

Internal Consistency

The 17-item NEI measure (Stephan et al., 2000) showed high Cronbach alpha coefficients across all three samples: Study 1: $\alpha = .91$; Study 2: $\alpha = .93$; Study 3: $\alpha = .90$. The reduced 14-item version of the measure also exhibited high alpha coefficients (Stephan et al., 2002, White sample: $\alpha = 0.94$; Black sample $\alpha = 0.95$; Aberson & Gaffney, 2008, $\alpha = 0.94$).

Test–Retest

No information is available indicating test–retest reliability to-date.

TABLE 23.1 Correlates of Negative Contact Experience

Variable	Study					
	CW1[1]	CW2[1]	CW3[1]	WGB[1]	WGW[1]	A&G[1]
Explicit Prejudice	.29*	.38*	.26*	.40*	.35*	− .35*[2]
Intergroup Anxiety	.16	.14	.18*	.41*	.34*	.25*
Intergroup Conflict	–	–	–	.11*	.24*	.15*
Status Difference	–	–	–	.17*	.14*	− .02
Realistic Threat	.37*	.28*	.17*	.39*	.34*	.34*
Symbolic Threat	.30*	.31*	.26*	.34*	.43*	.38*
Negative Stereotypes	.19*	.05	.11	.16*	.26*	.24*
Ingroup Identification	–	–	–	.13*	.19*	.07
Positive Contact	–	–	–	–	–	− .05
Implicit Attitudes	–	–	–	–	–	− .01

Notes:
[1]*CW1/2/3 = Stephan et al. (2000), Study 1/2/3; WGB/W = Stephan et al. (2002), Black/White sample; A&G = Aberson & Gaffney (2008).*
[2]*Reverse coded as compared with other studies, *$p < .05$.*

Validity

Convergent/Concurrent

Stephan et al. (2000) found that their 17-item measure of negative contact correlated significantly and positively with outgroup prejudice ($.26 \leq r \leq .38$), and realistic ($.17 \leq r \leq .37$) and symbolic ($.26 \leq r \leq .31$) threat (see Table 23.1 for all inter-item correlations). Stephan et al. (2002) reported weak (but significant) correlations between negative contact and ingroup identity (Black: $r \leq .13$; White: $r \leq .19$). Stephan et al. (2002) reported significant positive correlations between the shorter 14-item negative contact scale and prejudice ($.40 \leq r \leq .48$), intergroup anxiety ($.41 \leq r \leq .47$), and realistic ($.39 \leq r \leq .49$) and symbolic threat ($.34 \leq r \leq .45$). See Aberson and Gaffney (2008) for similar correlation coefficients.

Divergent/Discriminant

Aberson and Gaffney (2008) included both positive and negative contact in the same study and reported weak negative correlations between the two types of contact ($r \leq -.05$). They also reported a significant negative correlation between negative contact and explicit outgroup attitudes ($r = -.35$).

Construct/Factor Analytic

Aberson and Gaffney (2008) entered the 14 negative contact items into an EFA (alongside positive contact items) and found that all the negative contact items were loaded significantly by the negative contact factor (all factor loadings $\geq .40$). Furthermore, they reported low cross-construct correlations (all negative contact items were loaded $\leq .14$ by the positive contact factor; all positive contact items were loaded $\leq .20$ by the negative contact factor). Next, using a partial disaggregation approach (see Bagozzi & Edwards, 1998) to create three item parcels of between four to five items per parcel, they entered the negative contact items into a CFA (alongside a host of other variables). Once again, they found that the three negative contact item parcels were loaded significantly by their specified negative contact latent variable (all factor loadings $\geq .84$).

Criterion/Predictive

Stephan et al. (2000) reported that negative contact was significantly positively associated with outgroup prejudice (Studies 1, 2, and 3), symbolic threat (Studies 1 and 3), and intergroup anxiety (Study 3).[2] Stephan et al. (2002), for their Black and White sample, respectively, reported that the short-form negative contact scale was positively and significantly associated with outgroup prejudice ($\beta = .06$ and $.07$), symbolic threat ($\beta = .23$ and $.23$), realistic threat ($\beta = .23$ and $.32$), and intergroup anxiety ($\beta = .33$ and $.30$).

[2]We have not included regression weights as it is unclear whether the reported values are standardized or unstandardized.

Location

17 item scale:

Stephan, C.W., Stephan, W.G., Demitrakis, K.M., Yamada, A.M., & Clason, D.L. (2000). Women's attitudes toward men: An integrated threat theory approach. *Psychology of Women Quarterly*, *24*, 63–73.

14 item scale:

Stephan, W.G., Boniecki, K.A., Ybarra, O., et al. (2002). The role of threats in the racial attitudes of Blacks and Whites. *Personality and Social Psychology Bulletin*, *28*, 1242–1254.

Results and Comments

The early literature investigating the effects of negative contact treats the construct as a relatively independent entity. More recent studies seeking to redress the positivity bias in the field have tended to present negative contact in direct association with positive contact, envisaging the two variables as opposing but related, or sometimes as opposite ends of a single spectrum of a contact experience. A key question implied by contact valence asks which form of contact has the greater effect on our attitudes, biases and behaviors? This question presents a challenge for the measurement of valenced contact if the inventories are based on specific negative contact experiences; none of the studies discussed thus far have included directly comparable measures of specific positive contact experiences. If researchers wish to present positive and negative contact as similar but opposing factors, their chosen items need to reflect equivalent forms of valenced contact, or else risk compromising content validity.

NEGATIVE EXPERIENCES INVENTORY

Have [outgroupers] ever treated you as follows...

1. ...treated as inferior
2. ...verbally abused
3. ...ridiculed
4. ...manipulated
5. ...rejected
6. ...sexually harassed
7. ...sexually assaulted
8. ...threatened
9. ...physically harmed
10. ...exploited
11. ...forced to do something I didn't want to
12. ...unfairly criticized
13. ...made to feel unwanted
14. ...emotionally blackmailed
15. ...put down
16. ...intimidated
17. ...discriminated against

Notes:
Scored from 1 = *never* to 9 = *very frequently*.[3]
Reproduced with permission.

Valenced Contact (VC)

(Barlow et al., 2012; Pettigrew, 2008).

Variable

A few studies have utilized a 'valenced' contact scale, which includes similarly structured items for both positive and negative contact. We discuss two negative contact inventories that are accompanied by seemingly matched positive contact items, namely Pettigrew's (2008) multi-item positive/negative contact scale and Barlow et al.'s (2012) single item negative/positive contact items.

Description

Pettigrew (2008) reported preliminary results obtained using an eight-item measure of VC experience. The negative contact measure included one negative experience item and three negative feeling items. All four items were answered using a 4-point rating scale with higher numbers indicating more negative contact experiences. The positive contact measure included two positive experience items and two positive feeling items. All four items were answered on a 4-point rating scale. Higher scores represent more positive contact experiences.

[3]The items for the shorter 14-item measure could not be provided as they are not included in the original article.

Barlow et al. (2012, Study 2) also employed VC questions, although their scale only used single items. The wording of their questions was also more generalized than the experience-based items discussed thus far. The items were rated on a 7-point rating scale. Higher scores on the single negative contact item represent more negative contact experiences. Similarly, higher scores on the positive contact item represent more positive contact.

Sample

Pettigrew's (2008) eight items for positive and negative contact were conducted as part of a larger telephone survey taken from a 2004 sample of German citizens ($N = 1383$). Participants were 16 years and older.

The Barlow et al. (2012) items were designed specifically to assess the relative predictive power of positive versus negative contact on a range of dependent variables in an online sample of White Americans ($N = 416$, $M_{age} = 24.6$, $SD = 7.45$). Barlow et al. reported the mean value for positive contact as $M = 4.97$ ($SD = 1.36$), and for negative contact as $M = 3.01$ ($SD = 1.40$).

Reliability

Internal Consistency

Pettigrew's (2008) composite negative contact variable yielded a moderate Cronbach alpha coefficient ($\alpha = 0.78$), as did the positive contact scale ($\alpha = 0.78$).

Test–Retest

No information is currently available indicating test–retest reliability.

Validity

Convergent/Concurrent

Pettigrew's (2008) measure of negative contact correlated positively with anti-Muslim prejudice ($r = .30$). The Barlow et al. (2012) negative contact item correlated positively with modern racism ($r \leq .31$), old-fashioned racism ($r \leq .27$), active outgroup avoidance ($r \leq .25$), and issue avoidance ($r = .29$).[4]

Divergent/Discriminant

Pettigrew's (2008) measures of negative and positive contact correlated negatively and significantly with each other ($r \leq -.18$) as did the Barlow et al. (2012) negative and positive contact items ($r \leq -.26$). Furthermore, Pettigrew's (2008) positive contact scale correlated significantly and negatively with anti-Muslim prejudice ($r = -.41$) and the Barlow et al. positive contact item correlated negatively and significantly with modern racism ($r \leq -.22$), old-fashioned racism ($r \leq -.20$), active outgroup avoidance ($r \leq -.21$), and issue avoidance ($r = -.23$).

Construct/Factor Analytic

No factor analytic information has been provided to-date.

Criterion/Predictive

Individual threat ($\beta = .31$) and collective threat ($\beta = .18$) were both significantly positively associated with Pettigrew's (2008) measure of negative contact. The equivalent coefficients for the positive contact measure were also significant, but negative (individual threat $\beta = -.15$; collective threat $\beta = -.12$). The Barlow et al. (2012) negative contact item was positively associated with modern racism ($\beta = .27$), old-fashioned racism ($\beta = .24$), active outgroup avoidance ($\beta = .20$), and issue avoidance ($\beta = .25$). Their positive contact item, on the other hand, was negatively and significantly associated with modern racism ($\beta = -.15$), old-fashioned racism ($\beta = -.13$), active outgroup avoidance ($\beta = -.16$), and issue avoidance ($\beta = -.16$).

[4]Active outgroup avoidance, in this study, was conceptualized as the extent to which participants overtly avoided intergroup encounters whereas issue avoidance was conceptualized as the extent to which participants avoided intergroup topics that could be considered 'sensitive'.

Location

Multi-item scale:

Pettigrew, T.F. (2008). Future directions for intergroup contact theory and research. *International Journal of Intercultural Relations*, *32*, 187–199.

Single-item scale:

Barlow, F. K., Paolini, S., Pedersen, A., et al. (2012). The contact caveat: Negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin*, *38*, 1629–1643.

Results and Comments

Stephan and colleagues' (2000) negative experience measure has demonstrated good internal consistency and construct validity across several samples. Conceptually speaking, a number of the Stephan items could be distinguished as unique constructs in their own right, for example, measures of discrimination experiences (e.g., Broman, Mavaddat, & Hsu, 2000) show significant content overlap with items such as 'being treated with less respect than others'. It is arguable that 'discrimination' and 'negative contact' represent two labels for the same fundamental experience. To this end, multi-item measures of negatively valenced contact, such as Stephan and Stephan's (2000) inventory, have the potential to reveal possible subordinate dimensions of negative contact (e.g., active versus avoidant discrimination), but until the construct validity of these scales can be established, we would advise selecting the measure whose research pedigree relates most directly to the question at hand. In general, the Stephan et al. scale provides a valid and reliable psychometric test when negative contact is being treated as a stand-alone construct.

Negative contact is, however, not always treated as an isolated variable. Aberson and Gaffney (2008), for example, included measures for both positive and negative contact. Treating positive and negative contact as related but opposing constructs not only necessitates careful content matching of their respective scales (see above), but also invites questions as to the possible 'magnitude asymmetry' between positive and negative items. To illustrate, a negative contact scale assessing frequency of physical violence, for example, represents a more extreme experience than an ostensibly equivalent positive contact measure of feeling 'welcomed by', or 'receiving help from' the outgroup. In Aberson and Gaffney's (2008) study for example, the experience-based negative contact scale was used alongside a measure of contact quality with close outgroup friends. While Aberson and Gaffney do not claim equivalence between their two valenced contact measures, their study serves to highlight this particular issue for the comparison of positive and negative contact.

When comparison is required, we would recommend the use of some kind of valenced contact scale, such as that offered by Pettigrew's (2008). However, the valenced contact scales presented here require some improvement. Pettigrew's scale produces notably lower alpha scores than previous measures and several of its items also show a high degree of conceptual overlap with outgroup emotions. Pettigrew's scale does, however, benefit from experience items that seem reasonably matched in terms of intensity.

As an alternative, the two Barlow et al. (2012) single items of valenced contact avoid measuring specific (and thus potentially non-comparable) experiences. It should be noted, though, that single-item measures forego much of the richness of data that characterizes the experiential measures, and are limited in their use if any kind of latent variable modeling is intended. As with any questionnaire study, researchers should choose the measure of negative contact which most readily suits the hypothesis they wish to test. Finally, in making the above recommendations we would highlight the fact that although the studies discussed here have provided some correlational evidence as for convergent and divergent validity of their constructs, none has presented explicit tests of these criteria. This is clearly an imperative for future research.

VALENCED CONTACT (MULTI-ITEM SCALE)

Negative contact

1. How often has a foreigner pestered you?

Now think about encounters with foreigners in Germany. How often have you experienced the following feelings…

2. …angry
3. …irritated
4. …fearful

Scored from: 1 = *never*, 2 = *sometimes*, 3 = *often*, 4 = *very often*

Positive contact

1. How often has a foreigner helped you?
2. How often do you have interesting conversations with a foreigner?

Now think about encounters with foreigners in Germany. How often have you experienced the following feelings:

3. . . .cheerful
4. . . .satisfied

Scored from: 1 = *never*, 2 = *sometimes*, 3 = *often*, 4 = *very often*

Note: Reproduced with permission.

VALENCED CONTACT (SINGLE-ITEM SCALE)

Negative contact

On average, how frequently do you have NEGATIVE/BAD contact with Black people?

Scored from: 1 = *never* to 7 = *extremely frequently*

Positive contact

On average, how frequently do you have POSITIVE/GOOD contact with Black people?

Scored from: 1 = *never* to 7 = *extremely frequently*

Note: Reproduced with permission.

MEDIATORS OF THE EFFECTS OF INTERGROUP CONTACT ON OUTGROUP ATTITUDES

Intergroup Anxiety Scale (IAS)

(Stephan & Stephan, 1985).

Variable

Stephan and Stephan's (1985) IAS measure is the most popular means by which intergroup anxiety has been assessed. The scale is administered by asking participants to imagine how they would feel if they were the only member of their ingroup interacting with people from (an)other social group(s). Bearing this scenario in mind, participants are then asked to indicate the extent to which they would feel *awkward*, *self-conscious*, *happy* (reverse scored), *certain* (reverse scored), *accepted* (reverse scored), *confident* (reverse scored), *irritated*, *impatient*, *defensive*, *suspicious*, and *careful*. Stephan and Stephan originally used 10-point rating scales but other studies have successfully used seven-point rating scales (e.g., Islam & Hewstone, 1993). Some studies have also adapted the original rating scale format to a bipolar-adjective scale (e.g., Swart et al., 2011; Turner et al., 2008). Typically, higher scores represent higher experienced intergroup anxiety.

While a number of the short versions based on Stephan and Stephan's (1985) original scale exist — some using a subset of three (e.g., Greenland & Brown, 1999), four (e.g., Harwood et al., 2005, Study 2), or five (e.g., Voci & Hewstone, 2003, Study 1) items of the original 11 items — these shorter versions do not consistently use the same items. Because of these discrepancies, we will report on a short-form of the anxiety scale that has been used in two studies (Paolini et al., 2004, Study 1; Turner et al., 2008, Study 1). Both these studies employed a 6-item version of the original anxiety measure using the following adjectives in combination with the introduction blurb: *happy* (reverse scored), *awkward*, *self-conscious*, *confident* (reverse scored), *relaxed* (reverse scored), and *defensive*. While it is by no means an 'official' version, this 6-item version does provide, as we shall argue, an attractive shorter measure of intergroup anxiety for the researcher short on space in their questionnaire.

Lastly, we will also review a second short 6-item version (Swart et al., 2010; Swart et al. 2011) based on Stephan and Stephan's (1985) intergroup anxiety scale. This scale has been used on younger samples. This version differs slightly from the other two versions covered here not only in terms of its use of simpler language, making it better suited for younger samples, but also in its response format favoring a 5-point bipolar adjective response format over the typical rating scale format. The adjective pairs used in this version include *relaxed—nervous*, *pleased—worried*, *not scared—scared*, *at ease—awkward*, *open—defensive*, and *confident—unconfident*.

Description

A major development since Allport's (1954) pioneering work is that researchers have moved from simply demonstrating that contact works, to asking the more demanding question of *how* it works. Although the effects of contact may partly be due to mere exposure (i.e., the principle that familiarity fosters liking; Bornstein, 1989), the published research demonstrates that more sophisticated mechanisms are at work (see narrative review by Brown & Hewstone, 2005; and a meta-analysis specifically of mediators of contact by Pettigrew & Tropp, 2008). Because intergroup anxiety has consistently emerged as one of the strongest mediators of the effects of intergroup contact (Pettigrew & Tropp), we focus on it here.

Intergroup anxiety is defined as the affective response experienced when anticipating future, or expecting actual, contact with an outgroup member (Stephan & Stephan, 1985). It is a negative affective process that is integral to the contact situation (see Greenland & Brown, 1999), and is distinct from chronic levels (or trait measures) of anxiety (Britt, Boniecki, Vescio, Biernat, & Brown, 1996). Anxious responses include physiological responses of the autonomic system (such as an increased heart rate), behavioral responses (reduced performance and contact avoidance), and subjective feelings of anxiety (Mendes, Blascovich, Lickel, & Hunter, 2002); it is the latter feelings that social psychologists typically try to measure with scales tapping intergroup anxiety. Intergroup anxiety has numerous consequences for intergroup relations, notably biasing information processing (Stephan & Stephan, 1985). For example, intergroup anxiety has been shown to be related to greater stereotyping (see Bodenhausen, 1993), more homogeneous perceptions of the outgroup and more negative outgroup attitudes (Islam & Hewstone, 1993).

Intergroup anxiety has proven to be one of the strongest and most consistent predictors of negative attitudes in correlational studies (e.g., Bizman & Yinnon, 2001; Islam & Hewstone, 1993; Stephan et al., 2000). In their meta-analysis of mediators of contact, Pettigrew and Tropp (2008) identified intergroup anxiety as the most significant mediator between contact and attitudes (see also Paolini, Hewstone, Voci, Harwood, & Cairns, 2006). Furthermore, intergroup anxiety can also lead to avoidance of contact altogether (see Plant & Devine, 2003). We will now review Stephan and Stephan's (1985) intergroup anxiety scale.

Samples

Stephan and Stephan (1985) used a sample of 83 Hispanic college students to validate their full 11-item measure of intergroup anxiety.[5] Unfortunately, no participant details (e.g., mean age or mean score on the anxiety measure) were provided. Other studies have, however, used the scale. For example, Islam and Hewstone (1993) administered the full intergroup anxiety measure — using a seven-point rating scale — to Hindu and Muslim students at a university in Bangladesh.[6] The Muslim students (the majority group in Bangladesh) reported lower anxiety scores ($M_{\text{Muslim}} = 3.07$, $SD = 1.05$) than did the (minority) Hindu respondents ($M_{\text{Hindu}} = 4.77$, $SD = 1.16$).

Paolini et al. (2004, Study 1) administered the shorter six-item version to $N = 341$ Catholic ($M_{\text{age}} = 22.22$ years, $SD = 4.49$) and Protestant ($M_{\text{age}} = 24.14$ years, $SD = 6.84$) Northern Irish undergraduates. Turner et al. (2008, Study 1) administered the same six-item version to a group of British undergraduate students ($N = 142$, $M_{\text{age}} = 19.9$ years). Paolini and colleagues' sample, when asked about their feelings of anxiety towards the other religious group (Catholics for Protestants, and *vice versa*), reported low intergroup anxiety scores (Catholic $M = 1.91$, $SD = 1.00$; Protestant $M = 1.63$, $SD = 1.03$). Similarly, Turner et al.'s (2008) undergraduate sample, when asked about their feelings of anxiety towards Asians, reported low mean levels of intergroup anxiety ($M = 1.58$, $SD = .80$).

Swart and colleagues (2010, 2011) administered their six-item version to multiple high school samples in South Africa. In Studies 1 and 2, Swart et al. (2010) asked White (Study 1: $N = 186$, $M_{\text{age}} = 16.83$ years, $SD = 0.71$; Study 2: $N = 171$, $M_{\text{age}} = 16.80$ years, $SD = 0.71$) and mixed-race (Study 1: $N = 196$, $M_{\text{age}} = 16.98$ years, $SD = 0.96$; Study 2: $N = 191$, $M_{\text{age}} = 16.92$ years, $SD = 0.91$) South African respondents about their experienced intergroup anxiety when interacting with someone from the other race (Whites for mixed-race respondents and *vice versa*). Both White (Study 1: $M = 3.38$, $SD = 1.08$; Study 2: $M = 3.24$, $SD = 0.96$) and mixed-race (Study 1: $M = 2.57$, $SD = 1.05$; Study 2: $M = 2.25$, $SD = 0.79$) respondents scored either around or below the scale mid-point. In a three-wave

[5]In their article, Stephan and Stephan (1985) state that they used 10 adjectives when validating their intergroup anxiety scale, but, in fact, offer 11 items. We understand this to be a typographic error, and will thus refer to it as an 11-item scale.

[6]Islam and Hewstone (1993) actually administered 10 of the 11 original items, reporting that the word 'certainty' could not be accurately translated for the respondents in Bangladesh. This study is, however, the only other study to our knowledge that has used close to the full 11-item version offered by Stephan and Stephan's (1985). Bearing this in mind, we will report the results obtained by Islam and Hewstone (1993) in order to provide additional information, some of which was not provided by Stephan and Stephan's (1985) original analysis (e.g., means and standard deviations).

longitudinal study with six months separating measurement occasions, Swart et al. (2011) asked $N = 465$ mixed-race South African high school students (Time 1: $M_{age} = 14.69$ years, $SD = 1.07$) about their intergroup anxiety towards White South Africans. At all three time points, the respondents reported anxiety scores below the mid-point of the scale (Time 1: $M = 2.46$, $SD = 0.82$; Time 2: $M = 2.52$, $SD = 0.88$; Time 3: $M = 2.37$, $SD = 0.88$).

Reliability

Internal Consistency

Stephan and Stephan (1985) reported Cronbach alpha coefficients for the full intergroup anxiety scale ($\alpha = .86$), as did Islam and Hewstone (1993); Hindu sample $\alpha = .86$, Muslim sample $\alpha = .77$). Paolini et al. (2004) and Turner et al. (2008) reported alpha coefficients for the shorter six-item version ($\alpha = .90$ and .86, respectively) as did Swart and colleagues (2010); Study 1: White sample $\alpha = .93$, mixed-race sample $\alpha = .89$; Study 2: White sample $\alpha = .91$, mixed-race sample $\alpha = .79$). Lastly, Swart et al. (2011) reported alpha coefficients for their short version intergroup anxiety scale for younger samples across the three time points spaced six months apart ($\alpha_{time\ 1} = .78$, $\alpha_{time\ 2} = .78$, and $\alpha_{time\ 3} = .80$).

Test–Retest

Evidence for test–retest reliability of Stephan and Stephan's (1985) intergroup anxiety measure comes from the Swart et al. (2011) three-wave longitudinal study using their short six-item version. The six month test–retest correlations were significant (time 1 to time 2: $r = .31$, and time 2 to time 3: $r = .37$) as was the one year test–retest correlation ($r = .38$).

Validity

Convergent/Concurrent

Stephan and Stephan (1985) reported significant positive correlations between their measure of intergroup anxiety and intergroup dissimilarity ($r = .35$), xenophobia ($r = .23$), and stereotyping ($r = .26$). Paolini et al. (2004) reported a significant positive correlation between the short-form intergroup anxiety scale and outgroup prejudice ($r = .68$). Swart et al. (2010, Study 2) reported a significant positive correlation between their short-form intergroup anxiety scale for younger participants and negative action tendencies for their White sample ($r = .30$) – but not for their mixed-race sample ($r = -.03$).

Divergent/Discriminant

Stephan and Stephan (1985) offered evidence for divergent validity by reporting significant correlations between their intergroup anxiety scale and intergroup contact ($r = -.29$) Islam and Hewstone (1993) reported significant negative correlations between intergroup anxiety and perceived outgroup variability ($r = -.45$) and outgroup attitudes ($r = -.45$).

The two studies using the short form of the intergroup anxiety scale found it to correlate significantly and negatively with direct cross-group friendships with the outgroup (Paolini et al., 2004: $r = -.48$; Turner et al., 2008, Study 1: $r = -.39$), extended friendships with the outgroup (Paolini et al., 2004: $r = -.45$; Turner et al., 2008, Study 1: $r = -.50$), outgroup variability (Paolini et al., 2004, $r = -.35$), and attitudes towards the outgroup (Turner et al., 2008, $r = -.48$).

Using the short-form intergroup anxiety scale for younger participants across two studies, Swart et al. (2010) reported significant negative correlations between intergroup anxiety and cross-group friendships ($-.17 \leq r \leq -.44$), outgroup attitudes ($-.15 \leq r \leq -.55$), and perceived outgroup variability ($-.22 \leq r \leq -.39$; with the exception of the correlation between intergroup anxiety and perceived outgroup variability for the mixed-race sample in Study 2, which was non-significant, $r = -.09$). Swart et al. (2010, Study 2) did not find a significant correlation between their short-form intergroup anxiety scale for younger participants and negative action tendencies for their mixed-race sample ($r = -.03$).

Construct/Factor Analytic

Using a large sample of Colored high school children in South Africa, Swart et al. (2011) subjected their short intergroup anxiety measure to an exploratory maximum likelihood factor analysis with direct oblimin rotation at each time point in their three-wave longitudinal study (Time 1, $N = 465$; Time 2, $N = 394$; Time 3, $N = 351$) and found the intergroup anxiety construct to be unidimensional at each time point. Swart et al. (2011), in a series of CFAs, were also able to show that the item parcels were loaded by the intergroup anxiety latent variable strongly,

significantly, and consistently across the three waves (all factor loadings $\geq .66$, all $p < .001$).[7] Additionally, Swart et al. reported that their shorter intergroup anxiety scale displayed metric invariance over time (i.e., all item parcels were loaded equivalently by the latent variable when compared with the same parcel at a later time point). The CFA model presented in Swart et al. (2010), which included cross-group friendships, intergroup anxiety, outgroup attitudes, perceived outgroup variability (Study 1), and action tendencies (Study 2), displayed adequate model fit for both White and mixed-race samples, helped to establish the divergent validity of this short form of the scale. Furthermore, the factor loadings for the intergroup anxiety scale were invariant (i.e., statistically equivalent) between the White and mixed-race samples in both studies.

Criterion/Predictive

Stephan and Stephan (1985) found that higher levels of intergroup contact significantly predicted lower levels of intergroup anxiety. Furthermore, Islam and Hewstone (1993) also found that both contact quantity ($\beta = -.23$) and contact quality ($\beta = -.52$) were negatively associated with intergroup anxiety. In turn, Islam and Hewstone reported that higher levels of intergroup anxiety were significantly negatively associated with perceived outgroup variability ($\beta = -.37$) and outgroup attitudes ($\beta = -.16$). These results attest to the predictive validity of the full version of the intergroup anxiety scale.

Replicating the results with the short version of the intergroup anxiety scale, Paolini et al. (2004) found that direct cross-group friendships ($\beta = -.31$) and extended cross-group friendships ($\beta = -.26$) were negatively associated with intergroup anxiety. Intergroup anxiety was in turn significantly and positively associated with outgroup prejudice ($\beta = .57$) and was negatively associated with perceived outgroup variability ($\beta = -.17$). Turner et al. (2008, Study 1) also found in their multiple mediator model that extended cross-group friendships ($\beta = -.41$), and not direct cross-group friendships ($\beta = -.16$), were significantly associated with less intergroup anxiety. In turn, intergroup anxiety was negatively associated with outgroup attitudes ($\beta = -.19$). These results support the predictive validity of the short version of Stephan and Stephan's (1985) intergroup anxiety scale.

The short version of the intergroup anxiety scale (Swart et al., 2011) displayed good predictive validity, longitudinally. Cross-group friendships at time 1 (and time 2) were associated with less intergroup anxiety at time 2 (and time 3, respectively). Intergroup anxiety at earlier time points, itself, was associated with less perceived outgroup variability at later time points.

Location

Original scale

Stephan, W.G., & Stephan, C.W. (1985). Intergroup anxiety. *Journal of Social Issues*, *41*, 157–175.

Short form

Paolini, S., Hewstone, M., Cairns, E., & Voci, A. (2004). Effects of direct and indirect cross-group friendships on judgments of Catholics and Protestants in Northern Ireland: The mediating role of an anxiety reduction mechanism. *Personality and Social Psychology Bulletin*, *30*, 770–786.

Short form for younger sample

Swart, H., Hewstone, M., Christ, O., & Voci, A. (2010). The impact of cross-group friendships in South Africa: Affective mediators and multigroup comparisons. *Journal of Social Issues*, *66*, 309–333.

Results and Comments

Based on the evidence presented above, Stephan and Stephan's (1985) intergroup anxiety scale provides a good way in which researchers can assess intergroup anxiety. The full 11-item scale covers a wide range of adjectives that could be used to describe one's anxious feelings during an intergroup encounter. Some doubt is cast over the content validity of the measure as not all the items seem suited to estimating one's feelings of anxiety in an intergroup encounter. For example, feelings of being 'irritated' or 'impatient', while they may appear in anxiety provoking situations, seem to be more closely related to feelings of annoyance than anxiousness. Furthermore, neither Stephan and Stephan (1985) nor Islam and Hewstone (1993) performed an EFA on the full version of the scale. We, therefore, remain unsure as to its underlying dimensionality. The high reliability

[7]In both the 2010 and 2011 studies (Swart et al., 2010, 2011), Swart and colleagues used a partial disaggregation approach when setting up their latent variables. In their model, they used two-items per parcel yielding a total of three parcels. Results attesting to a construct's validity using item 'parcels' have been used in previous research (e.g., see Bagozzi & Edwards, 1998).

statistics for the full version of the scale, however, provide some comfort that an overall 'intergroup anxiety' construct is being measured. Paolini et al. (2004) and Turner et al. (2008), having performed a CFA on their short version of the intergroup anxiety scale, provided some assurance concerning the underlying factor structure of their short scale. It is the Swart et al. (2010, 2011) short version for younger participants, however, that provides the most certainty regarding underlying factor structure and behavior. Having performed both EFAs (demonstrating unidimensionality) and CFAs, they also demonstrated metric invariance across samples and time intervals.

As discussed earlier, self-report measures, on which contact researchers rely heavily, are often difficult to validate. How certain can we be that they measure what we intend them to measure? To this end, Greenland, Xenias, and Maio (2012) developed a measure of intergroup anxiety and used an abridged version of Stephan and Stephan's (1985) scale to help demonstrate convergent validity (Study 2). Greenland et al. (2012, Study 3) then demonstrated that their intergroup anxiety scale correlated with physiological responses that are associated with the experience of anxiety. This study thus helps ease concerns as to the validity of self-report measures, especially those of intergroup anxiety.

We believe that all three versions are vital to our toolkit. If the researcher is interested in tapping the broadest definition of intergroup anxiety, then the full item version should be used. If space is limited, however, then Paolini and colleagues' (2004) shorter version represents an appealing choice. Lastly, if the researcher is interested in assessing younger participant's feelings of intergroup anxiety, then the Swart et al. (2010) short version remains a good candidate. Either way, the multi-item nature of each of the scales allows for the construction of latent variables and thus, the estimation of measurement error associated with the construct. From this 'modeling' point of view, some may argue for the use of the shorter versions of the intergroup anxiety scale that use the set of items best suited to tapping the latent variable of intergroup anxiety (e.g., see Hayduk & Pazderka-Robinson, 2007).

INTERGROUP ANXIETY SCALE

All versions of the IAS use some form of introductory passage based on the one provided by Stephan and Stephan (1985):

> 'If you were the only member of your ethnic group and you were interacting with people from a different racial or ethnic group (e.g., talking with them, working on a project with them), how would you feel compared to occasions when you are interacting with people from your own ethnic group?'

1. certain (reverse coded)
2. awkward
3. self-conscious
4. happy (reverse coded)
5. accepted (reverse coded)
6. confident (reverse coded)
7. irritated
8. impatient
9. defensive
10. suspicious
11. careful when interacting with outgroup members

Scored from: 1 = *not at all* to 10 = *extremely*.

Short form (Paolini et al., 2004)

1. happy (reverse coded)
2. awkward
3. self-conscious
4. confident (reverse coded)
5. relaxed (reverse coded)
6. defensive

Scored from: 1 = *not at all* to 4 = *extremely*.

Short form for younger participants (Swart et al., 2010)

1. nervous/relaxed (reverse coded)
2. worried/pleased (reverse coded)
3. scared/not scared (reverse coded)
4. awkward/at ease (reverse coded)
5. defensive/open (reverse coded)
6. unconfident/confident (reverse coded)

Scored from: 1 = *not at all* to 5 = *very much*.

Note. Reproduced with permission.

MODERATORS OF THE INTERGROUP CONTACT TO OUTGROUP ATTITUDES RELATIONSHIP

Two Measures of Group Membership Salience during Contact

Intergroup contact research has made significant advances towards understanding the conditions under which intergroup contact is most effective. Here we focus on one of the key moderators of intergroup contact, group

membership salience. The recognition that group membership salience is a central moderating variable that helps us understand *when* contact will be more or less effective stems from a series of experimental and correlational studies, the latter of which we will focus on in greater detail below (see also Schmid & Hewstone, 2010). The question as to how contact should best be structured in ways that allow for changes in the cognitive representations of self and others to facilitate attitude change has long been disputed. Some authors argue that for intergroup contact to be successful, social categorizations (i.e., group affiliations) need to be deemphasized ('decategorization'; Brewer & Miller, 1984) while others have argued that they should be replaced with more inclusive group boundaries ('recategorization'; Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993). While evidence for both of these approaches exists, a key problem surrounds these two models: the lowering of group membership salience (i.e., de-emphasis of social category based information) may prevent generalization effects from the encountered outgroup individual to other outgroup members that have not as yet been encountered.

Brown and Hewstone (2005) thus argued that contact should be of an 'intergroup' nature, such that group membership salience during contact should be retained to ensure generalization from the encountered outgroup individual to the wider outgroup. They reasoned that keeping group membership low during contact (as advocated in the decategorization and recategorization approaches) is inadvisable since individuals are unlikely to notice group affiliation and may not process stereotype-disconfirming information, which may then impede positive attitude generalization to the outgroup as a whole. Both experimental (e.g., Van Oudenhoven, Groenewoud, & Hewstone, 1996) and correlational evidence now offers consistent support that the relationship between contact and outgroup attitudes is stronger when group membership salience is high (see Brown & Hewstone, 2005). Many studies now show that salience moderates the contact-attitude link for a wide range of attitude measures (including outgroup trust, forgiveness, perceived outgroup homogeneity, and the desire to live in the majority group country; Harwood et al., 2005).

Below we provide an overview of two measures used to capture group membership salience during contact. It should be noted that we include here only measures that were used in studies that explicitly examined salience as a moderator of contact's effects on an outcome measure.

Group Membership Salience during Contact: Measure 1

(Voci & Hewstone, 2003).

Variable

Voci and Hewstone (2003) used different operationalizations of group salience in two cross-sectional studies in the context of intergroup relations between Italians and immigrants. For the purpose of this toolkit, we will focus on the three items used in Study 1 (it is more advisable, from a latent variable modelling perspective, to use these three reliable items than the two items used in Study 2).

Description

Voci and Hewstone (2003) used three items to measure salience (see below for sample details). Responses were recorded on a 5-point Likert-type rating scale. Higher scores represented more perceived group salience during the interaction. When considering the moderation by salience, Voci and Hewstone's (2003) results revealed significant effects, and in the expected direction, between contact and intergroup anxiety (a negative effect), and attitudes towards co-workers (a positive effect). Contact was more strongly associated with less intergroup anxiety when salience was high than when salience was low; similarly, contact was more strongly associated with positive attitudes when salience was high than low.

Sample

Study 1 examined the association of intergroup contact (quantity and quality) with immigrants on intergroup anxiety, perceived variability, outgroup attitudes, and subtle prejudice among $N = 310$ Italian students ($M_{age} = 22.20$ years, $SD = 2.39$). Voci and Hewstone (2003) reported a mean group salience statistic of $M = 2.17$.

Reliability

Internal Consistency

The three items exhibited a Cronbach alpha of .63 (Voci & Hewstone, 2003).

Test–Retest

No evidence for test–retest reliability is currently available.

Validity

Convergent/Concurrent

No correlations were provided and so convergent validity cannot be assessed.

Divergent/Discriminant

No correlations were provided and so divergent validity cannot be assessed.

Construct/Factor Analytic

No exploratory or confirmatory factor analyses have been reported to-date.

Location

Voci, A., & Hewstone, M. (2003). Intergroup contact and prejudice towards immigrants in Italy: The mediational role of anxiety and the moderational role of group salience. *Group Processes and Intergroup Relations*, *6*, 37–54.

SAMPLE ITEMS (MEASURE 1)

When you met people from [the outgroup]...

1. How aware were you that you belonged to different communities?
2. Did you perceive the other person as a typical [outgroup member]?
3. Did you feel that you were two people representing their respective membership groups?

Scored from: 1 = *not at all* to 5 = *very*.

Group Membership Salience during Contact: Measure 2

(Harwood et al., 2005).

Variable

In two studies focusing on intergenerational contact and attitudes, Harwood et al. (2005) used the most comprehensive and reliable set of measures of *group salience* to date.

Description

To assess participants' group membership salience during contact, Harwood et al. (2005) asked participants to rate, on a 7-point Likert-type response scale, their awareness of the age difference between themselves and each grandparent, how much they thought about each grandparent's age, how much their age mattered when talking, and the extent to which each grandparent was 'typical' of other older people (people older than 65 years; Harwood et al., 2005). Higher scores on the salience measures usually indicate higher group salience. Results from Study 1 and 2 revealed that quality of contact with grandparents was significantly associated with more positive attitudes toward the elderly when group salience was high, but not when it was low. Moreover, in Study 2, a number of moderated-mediation effects of contact on attitudes emerged, such that the indirect effects (e.g., via anxiety and perspective taking) were only significant when salience was high.

Samples

Study 1 examined $N = 192$ American students' ($M_{age} = 19.86$ years, $SD = 1.71$) contact (quantity and quality) with grandparents, and attitudes toward older adults (no mean scores were reported). Study 2 examined $N = 100$ British students' ($M_{age} = 19.92$, $SD = 1.59$) the effect of contact with their grandparents (but focusing on the grandparent with whom they interacted most regularly) on a range of outcome variables (e.g., group variability, anxiety, and perspective taking). The same four items used in Study 1 were used in this study, as well, but asked about one grandparent only ($M = 4.01$, $SD = 1.17$).

Reliability

Internal Consistency

In Study 1, the four items were asked of each grandparent, paternal and maternal, and all exhibited Cronbach alpha coefficients: paternal grandfather, $\alpha = .76$; paternal grandmother, $\alpha = .80$; maternal grandfather, $\alpha = .72$; maternal grandmother, $\alpha = .72$. In Study 2, the scale yielded an alpha coefficient ($\alpha = .72$).

Test–Retest

No evidence is currently available for test–retest reliability.

Validity

Convergent/Concurrent

In Study 1, Harwood et al. (2005) reported that their four measures of group salience correlated significantly with each other ($.27 < r < .59$). These correlations are strong enough to demonstrate good convergent validity, but are not too strong to raise concerns that they measure an overall 'salience' construct. Furthermore, correlations between salience measures were stronger within lineage ties (i.e., correlations between salience scores for paternal and maternal grandfather and grandmother were .58 and .59, respectively) than they were between lineage ties (the correlation between paternal grandmother and maternal grandfather, for example, was .29).

Divergent/Discriminant

Group salience is only hypothesized to influence the relationship between contact and attitudes and is not hypothesized to influence attitudes itself. Therefore, the non-significant correlations between group salience and outgroup attitudes ($r = -.18$) and perceived variability ($r = -.10$) reported in Study 2 provide evidence for divergent validity. The correlation between intergroup contact and salience was small, but significant ($r = -.25$, Study 2).

Construct/Factor Analytic

Soliz and Harwood (2006) used the four salience items ($N = 369$ American university students) in a study investigating communicative and relational facets of the grandchild–grandparent relationship. The authors parcelled the four-item salience measure into two parcels and entered them into a CFA. Across two CFAs, the authors found that the salience parcels loaded strongly and significantly onto the designated latent variable (factor loadings ranging between .80 and .83).

Location

Harwood, J., Hewstone, M., Paolini, S., & Voci, A. (2005). Grandparent–grandchild contact and attitudes towards older adults: Moderator and mediator effects. *Personality and Social Psychology Bulletin*, *31*, 393–406.

Results and Comments

We recommend that researchers seeking to examine group membership salience as a moderator of intergroup contact employ the scale used by Harwood et al. (2005). In both their studies, this scale showed high reliability, with strong correlations between the salience measures for each grandparent in Study 1. Moreover, the scale captures the key components of group membership salience during contact, such as awareness and typicality (see Brown & Hewstone, 2005) and thus provides a superior operationalization of salience to single item measures. As with the other variables considered in this chapter, the multi-item nature of Harwood and colleagues' salience items make them an ideal candidate for latent variable modeling and the estimation of associated error – another benefit over single item measures. The fact that the scale is composed of multiple items also holds additional advantages, since it offers the researcher greater flexibility with regard to statistical/analytical procedures. One can thus test moderation using latent modeling procedures, estimating both contact (as independent variable) and salience (as moderator variable) as latent variables, and then also estimate the interaction within a latent variable framework, using a random slope (latent interaction) modeling approach.

GROUP MEMBERSHIP SALIENCE DURING CONTACT SCALE (MEASURE 2)

When communicating with your grandparent:

1. How aware are you of the age difference between you and this grandparent?
2. How much do you think about this grandparent's age?
3. How much does the age difference between you and this grandparent matter?
4. To what extent is this grandparent typical of older people (people over 65)?

Scored from: 1 = *not at all* to 7 = *very*.
Note. Reproduced with permission.

Measuring Attitudes Towards Outgroups

Variable

We now turn our attention to the assessment of the main outcome variable concerning contact theory, attitudes towards outgroups. Psychologists have long been interested in both defining the attitude construct (e.g., Thurstone, 1931) and in developing ways to measure it (e.g., Bogardus, 1925; Likert, 1932). In this section we investigate the psychometric properties of two popular measures typically used to study outgroup attitudes in research on intergroup contact: the feeling thermometer (see Converse et al., 1980) and the General Evaluation Scale (Wright et al., 1997). Before proceeding, however, it is important to clarify what we mean by an *attitude*. Attitudes can have cognitive, affective, or behavioral components. In this chapter, we focus on the affective component of attitudes. For the purposes of this chapter, we will define attitude as a positive or negative (i.e., valenced) evaluation of an entity (in our case, a social group), one that can also vary in extremity (Eagly & Chaiken, 1993). Given this definition, our attitude measures must be able to catch both the valence and extremity of the attitude held towards the social group.

Feeling Thermometer (FT)

(Converse et al., 1980).

Description

The FT has been a mainstay of many studies investigating attitudes towards a variety of social groups. The FT works by asking participants to rate a target group on a thermometer that typically runs from 0 to 100 degrees (though feeling thermometers that use 10-points are not uncommon; e.g., De Tezanos-Pinto, Bratt, & Brown, 2010). Ratings that fall between 51 and 100 'degrees' are indicative of 'warmer' or more favorable attitudes towards the target group, while ratings that fall between 49 and 0 'degrees' are taken to indicate progressively 'colder' or more negative attitudes. Neutral attitudes are indicated by rating the group at 50 'degrees'. Often a picture of a thermometer is provided with the question to aid interpretation. Thus, the feeling thermometer adequately allows for the expression of both attitude valence (cold, neutral, warm) and extremity (i.e., a score of 81 represents a 'warmer' or more 'favorable' attitude than does a rating of 63).

Samples

Feeling thermometers first appeared in the 1964 round of the American National Election Study (see Converse et al., 1980) asking about respondents' attitudes towards a number of groups (e.g., from farmers to conservatives to big businesses). Since its introduction, the feeling thermometer has been used extensively in politics (e.g., Conover & Feldman, 1981), medicine (e.g., Juniper et al., 1996), and aggression studies (e.g., Uhlmann & Swanson, 2004) as a measure of attitudes.

In social psychology, the feeling thermometer has been successfully used with young participants (e.g., $M_{age} = 13.8$ years, De Tezanos-Pinto et al., 2010), student samples (e.g., $M_{age} = 20$ years, Schmid, Hewstone, Tausch, Cairns, & Hughes, 2009), as well as adult samples (e.g., $M_{age} = 45.27$ years, Tausch et al., 2010, Study 2). In addition to American samples, the feeling thermometer has been used in a wide range of national contexts, including South Africa (Lolliot, 2013), Cyprus (Tausch et al., 2010, Study 1), Northern Ireland (Tausch et al., 2010, Studies 2 and 4), Norway (De Tezanos-Pinto et al., 2010), and Canada (Haddock, Zanna, & Esses, 1993).

Reliability

Internal Consistency

Turner and Feddes (2011) used multiple feeling thermometers (14 in total) to tap into a generalized prejudice score. They asked their undergraduate sample about their attitudes towards a series of social groups at two measurement occasions spaced six weeks apart. They found that the feeling thermometers formed a reliable measure of generalized outgroup attitudes at both measurement occasions (Cronbach alpha coefficients for Time 1: $\alpha \geq .91$, and Time 2: $\alpha \geq .96$). De Tezanos-Pinto et al. (2010), using a sample of Norwegian school students, found that the correlations between the feeling thermometer ratings of attitudes towards Turkish / Pakistani / Indian boys and girls were strong ($.82 \leq r \leq .87$). In a three-wave longitudinal study (measurement occasions were spaced six months apart; Lolliot, 2013), the feeling thermometer formed a reliable measure of White South African secondary school students' attitudes towards mixed-race South Africans (the Cronbach alpha coefficient for Time 1, 2, and 3 combined = .80), Black South Africans (alpha coefficient for Time 1, 2, and 3 combined = .85), African immigrants (alpha coefficient for Time 1, 2, and 3 combined = .80), and Indian South Africans (alpha coefficient for Time 1, 2, and 3 combined = .75).

Test–Retest

Feeling thermometers display good test–retest reliability across different time lags. Dasgupta and Greenwald (2001) reported test–retest correlations of .87 and .88 with measurement occasions spaced 24 hours apart (Experiment 1). Tausch et al. (2010) also reported one-year feeling thermometer autocorrelations ranging from .25 (ingroup feeling thermometer) to .30 and .51 (outgroup feeling thermometers), and Lolliot (2013) reported six-month autocorrelations for four outgroups that ranged from .49 to .76, and one-year autocorrelations ranging from .44 to .56.

The latter three-wave longitudinal study (Lolliot, 2013) also provided the chance to test for autoregressive stationarity of the feeling thermometer.[8] Results showed that the feeling thermometer ratings of attitudes towards the four outgroups exhibited autoregressive stationarity.

Validity

Convergent/Concurrent

Dasgupta and Greenwald (2001) reported significant correlations between their feeling thermometer and semantic differential scales (similar but not identical to the one discussed in the next section) for attitudes towards White ($r = .51$) and Black ($r = .43$) Americans (Experiment 1), and for attitudes towards the young ($r = .70$) and old ($r = .71$; Experiment 2). Turner and Feddes (2011) found that their respondents' outgroup feeling thermometer ratings correlated strongly with their ratings of the same outgroups using a different attitudinal measure ($.73 \leq .91$). These same-trait multimethod correlations indicate that the feeling thermometer shows good convergent validity.

Divergent/Discriminant

Tausch et al. (2010, Study 3) reported weak correlations with the Biased Inventory of Desirable Responding for their feeling thermometer measures of outgroup attitudes ($r \leq .17$). Sharp, Voci, and Hewstone (2011) reported weak correlations between their feeling thermometers and their measures of public self-consciousness ($|r| \leq .14$) and social comparison ($r \leq -.04$). Tausch, Tam et al. (2007, Study 1) reported significant negative correlations between the feeling thermometer and intergroup anxiety ($r = -.41$), symbolic ($r = -.37$) and realistic ($r = -.31$) threat.

Location

For the original feeling thermometer, consult the Inter-university Consortium for Political and Social Research's website at: www.icpsr.umich.edu (Retrieved May 25, 2014).

Results and Comments

The results reported here attest to the feeling thermometer being a psychometrically sound measure of outgroup attitudes. Given its analogous ties to a temperature thermometer, the feeling thermometer is easy and

[8]Testing for autoregressive stationarity tests the hypothesis that the degree to which a variable produces a change in later iterations of itself is consistent across consecutive time points.

intuitive to use (Converse & Presser, 1986). It also allows for the quick assessment of attitudes and represents an easy way to assess attitudes towards multiple groups, including the ingroup (e.g., Tausch et al., 2010).

While we have presented evidence in favor of the feeling thermometer, we should highlight some concerns to take into account before deciding on whether or not to use the feeling thermometer. The researcher interested in using the feeling thermometer to assess attitudes towards multiple outgroups should be aware of shared method variances (see Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) as using the same scale to measure attitudes towards different groups could artificially inflate the correlation(s) found between them. While a few studies have demonstrated that this is not a serious issue (Alwin, 1997; see also Lolliot, 2013; Tausch et al., 2010), shared method effects should nonetheless always be kept in mind when designing the questionnaire.

Another concern relates to the scaling properties of the feeling thermometer. Although 11-point feeling thermometers have been shown to be more reliable than 7-point rating scales (Alwin, 1997) and research has demonstrated that reliability goes up with increasing response categories (Alwin, 1992; Andrews, 1984), when answering the 100-point feeling thermometer, many respondents rely on using a restricted subset of the feeling thermometer, typically response categories that are multiples of 5 (i.e., 25, 30, 35, and so forth) with answers clustering around the mid-point of the scale (i.e., 50; Schaeffer & Presser, 2003). It is therefore unsurprising that standard deviations associated with feeling thermometers are typically large. It should, however, be noted that social scientists are more frequently distributing their surveys on the internet using online survey tools, most of which include online feeling thermometers where the participant can use a sliding bar to indicate their feelings towards an outgroup (their rating typically appears in a box to the right of the thermometer). Evidence from our online data indicates that individuals do make fuller use of the whole range of response options on these online feeling thermometers, and tend to rely less heavily on the mid-point.

Third, using single-item measures (such as a single feeling thermometer to assess attitudes towards outgroups) does not allow the researcher to take advantage of latent variable modeling techniques. By using multiple indicators for a latent variable, modeling techniques such as structural equation modeling are able to estimate and thus partial out measurement error allowing for more accurate parameter estimation. Sometimes, however, there are advantages to using single item measures. We will touch on these shortly.

Keeping these considerations in mind, we turn our attention to the psychometric properties of a second popular outgroup attitude measure: the general evaluation scale (Wright et al., 1997).

FEELING THERMOMETER

Below is something that looks like a thermometer (Fig. 23.1). We call it a 'feeling thermometer' because it measures your feelings towards groups. Here's how it works. If you don't know too much about a group, or don't feel particularly warm or cold toward them, then you should place them in the middle, at the 50 degree mark. If you have a warm feeling toward a group, or feel favorably toward it, you would give it a score somewhere between 50 and 100 depending on how warm your feeling is toward the group. On the other hand, if you don't feel very favorably toward some of these groups — if there are some you don't care for too much — then you would place them somewhere between the 0 and 50 mark.

How do you feel towards...

...[outgroup]?

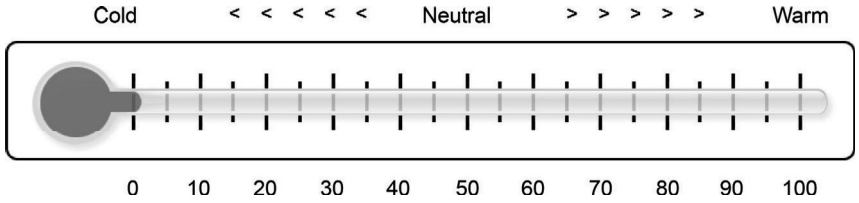

FIGURE 23.1 The Feeling Thermometer.

For more abbreviated instructions, see Schmid, Hewstone, and Tausch (2013).

General Evaluation Scale (GES)

(Wright et al., 1997).

Description

The Wright et al. (1997) article on the extended contact effect introduced the GES as a means of tapping outgroup attitudes. A semantic differential scale, the GES consists of six bipolar adjective pairs. These items are presented on opposite ends of two anchors (e.g., 1 = *negative* to 7 = *positive*). Thus, the GES is a measure of intergroup attitudes that captures both the valence and extremity of the attitude (for instance, taking the above negative-positive example, a score of 2 represents a more negative attitude than does a score of 4). Furthermore, the GES is approximately balanced containing both positively and negatively directed adjective pairs.

In addition to considering the full six-item GES, we will also investigate the reliability and validity evidence for a shorter, four-item version of the GES (Swart et al., 2011). The Swart et al. short-form version used a 5-point rating scale. For both the full and short-form versions, higher scores indicate more favorable attitudes.

Sample

The GES scale was first used on three American undergraduate samples (Wright et al., 1997, Studies 1 and 2). Study 1 ascertained the attitudes of 125 White Americans towards minority groups. The mean GES scores for this study ranged from 3.79 to 4.07 (no standard deviations were reported). In Study 2, Wright and colleagues used the GES with a second White American sample ($N = 132$) and a minority American sample ($N = 120$); the latter sample included Asian American, African American, and Latino/Latina respondents (means not reported). In a three-wave longitudinal study with six months separating measurement occasions, Swart et al. (2011) used a four-item adaptation of the GES to ascertain the attitudes of $N = 465$ mixed-race South African high school students (Time 1: $M_{age} = 14.69$ years, $SD = 1.07$) towards White South Africans. In their study, the mean GES scores were reported as Time 1: $M = 3.86$ ($SD = 0.76$); Time 2: $M = 3.61$ ($SD = 0.90$); Time 3: $M = 3.77$ ($SD = 0.84$).

Reliability

Internal Consistency

Wright et al. (1997, Study 1) reported a Cronbach alpha coefficient of .90. Other studies have reported Cronbach alphas that range from .68 (Stone & Crisp, 2007, Study 2) to .91 (Eller & Abrams, 2004).

While Wright et al. (1997, Study 2) did not report Cronbach alpha coefficients for their minority sample, Eller, Abrams, and Zimmermann (2011) reported good GES internal consistency results based on a sample of international students studying in the UK ($.79 \leq \alpha \leq .82$). Similarly, Swart et al.'s (2011) shortened version proved internally consistent in a sample of mixed-race South African students ($.68 \leq \alpha \leq .73$)

The GES has exhibited Cronbach alpha coefficients over time frames of two weeks (Time 1: $\alpha = .79$, Time 2 $\alpha = .78$, Eller & Abrams, 2003); 12 weeks (Time 1: $\alpha = .80$, Time 2: $\alpha = .81$, Gómez, Eller, & Vázquez, 2013); six months (Time 1: $\alpha = .82$, Time 2: $\alpha = .88$, Eller & Abrams, 2004); and one year (Time 1: $\alpha = .79$, Time 2: $\alpha = .82$, Eller et al., 2011).[9] The Swart et al. (2011) short-form GES showed adequate six-month alpha coefficients (Time 1: $\alpha = .68$, Time 2: $\alpha = .68$, Time 3: $\alpha = 73$).

While internally consistent among student samples (e.g., $M_{age} = 21.52$, $SD = 2.92$, $\alpha = .90$, Tausch, Hewstone et al., 2007), the full GES has also been shown to be internally consistent amongst elementary school (e.g., $M_{age} = 13.6$ years, $\alpha = .91$, Turner et al., 2007a, Study 3), high school (e.g., $M_{age} = 16.5$ years, $SD = 0.60$, $\alpha \geq .80$, Gómez et al., 2013), and retired adult (e.g., $M_{age} = 74.81$ years, $SD = 7.43$, $\alpha = 89$, Abrams, Eller, & Bryant, 2006) samples.

Test–Retest

Although the full six-item GES has been used in numerous longitudinal studies (see Eller & Abrams, 2003, 2004; Eller et al., 2011; Gómez et al., 2013), no test–retest correlations have been reported. For the short-form version, however, Swart et al. (2011) reported six-month stability coefficients ranging from $r = .13$ (time 1 to 2) to $r = .32$ (time 2 to 3); the one-year stability coefficients ($r = .37$).

Validity

Convergent/Concurrent

Eller and Abrams (2004) reported significant correlation coefficients between GES ratings of French people and Algerians ($r = .45$, Study 1), as well as between GES ratings of Americans and Canadians ($r = .30$, Study 2).

[9]Wright et al. (1997, Study 3) assessed outgroup attitudes at three time points spaced approximately two hours apart using an eight-item version of the GES, which showed good to excellent reliability (Time 1: $\alpha = .94$, Time 2: $\alpha = .85$, Time 3: $\alpha = .93$).

Furthermore, Tausch, Hewstone et al. (2007) reported significant positive correlations between contact quantity and quality and the GES ($r = .42$ and .46, respectively). Tam et al. (2007) reported significant positive correlations between the GES and intergroup contact ($r = .26$), intergroup forgiveness ($r = .34$), and positive emotions towards the outgroup ($r = .33$). Swart et al. (2011) found, cross-sectionally, that their short-form GES correlated significantly with cross-group friendships at time 1 ($r = .17$) and time 3 ($r = .21$), but not at time 2 ($r = .09$). Similarly, Swart et al. found their GES correlated significantly with outgroup variability at time 2 and 3 ($r = .12$ and .25, respectively), but non-significantly at time 1 ($r = .05$). Furthermore, Swart et al. reported positive, significant cross-sectional correlations between their GES and intergroup empathy ($r_{time\ 1} = .26$, $r_{time\ 2} = .17$, $r_{time\ 3} = .39$).

Divergent/Discriminant

Wright et al. (1997) reported that the GES correlated moderately and negatively with their Affective Prejudice scale ($r = -.31$; Study 1).[10] Eller et al. (2011) reported a significant negative correlation between the GES and intergroup anxiety ($r = -.43$; see Tausch, Hewstone et al., 2007; Voci & Hewstone, 2003). Tausch, Hewstone et al. (2007) reported significant negative correlations between the GES and symbolic ($r = -.47$) and realistic ($r = -.36$) threats, whereas Tam et al. (2007) reported significant negative correlations between the GES and fear of ($r = -.29$) and anger towards ($r = -.44$) the outgroup. Harwood et al. (2005, Study 2) reported a non-significant correlation between the GES and intergroup salience (e.g., $r = -.18$). The Swart et al. (2011) short-form GES correlated negatively and significantly with outgroup anxiety ($r_{time\ 1} = -.30$, $r_{time\ 2} = -.29$, $r_{time\ 3} = -.38$) and negative action tendencies ($r_{time\ 1} = -.32$, $r_{time\ 2} = -.29$, $r_{time\ 3} = -.37$).

Construct/Factor Analytic

Tausch, Hewstone et al. (2007), in a confirmatory analysis, included the GES as part of their measurement model. They reported satisfactory model fit statistics, thus helping to demonstrate that the GES is empirically distinguishable from contact quantity and quality, threats to the ingroup, and intergroup anxiety (Tausch, Hewstone et al., 2007).

Swart et al. (2011) provided EFA (maximum likelihood extraction with direct oblimin rotation) and CFA support for the factor structure of the short-form of the GES, and reported the short-form to be unidimensional. In a series of CFAs, Swart et al. found that the individual GES items were loaded significantly and strongly by the GES latent variable at all three time points (all factor loadings $\geq .49$). Furthermore, the factor loadings were invariant across measurement occasions (Swart et al., 2011).

Criterion/Predictive

Voci and Hewstone (2003) reported a positive association between intergroup contact and the GES. Eller and Abrams (2004, Study 1) reported significant positive regression weights between cross-group friendships with an outgroup and GES scores towards that outgroup ($\beta = = .39$) and GES scores towards an uninvolved outgroup ($\beta = .44$). Furthermore, Eller et al. (2011) also showed that both the quality of contact and extended contact with an outgroup was predictive of later, more favorable GES ratings of that outgroup. Tausch, Hewstone et al. (2007) reported significant negative associations between threats to the ingroup and the GES ($\beta = -.26$) as well as between intergroup anxiety and the GES ($\beta = -.31$; see also Eller et al., 2011; Voci & Hewstone, 2003).

For the short-form GES, Swart et al. (2011) reported that empathy towards the outgroup was positively associated with later outgroup GES scores.

Location

Original six-item scale:

Wright, S.C., Aron, A., McLaughlin-Volpe, T., & Ropp, S.A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*, 73–90.

[10]Across three time points spaced roughly two hours apart, Wright et al.(1997, Study 3) found their eight-item version of the GES, which was based on the original six-item version, correlated with two other measures of outgroup attitudes, namely a scale measuring intergroup differentiation ($r_{time\ 1} = -.25$; $r_{time\ 2} = -.46$; $r_{time\ 3} = -.49$) and a measure of intergroup bias in resource allocation ($r_{time\ 1} = -.36$; $r_{time\ 2} = -.25$; $r_{time\ 3} = -.26$).

Short version (four items):

Swart, H., Hewstone, M., Christ, O., & Voci, A. (2011). Affective mediators of intergroup contact: A three-wave longitudinal study in South Africa. *Journal of Personality and Social Psychology*, *101*, 1221–1238.

Results and Comments

The GES represents a psychometrically sound multi-item measure of outgroup attitudes that has yielded reliable results across multiple samples spanning different ages, population groups, and contexts. The GES is easy to understand and simple to use. As far as which GES should be used (the long or short form), a number of options have to be taken into consideration. For example, how much space is available in the questionnaire will guide whether the six-item or four-item GES should be used. Hayduk and Pazderka-Robinson (2007) argued that the fewest best items from a scale should be used (cf. Little, Cunningham, Shahar, & Widaman, 2002, on item-parcelling). For younger participants, we would suggest using Turner and colleagues' (2007a) version of the GES. If using the full six items for younger samples is likely to lead to participant fatigue, we would suggest randomly selecting an appropriate subset of the full GES. From a latent variable perspective, however, we would advise using no fewer than three items (see Little et al., 2002). The question, then, remains: Which attitudinal measure to use? Pettigrew and Tropp (2006) found that single-item outgroup attitude measures and (reliable) multi-item measures performed similarly in their associative power with contact. This, to a large extent, is consistent with Hayduk and Pazderka-Robinson's (2007) contention that researchers should use the fewest 'best' items of a construct, even if this entails using a single item. To this end, the feeling thermometer is, we believe, a good candidate. In the same breath, the GES is also attractive as its multi-item nature allows the researcher to form latent variables that are able to control for measurement error. Remember that Krosnick, Judd, and Wittenbrink (2005) contended that multi-item measures are more instructive, especially with regards to construct validity. The answer to the question, 'Which measure to use?', then boils down to the researchers' intention(s). For example, if one is aiming to assess attitudes towards multiple outgroups, as research on the secondary transfer effect of contact requires (see Lolliot et al., 2013; Tausch et al., 2010), or if space is limited in the survey instrument, then feeling thermometers may prove optimal. On the other hand, if space allows, the GES – with its ability to form latent attitude variables that are internally consistent and show good construct validity – should be used.

GENERAL EVALUATION SCALE

Please describe how you feel about [outgroup] in general:

1. warm-cold (reverse coded)
2. negative–positive
3. friendly–hostile (reverse coded)
4. suspicious–trusting
5. respect–contempt (reverse coded)
6. admiration–disgust (reverse coded)

Short form

1. negative-positive
2. hostile–friendly
3. suspicious–trusting
4. contempt–respect

Note: Reproduced with permission.

Future Research Directions

We have endeavored to provide a well-functioning 'toolkit' consisting of measures of the most important theoretical constructs to researchers interested in contact theory. We call for future researchers to investigate the psychometric properties of the measures covered in this chapter *as well as* other intergroup contact measures. Where possible, we have contrasted two or more measures of a construct in an attempt to provide multiple options when considering hypotheses and questionnaire design and highlighted what should be kept in mind when choosing a specific measure.

Given the burgeoning diversity of many societies, research on intergroup contact and its ameliorative effects is needed now more than ever, so we emphasize that the 'tools' we use in our investigations need to be finely tuned. We hope that this chapter will help future researchers to choose the best measures to test theoretical advances and practical applications of intergroup contact theory.

We believe that we have provided compelling evidence for the psychometric properties for the constructs that appear in this chapter. It should be nonetheless kept in mind that we strongly advocate that future research seeks to formally establish the results presented in this chapter.

References

Aberson, C. L., & Gaffney, A. M. (2008). An integrated threat model of explicit and implicit attitudes. *European Journal of Social Psychology*, *39*, 808–830.

Abrams, D., Eller, A., & Bryant, J. (2006). An age apart: The effects of intergenerational contact and stereotype threat on performance and intergroup bias. *Psychology and Aging*, *21*, 691–702. Available from http://dx.doi.org/doi:10.1037/0882-7974.21.4.691.

Allport, G. W. (1954). *The nature of prejudice*. Cambridge/Reading, MA: Addison-Wesley.

Alwin, D. F. (1992). Information transmission in the survey interview: Number of response categories and the reliability of attitude measurement. *Sociological Methodology*, *22*, 83–118.

Alwin, D. F. (1997). Feeling thermometers versus 7-point scales: Which are better? *Sociological Methods Research*, *25*, 318–340.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*, 411–423.

Andrews, F. M. (1984). Construct validity and error components of survey measures: A structural equation modeling approach. *Public Opinion Quarterly*, *48*, 409–442.

Bagozzi, R. P., & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, *1*, 45–87.

Baker, P. E. (1934). Negro-White Adjustment in America. *Journal of Negro Education*, *3*, 194–204.

Barlow, F. K., Paolini, S., Pedersen, A., et al. (2012). The contact caveat negative contact predicts increased prejudice more than positive contact predicts reduced prejudice. *Personality and Social Psychology Bulletin*, *38*, 1629–1643.

Binder, J., Zagefka, H., Brown, R., et al. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis amongst majority and minority groups in three European countries. *Journal of Personality and Social Psychology*, *96*, 843–856.

Bizman, A., & Yinnon, Y. (2001). Intergroup and interpersonal threats as determinants of prejudice: The moderating role of in-group identification. *Basic and Applied Social Psychology*, *23*, 191–196.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology*, *80*, 253–267.

Bodenhausen, G. V. (1993). Emotion, arousal, and stereotypic judgement: A heuristic model of affect and stereotyping. In D. Mackie, & D. Hamilton (Eds.), *Affect, cognition, and stereotyping: Interactive processes in intergroup perception* (pp. 13–37). San Diego, CA: Academic.

Bogardus, E. S. (1925). Social distance and its origins. *Journal of Applied Sociology*, *9*, 216–226.

Bogardus, E. S. (1928). *Immigration and race attitudes*. Boston: Heath.

Bornstein, R. F. (1989). Exposure and affect: Overview and met-analysis of research, 1968–1987. *Psychological Bulletin*, *106*, 265–289.

Brewer, M. B., & Miller, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller, & M. B. Brewer (Eds.), *Groups in Contact: The psychology of desegregation* (pp. 281–302). Orlando, FL: Academic.

Britt, T. W., Boniecki, K. A., Vescio, T. K., Biernat, M. R., & Brown, L. M. (1996). Intergroup anxiety: A Person × Situation approach. *Personality and Social Psychology Bulletin*, *22*, 1177–1188.

Broman, C. L., Mavaddat, R., & Hsu, S. Y. (2000). The experience and consequences of perceived racial discrimination: A study of African Americans. *Journal of Black Psychology*, *26*, 165–180.

Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 255–342). San Diego, CA: Elsevier.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81–105.

Cernat, V. (2010). Intergroup contact in Romania: When minority size is positively related to intergroup conflict. *Journal of Community & Applied Social Psychology*, *20*, 15–29.

Christ, O., Hewstone, M., Tausch, N., Wagner, U., Voci, A., Hughes, J., & Cairns, E. (2010). Direct contact as a moderator of extended contact effects: Cross-sectional and longitudinal impact on outgroup attitudes, behavioral intentions, and attitude certainty. *Personality and Social Psychology Bulletin*, *36*, 1662–1674.

Conover, P. J., & Feldman, S. (1981). The origins and meaning of Liberal/Conservative self-identifications. *American Journal of Political Science*, *25*, 617–645.

Converse, J. M., & Presser (1986). *S. Survey questions: Handcrafting the standardized questionnaire*. California, USA: Sage.

Converse, P. E., Dotson, J. D., Hoag, W. J., & McGee, W. H. (1980). *American social attitudes data sourcebook 1947–1978*. Cambridge, Mass: Harvard University Press.

Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*, 800–814.

Davies, K., Tropp, L. R., Aron, A., Pettigrew, T. F., & Wright, S. C. (2011). Cross-group friendships and intergroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, *15*, 332–351.

De Tezanos-Pinto, P., Bratt, C., & Brown, R. (2010). What will the others think? In-group norms as a mediator of the effects of intergroup contact. *British Journal of Social Psychology*, *49*, 507–523.

Dhont, K., Van Hiel, A., De Bolle, M., & Roets, A. (2012). Longitudinal intergroup contact effects on prejudice using self- and observer-reports. *British Journal of Social Psychology*, *51*, 221–238.

Dijker, A. J. (1987). Emotional reactions to ethnic minorities. *European Journal of Social Psychology*, *17*, 305–325.

Dovidio, J., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510–540.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace.

Eller, A., & Abrams, D. (2003). 'Gringos' in Mexico: Cross-sectional and longitudinal effects of language school-promoted contact on intergroup bias. *Group Processes and Intergroup Relations*, *6*, 55–75.

Eller, A., & Abrams, D. (2004). Come together: Longitudinal comparisons of Pettigrew's reformulated intergroup contact model and the common ingroup identity model in Anglo-French and Mexican-American contexts. *European Journal of Social Psychology*, *34*, 229–256.

Eller, A., Abrams, D., & Zimmermann, A. (2011). Two degrees of separation: A longitudinal study of actual and perceived extended international contact. *Group Processes & Intergroup Relations*, *14*, 175–191.

Esses, V. M., Jackson, L. M., Dovidio, J. F., & Hodson, G. (2005). Instrumental relations among groups: Group competition, conflict and prejudice. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 227–243). Malden, MA: Blackwell.

Gallagher, A. M. (1995). The approach of government: Community relations and equity. In S. Dunn (Ed.), *Facets of the conflict in Northern Ireland* (pp. 27–43). New York: St. Martin's Press.

Gaertner, S. L., Dovidio, J., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology*, *4*, 1–26.

Gómez, Á., Eller, A., & Vázquez, A. (2013). Verification of ingroup identity as a longitudinal mediator betweeen intergroup contact and outgroup evaluation. *Spanish Journal of Psychology*, *16*, 1–11.

Greenland, K., & Brown, R. (1999). Categorization and intergroup anxiety in contact between British and Japanese nationals. *European Journal of Social Psychology*, *29*, 503–521.

Greenland, K., Xenias, D., & Maio, G. (2012). Intergroup anxiety from the self and other: Evidence from self-report, physiological effects, and real interactions. *European Journal of Social Psychology*, *42*, 150–163.

Haddock, G., Zanna, M. P., & Esses, V. M. (1993). Assessing the structure of prejudicial attitudes: The case of attitudes toward homosexuals. *Journal of Personality and Social Psychology*, *65*, 1105–1118.

Harwood, J., Hewstone, M., Paolini, S., & Voci, A. (2005). Grandparent-grandchild contact and attitudes towards older adults: Moderator and mediator effects. *Personality and Social Psychology Bulletin*, *31*, 393–406.

Hayduk, L. A., & Pazderka-Robinson, H. (2007). Fighting to understand the world causally: Three battles connected to the causal implications of structural equation models. In W. Outhwaite, & S. Turner (Eds.), *Sage handbook of social science methodology* (pp. 147–171). London: Sage.

Hewstone, M. (2009). Living apart, living together? The role of intergroup contact in social integration. *Proceedings of the British Academy*, *162*, 243–300.

Hewstone, M., Judd, C. M., & Sharp, M. (2011). Do observer ratings validate self-reports of intergroup contact?: A round-robin analysis. *Journal of Experimental Social Psychology*, *47*, 599–609.

Hooghe, M., Reeskens, T., Stolle, D., & Trappers, A. (2009). Ethnic diversity and generalized trust in Europe: A cross-national multilevel study. *Comparative Political Studies*, *42*, 198–223.

Islam, M. R. (1992). *Hindu-Muslim intergroup relations in Bangladesh: A cognitive-intergroup analysis. (Unpublished doctoral dissertation)*. Bristol, UK: University of Bristol.

Islam, M. R., & Hewstone, M. (1993). Dimensions of contact as predictors of intergroup anxiety, perceived out-group variability, and outgroup attitude: An integrative model. *Personality and Social Psychology Bulletin*, *19*, 700–710.

Juniper, E. F., Guyatt, G. H., Feeny, D. H., Ferrie, P. J., Griffith, L. E., & Townsend, M. (1996). Measuring quality of life in children with asthma. *Quality of Life Research*, *5*, 35–46.

Khmelkov, V. T., & Hallinan, M. T. (1999). Organizational effects on race relations in schools. *Journal of Social Issues*, *55*, 627–645.

Krosnick, J. A., Judd, C. M., & Wittenbrink, B. (2005). Attitude measurement. In D. Albarracin, B. T. Johnson, & M. P. Zanna (Eds.), *Handbook of attitudes and attitude change*. Mahwah, NJ: Erlbaum.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*, 1–55.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, *9*, 151–173.

Lolliot, S. (2013). *The secondary transfer effect of contact (Unpublished doctoral dissertation)*. Oxford, UK: Oxford University.

Lolliot, S., Schmid, K., Hewstone, M., Al Ramiah, A., Tausch, N., & Swart, H. (2013). Generalized effects of intergroup contact: The secondary transfer effect. In G. Hodson, & M. Hewstone (Eds.), *Advances in intergroup contact* (pp. 81–112). London, UK: Psychology Press.

Mähönen, T. A., Jasinskaja-Lahti, I., & Liebkind, K. (2011). The impact of perceived social norms, gender, and intergroup anxiety on the relationship between intergroup contact and ethnic attitudes of adolescents. *Journal of Applied Social Psychology*, *41*, 1877–1899.

Marsh, H. W., & Hocevar, D. (1988). A new, more powerful approach to multitrait multimethod analyses: Application of 2nd-order confirmatory factor-analysis. *Journal of Applied Psychology*, *73*, 107–117.

Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interactions with White and Black men. *Personality and Social Psychology Bulletin*, *28*, 939–952.

Nightingale, C. H. (2012). *Segregation: A global history of divided cities*. Chicago, IL: University of Chicago Press.

Paolini, S., Harwood, J., & Rubin, M. (2010). Negative intergroup contact makes group memberships salient: Explaining why intergroup conflict endures. *Personality and Social Psychology Bulletin*, *36*, 1723–1738.

Paolini, S., Hewstone, M., Cairns, E., & Voci, A. (2004). Effects of direct and indirect cross-group friendships on judgments of Catholics and Protestants in Northern Ireland: The mediating role of an anxiety reduction mechanism. *Personality and Social Psychology Bulletin*, *30* 770–786.

Paolini, S., Hewstone, M., Voci, A., Harwood, J., & Cairns, E. (2006). Intergroup contact and the promotion of intergroup harmony: The influence of intergroup emotions. In R. Brown, & D. Capozza (Eds.), *Social identities: Motivational, emotional and cultural influences* (pp. 209–238). Hove, UK: Psychology Press.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*, 65–85.

Pettigrew, T. F. (2008). Future directions for intergroup contact theory and research. *International Journal of Intercultural Relations*, *32*, 187–199.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytical test of the intergroup contact theory. *Journal of Personality and Social Psychology*, *90*, 751–783.

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*, 922–934.

Plant, E. A., & Devine, P. G. (2003). The antecedents and implications of interracial anxiety. *Personality and Social Psychology Bulletin*, *29*, 790–801.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903.

Prestwich, A., Kenworthy, J. B., Wilson, M., & Kwan-Tat, N. (2008). Differential relations between two types of contact and implicit and explicit racial attitudes. *British Journal of Social Psychology*, *47*, 575–588.

Quillian, L. (1995). Prejudice as a response to perceived group threat: Population composition and anti-immigrant and racial prejudice in Europe. *American Sociological Review*, *60*, 586–611.

Schaeffer, N. C., & Presser, S. (2003). The science of asking questions. *Annual Review of Sociology*, *29*, 65–88.

Schmid, K., & Hewstone, M. (2010). Combined effects of intergroup contact and multiple categorization: Consequences for intergroup attitudes in diverse social contexts. In R. Crisp (Ed.), *The psychology of social and cultural diversity* (pp. 297–321). Chichester, UK: Wiley-Blackwell.

Schmid, K., Hewstone, M., & Tausch, N. (2013). Secondary transfer effects of intergroup contact via social identity complexity. *British Journal of Social Psychology*, Advanced online publication.

Schmid, K., Hewstone, M., Tausch, N., Cairns, E., & Hughes, J. (2009). Antecedents and consequences of social identity complexity: Intergroup contact, distinctiveness threat, and outgroup attitudes. *Personality and Social Psychology Bulletin*, *35*, 1085–1098.

Sharp, M., Voci, A., & Hewstone, M. (2011). Individual difference variables as moderators of the effect of extended cross-group friendship on prejudice: Testing the effects of public self-consciousness and social comparison. *Group Processes & Intergroup Relations*, *14*, 207–221.

Soliz, J., & Harwood, J. (2006). Shared family identity, age salience, and intergroup contact: Investigations of the grandparent-grandchild relationship. *Communication Monographs*, *73*, 87–107.

Steenkamp, C. (2009). Xenophobia in South Africa: What does it say about trust? *The Round Table*, *98*, 439–447.

Stephan, C. W., Stephan, W. G., Demitrakis, K. M., Yamada, A. M., & Clason, D. L. (2000). Women's attitudes toward men: An integrated threat theory approach. *Psychology of Women Quarterly*, *24*, 63–73.

Stephan, W. G., Boniecki, K. A., Ybarra, O., Bettencourt, A., Ervin, K. S., Jackson, L. A., McNatt, P. S., & Renfro, C. L. (2002). The role of threats in the racial attitudes of Blacks and Whites. *Personality and Social Psychology Bulletin*, *28*, 1242–1254.

Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, *41*, 157–175.

Stephan, W. G., & Stephan, C. W. (2000). An integrated threat theory of prejudice. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 23–46). Hillsdale, NJ: Erlbaum.

Stephan, W. G., Stephan, C. W., & Gudykunst, W. B. (1999). Anxiety in intergroup relations: A comparison of anxiety/uncertainty management theory and integrated threat theory. *International Journal of Intercultural Relations*, *23*, 613–628.

Stone, C., & Crisp, R. J. (2007). Superordinate and subgroup identification as predictors of intergroup evaluation in common ingroup contexts. *Group Processes & Intergroup Relations*, *10*, 493–513. Available from http://dx.doi.org/doi:10.1177/1368430207081537.

Swart, H., Hewstone, M., Christ, O., & Voci, A. (2010). The impact of cross-group friendships in South Africa: Affective mediators and multigroup comparisons. *Journal of Social Issues*, *66*, 309–333.

Swart, H., Hewstone, M., Christ, O., & Voci, A. (2011). Affective mediators of intergroup contact: A three-wave longitudinal study in South Africa. *Journal of Personality and Social Psychology*, *101*, 1221–1238.

Tam, T., Hewstone, M., Cairns, E., Tausch, N., Maio, G., & Kenworthy, J. (2007). The impact of intergroup emotions on forgiveness in Northern Ireland. *Group Processes & Intergroup Relations*, *10*, 119–135.

Tam, T., Hewstone, M., Kenworthy, J., & Cairns, E. (2009). Intergroup trust in Northern Ireland. *Personality and Social Psychology Bulletin*, *35*, 45–59.

Tausch, N., Hewstone, M., Kenworthy, J., Cairns, E., & Christ, O. (2007). Cross-community contact, perceived status differences, and intergroup attitudes in Northern Ireland: The mediating roles of individual-level versus group-level threats and the moderating role of social identification. *Political Psychology*, *28*, 53–68.

Tausch, N., Hewstone, M., Kenworthy, J. B., Psaltis, C., Schmid, K., Popan, J. R., Cairns, E., & Hughes, J. (2010). Secondary transfer effects of intergroup contact: Alternative accounts and underlying processes. *Journal of Personality and Social Psychology*, *99*, 282–302.

Tausch, N., Hewstone, M., Schmid, K., Hughes, J., & Cairns, E. (2011). Extended contact effects as a function of closeness of relationship with ingroup contacts. *Group Processes and Intergroup Relations*, *14*, 239–254.

Tausch, N., Tam., T., Hewstone, M., Kenworthy, J., & Cairns, E. (2007). Individual-level and group-level mediators of contact effects in Northern Ireland: The moderating role of social identification. *British Journal of Social Psychology*, *46*, 541–556.

Thurstone, L. L. (1931). The measurement of social attitudes. *The Journal of Abnormal and Social Psychology*, *26*, 249–269.

Turner, R. N., & Feddes, A. R. (2011). How intergroup friendship works: A longitudinal study of friendship effects on outgroup attitudes. *European Journal of Social Psychology*, *41*, 914–923.

Turner, R. N., Hewstone, M., & Voci, A. (2007a). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. *Journal of Personality and Social Psychology*, *93*, 369–388.

Turner, R. N., Hewstone, M., Voci, A., Paolini, S., & Christ, O. (2007b). Reducing prejudice via direct and extended cross-group friendship. In W. Stroebe, & M. Hewstone (Eds.), *European review of social psychology* (Vol. 18, pp. 212–255). Hove, UK: Psychology Press.

Turner, R. N., Hewstone, M., Voci, A., & Vonofakou, C. (2008). A test of the extended intergroup contact hypothesis: The mediating role of intergroup anxiety, perceived ingroup and outgroup norms, and inclusion of the outgroup in the self. *Journal of Personality and Social Psychology*, *95*, 843–860.

Uhlmann, E., & Swanson, J. (2004). Exposure to violent video games increases automatic aggressiveness. *Journal of Adolescence*, *27*, 41–52.

Van Oudenhoven, J. P., Groenewoud, J. T., & Hewstone, M. (1996). Cooperation, ethnic salience and generalization of interethnic attitudes. *European Journal of Social Psychology*, *26*, 649–661.

Visser, P. S., & Krosnick, J. A. (1998). Development of attitude strength over the life cycle: Surge and decline. *Journal of Personality and Social Psychology*, *75*, 1389–1410.

Voci, A., & Hewstone, M. (2003). Intergroup contact and prejudice towards immigrants in Italy: The mediational role of anxiety and the moderational role of group salience. *Group Processes and Intergroup Relations*, *6*, 37–54.

Vonofakou, C., Hewstone, M., Voci, A., Paolini, S., Turner, R., Tausch, N., Tam, T., Harwood, J., & Cairns, E. (2008). The impact of direct and extended cross-group friendships on improving intergroup relations. In U. Wagner, L. R. Tropp, G. Finchilescu, & C. Tredoux (Eds.), *Improving intergroup relations — Building on the legacy of Thomas F. Pettigrew* (pp. 107—124). Oxford, UK: Blackwell.

Wagner, U., Christ, O., Pettigrew, T. F., Stellmacher, J., & Wolf, C. (2006). Prejudice and minority proportion: Contact instead of threat effects. *Social Psychology Quarterly*, *69*, 380—390.

Wagner, U., Hewstone, M., & Machliet, U. (1989). Contact and prejudice between Germans and Turks: A correlational study. *Human Relations*, *42*, 561—574.

Williams, R. M. (1947). *The reduction of intergroup tensions*. New York: Social Science Research Council.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101—126.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*, 73—90.