Meta-analytic evidence for higher implicit affiliation and intimacy motivation scores in women, compared to men

Amely Drescher, Oliver C. Schultheiss

Friedrich-Alexander University, Erlangen, Germany

ABSTRACT

We used meta-analysis to test for gender differences in implicit needs for affiliation/intimacy, assessed via story-coding methods. We included thirty-three effect sizes from 26 publications and 2 unpublished studies, covering a total of 5962 research participants (58% female). Across studies, women scored higher than men in measures of implicit affiliation motivation ($d^* = 0.45, 95\% CI = [0.37; 0.53]$). This finding was not moderated by the coding system used, gender congruence of the picture cues presented, or correction for protocol length. Men and women did not differ in their implicit needs for power ($N = 2493, k = 15, d^* = −0.19, 95\% CI = [−0.44; 0.05]) or achievement ($N = 2235, k = 13, d^* = 0.14, 95\% CI = [−0.03; 0.30]$).

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

Research on implicit motives, that is, nonconsciously operating affective preferences for specific classes of incentives, has a long tradition and remains an active area of research (Schultheiss & Brunstein, 2010), with recent studies exploring such diverse phenomena as cross-cultural patterns of sociosexuality (Hofer et al., 2010), relationship satisfaction (Hagemeyer, Schönbrodt, Neyer, Neberich, & Asendorpf, 2015; Job, Bernecker, & Dweck, 2012), or hormonal responses to stress (Schultheiss, Wiemers, & Wolf, 2014). Yet despite the centrality of the implicit motive construct and its measures for understanding personality and motivation, there is little systematic research on whether and how women and men differ in their implicit motivational needs. So far, only two qualitative literature reviews have attempted to address this issue. They have come to somewhat different conclusions: Whereas Stewart and Chester (1982) saw no evidence for fundamental gender differences in implicit motives and their sensitivity to motivational arousal, Duncan and Peterson (2010) have noted that research published since 1982 suggests that women score higher than men on measures of the implicit need for affiliation, but not on other motive measures. However, they did not provide a quantitative estimate for this gender difference.

In the present research, we aim to fill this gap through a meta-analysis focused on gender differences in implicit motive measures related to affiliation and potential moderators of gender differences. We also report findings for measures of the motivational needs for power and achievement whenever these were included in the studies that resulted from our targeted literature search.

1.1. The family of implicit affiliation motive measures

Researchers in the McClelland-Atkinson tradition have developed several distinct measures tapping various aspects of the need for affiliation, broadly defined here as a capacity for deriving pleasure from being with others and experiencing social separation as aversive (see Schultheiss, 2008). These measures share two crucial features. First, they are based on story-telling methods of assessment. To obtain measurements of participants’ implicit affiliative needs, researchers present ambiguous picture cues to their study participants, a procedure called picture story exercise (PSE; see McClelland, Koestner, & Weinberger, 1989; sometimes sentence cues are used instead of pictures; see French, 1956). On the PSE, participants are required to write an imaginative story about each
cure. These stories are then coded, based on coding systems featuring clearly defined coding rules, for affiliation-related imagery, and a score is derived for each participant by summing up all instances of coded imagery.

The second crucial feature is that affiliation motive measures were derived from experimental studies in which a motivational state was induced in an experimental group, but not a control group. Participants of both groups then wrote imaginative stories about pictures suggestive of affiliative contact. Researchers distill differences between experimental-group and control-group stories into content coding systems that aimed to capture the themes that were unique to aroused affiliation motivation (Winter, 1998). Due to way they were derived, these measures therefore all fulfill a core requirement for validity; that is, it has been demonstrated that changes in the targeted attribute (here: affiliation motivation) have a causal impact on changes in its measure (here: frequency of affiliative imagery) (see Borsboom, Mellenbergh, & van Heerden, 2004; McClelland, 1958, 1987).

The combination of these defining features differentiate affiliation motive measures in the McClelland-Atkinson tradition of motive research from self-report measures of affiliation motivation and older or other contemporary thematic apperceptive methods that were not derived through experimental motive arousal and which therefore are not the focus of our study.1 Their independence from measures of self-attributed affiliation motivation has been demonstrated meta-analytically, with the variance overlap between empirically derived picture-story methods and self-report measures of affiliation motivation being less than 2% (Köllner & Schultheiss, 2014). In the following, we provide brief descriptions of each of the implicit affiliation motive measures that have been described and used in research reports since the 1950s.

1.1. n Affiliation

The first measure of the need for affiliation (abbreviated n Affiliation) was introduced by Atkinson, Hemys, and Veroff (1958); see also Heyns, Veroff, & Atkinson, 1992) and built on an earlier study by Shipley and Veroff (1958). Atkinson et al. (1958) defined n Affiliation as a strong concern with establishing, maintaining or restoring a positive relationship with another person or group. According to the coding system, people high in n Affiliation are driven either by the search for the pleasure of close, harmonious social contact or the avoidance of social rejection and exclusion. Perhaps as a consequence of this inherent duality of the measure, research conducted with it has painted a mixed picture of the affiliation-motivated person, with some evidence supporting the idea that affiliation-motivated individuals seek proximity to others, but also evidence that they shun others once they view them as too dissimilar to themselves. Moreover, their fear of rejection may make them anxious, demanding, and ultimately unpopular with others (for summaries, see Weinberger, Cotler, & Fishman, 2010; Winter, 1996). Boyatzis (1973) therefore called for the development of measures targeting more specifically the capacity for interpersonal closeness and love.

1.1.2. n Intimacy

McAdams (1980) developed a measure for the need for intimacy (n Intimacy), defined as a constant preference or readiness to seek experiences of warm, close, and communicative social interaction. The core experience of intimacy motivation is a noninstrumental, reciprocal sharing of desires, feelings and thoughts. Close dyadic relationships are more important to highly intimacy-motivated individuals than mere belonging to a social group. In contrast to affiliation motivation, intimacy motivation does not appear to be characterized by a fear of rejection (McAdams, Jackson, & Kirshnit, 1984; but see Hofer & Busch, 2011). In general, intimacy-motivated individuals are happy and satisfied in dyadic interactions and in life more generally and are liked by others (McAdams et al., 1984; Weinberger et al., 2010).

1.1.3. n Affiliation-Intimacy

In an attempt to devise a comprehensive coding system integrating previous measures, Winter (1991, 1994) combined affiliation and intimacy motivation into one category. This step was based upon the conceptual overlap of both motives in terms of shared social interaction and warm feelings towards others and also on the substantial variance overlap between their measures (see Hofer & Busch, 2011; McAdams et al., 1984). Winter’s integrated coding system does not allow differentiating between n Affiliation and n Intimacy.

1.1.4. Affiliative trust-mistrust

McKay’s (1992) coding system for affiliative trust-mistrust focuses on a balanced assessment of positive and negative aspects of dispositional affiliation motivation. It contains two independent scales that quantify the degree to which a person describes close relationships as dependable, warm, and rewarding (trusting) or not dependable, warm, and rewarding (mistrusting). Both scales can be combined into an overall trust-mistrust difference score. This measure has been linked to immune system functions (McKay, 1991; McKay et al., 1997), reflecting a key role of affiliation motivation in health and disease (see McClelland, 1989).2

Although there is no single, comprehensive study that has assessed n Affiliation, n Intimacy, n Affiliation-Intimacy, and affiliative trust-mistrust in one sample and determined their shared overall variance, several studies exist that document substantial variance overlap between individual members of this family of measures (e.g. Hofer & Busch, 2011: r(271) Affiliation × n Intimacy = 0.67; McKay, 1992: r(70) affiliative trust × n Affiliation = 0.50, r(70) affiliative trust × n Intimacy = 0.32; Winter, 1991: r(42) Affiliation-Intimacy × n Affiliation = 0.40, r(42) n Affiliation-Intimacy × n Intimacy = 0.41). We suggest that this, along with the close conceptual relationships between the original studies that derived the coding systems from experimental arousal experiments, justifies the inclusion of findings obtained with these different measures in our meta-analysis.

1.2. Potential moderators of gender differences in implicit affiliation motive scores

Because the literature reviews by Stewart and Chester (1982) and Duncan and Peterson (2010) came to different conclusions regarding the existence of gender differences in affiliation motivation, our first goal in this study is to settle the issue systematically and quantitatively through meta-analytic techniques. But even if we find a gender difference in affiliation motivation favoring women, as suggested by Duncan and Peterson (2010), substantive psychological interpretations of such differences (see Section 4)

---

1. Contemporary motive measures that did not meet these criteria and whose conceptual and empirical convergence with PSE-type implicit motive measures has been questioned in recent research include the Operant Motive Test and the Multi-Motive Grid (see Schuler, Brandstätter, Wegner, & Baumann, 2015; Schultheiss, Yankova, Dirlikov, & Schad, 2009).

2. Siegel and Weinberger (1998) have introduced the concept and measurement of the oneness motive, defined as the need to become part of, be at one with, or belong to a larger whole, as another addition to the family of affiliation motivation measures. Similar to n Intimacy, the oneness motive can be characterized as largely positive and emerges most clearly in interpersonal relationships (Weinberger et al., 2010). However, at the time we conducted this meta-analysis, only few studies had been published using this measure and none had considered gender differences. We therefore did not consider this measure in our conceptual review and meta-analysis.
should only be considered if it can be shown that the difference is not primarily or solely due to methodology. One such methodological factor may be picture set differences. Stewart and Chester (1982) note that in research up until 1982, many studies on affiliation motivation had used different picture sets for female and male research participants. Duncan and Peterson (2010) pointed out that if pictures of women elicit higher levels of affiliation imagery in both men and women than pictures of men, but the former are only shown to women and the latter only to men, then an apparent gender difference in motive score levels could merely be an artifact of the use of gender-congruent picture sets. We therefore examined picture set (same or different for men and women) as a possible moderator of gender differences in affiliation motives in the present study.

Another possible influence on apparent motivational gender differences is overall PSE story length. For two large samples, Schultheiss and Brunstein (2001) and Pang and Schultheiss (2005) reported a gender difference not only in raw affiliation motive scores, but also in overall PSE word count, with women having higher scores in both. When these researchers regressed word count out of motive scores, however, the gender difference in affiliation motive levels persisted. Nevertheless, because overall PSE protocol length – that is, the matrix in which motives are scored – and affiliation motive imagery are correlated and both show a similar gender difference, statistical control for PSE protocol length (or lack thereof) may moderate whatever association there is between gender and affiliation motive levels, and we therefore took this variable into account in our analysis.

A third potential methodological source of gender differences is the coding system used. As sketched out above, one key difference between coding systems is whether they capture both approach (hope of closeness) and avoidance (fear of rejection) aspects of affiliation motivation (i.e., n Affiliation, n Affiliation-Intimacy, affiliative trust-mistrust) or focus primarily on the approach side of affiliation motivation (i.e., n Intimacy). From a review of individual studies, it is difficult to determine whether differences in coding systems are systematically associated with gender differences in motive scores. For instance, McAdams (1980) and McAdams and Constantian (1983) failed to find significant or consistent sex differences in n Intimacy and n Affiliation across several studies, but McAdams, Lester, Brand, McNamara, and Lensky (1988) found higher levels of n Intimacy in women than men across two studies, and so did Schultheiss and Brunstein (2001) and Pang and Schultheiss (2005) for n Affiliation-Intimacy. It is presently unclear whether these results can be interpreted as evidence for a stronger sex difference for coding systems incorporating fear-of-rejection components than for coding systems focusing on hope of closeness only, which would suggest that men and women do not differ in affiliation motivation per se, but perhaps only in its approach and avoidance aspects (for gender differences in traits broadly related to approach and avoidance, see for instance Lynn & Martin, 1997). We hoped to gain more clarity about this issue by looking at coding systems as a moderator variable.3

While our meta-analysis specifically targets measures of affiliation motivation, many of the included studies also featured measures of the needs for power (n Power), defined as an affective preference for having impact on others (Winter, 1973), and achievement (n Achievement), defined as an affective preference for the autonomous mastery of challenging tasks (McClelland & Koestner, 1992; Schultheiss & Brunstein, 2005), that fulfilled the same criteria we applied to our selection of affiliation motive measures (i.e., coding systems based on experimental arousal studies; assessment via picture-story and content-coding methods). We included these measures in our meta-analysis primarily because they helped us determine whether observed gender differences are specific to affiliation motivation measures within a given sample or extend to other motive measures, too. In the latter case, substantial sex differences could be interpreted as a general gender bias of picture-story based assessment methods and not as a phenomenon specifically associated with affiliation motivation measures. A secondary benefit of including measures of n Power and n Achievement is, of course, that this provides for the first time quantifiable effect size information for potential sex differences in these motives, too. However, given the absence of such differences in the large-sample studies by Schultheiss and Brunstein (2001) and Pang and Schultheiss (2005), we did not expect to find substantial sex differences for these motive measures.

2. Method

Our methodological approach in this meta-analysis is based on Lipsey and Wilson (2001), except where indicated. Fig. 1 provides an overview of the literature search and coding process.

2.1. Literature search

To retrieve literature relevant to our topic, we performed a detailed search in the psychological database PsyCINFO. As only few studies have explicitly dealt with gender differences in implicit affiliation motivation so far, we opted for a broad and method-oriented search to detect as many relevant articles as possible. Accordingly, we developed a search term that combined terms representing different methods for the assessment of implicit motives and terms representing the entire family of affiliation motive measures. The resulting term was: “(motivation OR imaginative stor OR intimacy need OR n Affiliation OR n Affiliation-Intimacy) AND (PSE OR picture-story exercise OR content coding OR thematic apperception test OR PSE OR TAT OR PSE OR thematic analysis OR thematic apperception OR thematic apperception test OR content coding OR ‘Thematic content OR projective personality measur”) AND (affiliation OR intimacy OR affiliation-intimacy)”.

The search was performed on September 22nd 2012 and produced 2869 hits. To ensure the quality of the search, the result list was compared to a literature list of thirteen publications that dealt with gender differences in implicit affiliation motivation and were therefore relevant for the meta-analysis. The list was created by the second author, who is well-acquainted with the literature in this area of research. As all publications of this literature list were part of the search results, it can be expected that the search term was sufficiently appropriate to detect relevant literature.

Because it was economically impossible to screen all 2869 detected articles for relevance, we refined our result list by performing two additional searches within the meta-analytic database. First, we selected those articles that had used the described assessment methods or explicitly dealt with gender differences by using the search term: “TAT OR PSE OR picture story exercise OR picture-story exercise OR content coding OR thematic apperception OR gender OR sex”.

Secondly, we tried to identify studies that had actually used a measure of implicit affiliation motivation by using the search term: “affiliation motivation OR intimacy motivation OR affiliation-intimacy OR need affiliation OR need intimacy OR affiliation need OR intimacy need OR n affiliation OR n intimacy OR implicit motives”.

This procedure reduced the data pool to 828 entries. All publications of our previously created literature list were retained in these entries.\footnote{Central tendency measures (mean, median) of participant age did not significantly moderate the effect size of the gender difference in affiliation motivation measures.}
To avoid a possible publication bias, we also searched for unpublished papers by sending a request to the electronic mailing list of the Society for Personality and Social Psychology, asking for relevant gray papers. In response to this, we received three articles by the first author. A coding manual was written to ensure coding procedure and were coded regarding their study characteristics and result sections of each study. 28 studies remained after this paper request were rechecked for relevance focusing on method quality. The development of the coding manual included items to code study descriptors (e.g. methods, samples, context) as well as effect size information (empirical findings of a study). Inter-rater agreement was very high for both nominal data (98% average agreement) and interval-scaled data ($r = 0.998$) in a sample of eight studies.

2.2. Inclusion criteria for study eligibility

Title, abstract, and keywords of each of the remaining 828 studies were scanned individually for their relevance by the first author. The following criteria were applied to determine the eligibility of each study for inclusion in the meta-analysis:

1. Studies must report at least one measure of implicit affiliation motivation as defined in the introduction. Based on this criterion, 349 studies were excluded from further analyses.
2. Study samples must consist of men and women to allow the calculation of a gender difference effect size. 124 studies did not fulfill this criterion.
3. The analysis was not limited to a certain time period, but a natural historical onset was provided by the publication of the first studies based on the implicit motive assessment technique in the late 1940s (Schultheiss & Brunstein, 2010). Eight studies that were published before 1940 were therefore excluded from the analysis.
4. Only articles written in English or German were included. Eighteen articles did not meet this criterion.
5. To avoid duplication, studies were excluded if they had already been published in other works included in the meta-analysis ($n = 5$).

Some articles did not provide enough information to ensure their relevance ($n = 187$), did not show any thematic relevance at all ($n = 24$), or could not be procured ($n = 9$) and were therefore also excluded from the analysis.

2.3. Coding of study characteristics

Before coding, all remaining studies and the results of our gray paper request were rechecked for relevance focusing on method and result sections of each study. 28 studies remained after this procedure and were coded regarding their study characteristics by the first author. A coding manual was written to ensure coding quality. The development of the coding manual included items to code study descriptors (e.g. methods, samples, context) as well as effect size information (empirical findings of a study). Inter-rater agreement was very high for both nominal data (98% average agreement) and interval-scaled data ($r = 0.998$) in a sample of eight studies.

2.4. Computation of effect size

We used Cohen’s $d$, the standardized mean difference, as effect size measure. Effect sizes were calculated per study, quantifying the mean difference in implicit affiliation motivation scores between men and women. If a study also reported findings for n Achievement and n Power, we additionally calculated effect sizes for these measures. If studies failed to report means and standard deviations but provided other statistical data (e.g. $t$ values, $F$ ratios), we used these to estimate the effect size (see Lipsey & Wilson, 2001, pp. 198–200).

Some articles reported statistical data of multiple independent samples. In this case we defined each sample to be an individual study and calculated an effect size for each of them. If a study reported various outcome variables (e.g. measure of need for intimacy and measure of need for affiliation) concerning the same sample, we coded effect sizes separately for each measure and then averaged them into a single study effect size. The same procedure was applied if a study reported its results separately for several subgroups (Hunter & Schmidt, 2004).

2.5. Correction of effect sizes

Because all samples sizes were larger than 20 participants, there was no need to adjust effect sizes for small sample bias. Following Hunter and Schmidt (2004), we corrected individual effect sizes for attenuation due to unreliability of the implicit measure. The adjusted effect size ($d'$) results from the division of the observed effect size by the square root of the reliability of the implicit measure. Unfortunately, information about reliability was available only sporadically: Merely 39% of the studies ($n = 12$) reported useful reliability coefficients. Therefore it was not possible to adjust each individual effect size. So we first did a bare-bones meta-analysis and then corrected the effect sizes from all studies for attenuation with the help of an average reliability coefficient ($r_{yy} = 0.89$) from the data available.
2.6. Integration and statistical analyses

To integrate effect sizes we applied a random-effects model, which is based on the assumption that the meta-analyzed studies do not come from the same population and therefore do not share the same true effect size (Borenstein, Hedges, Higgins, & Rothstein, 2009). We decided to apply this approach because (a) it is an alternative, fixed-effects model is based on the strong assumption that all effect sizes come from the same population, a notion that is often viewed critically by statisticians (see Lipsey & Wilson, 2001), (b) sample size and motive measurements varied across the primary studies, and (c) all studies had been performed by independent researchers and laboratories and in more than one country.

We followed Borenstein et al. (2009) and computed the within-study variance (\(V\)) as well as the between-study variance (\(T^2\)) for the random-effects model. All individual effect sizes were then weighted by the inverse of its within-study variance plus the between-study variance (\(w = \frac{1}{V + T^2}\)). The weighted mean effect size (\(d\)) was computed by dividing the sum of the weighted effect sizes by the sum of the weights.

Because all potential moderators were categorical, we ran moderator analyses as subgroup analyses with z-tests (Borenstein et al., 2009), testing the moderators coding system, picture cues, and verbal fluency. Concerning coding systems, we aggregated studies that had used coding manuals for assessing n Affiliation (Heyns, Veroff, & Atkinson, 1958; Shipley & Veroff, 1958), n Affiliation-Intimacy (Winter, 1991), and affiliative trust-mistrust (McAdams, 1980) into one category, as they all measure hope and fear aspects of affiliation motivation. This category was contrasted with studies that had used coding systems for the assessment of n Intimacy (McAdams, 1980). To examine the possible influence of picture cues, we compared studies that had used gender-congruent picture sets for men and women with those that had used the same picture sets for both. Lastly we contrasted studies that reported only raw motive imagery scores with studies that had corrected motive scores for PSE protocol length to investigate the influence of women’s longer stories on the gender difference in affiliation motivation.

3. Results

3.1. Outliers

Prior to data analysis, we used box plots to identify three effect size outliers for affiliation motivation (McKay, 1987; Schroth, 1985; Touliatos & Lindholm, 1975) and two for power motivation (Chusmir, 1983; Schroth, 1985). We explored how our results would be affected by recoding outlier effect sizes to those of the closest non-outlier of the effect size distribution (i.e., winsorizing). For the sake of completeness, we performed all analyses with the original datasets (including outliers) and the winsorized datasets. As the estimated original and winsORIZED mean effect sizes did not differ significantly from each other (\(z_{\text{diff}} = 0.40, p = 0.66\) for affiliation; \(z_{\text{diff}} = 1.12, p = 0.13\) for power) we will present only the results of the original datasets with outlier inclusion in the following.

3.2. Descriptive characteristics

A total of 33 independent effect sizes for gender differences in affiliation motivation scores retrieved from 26 publications and 2 unpublished datasets were integrated into the meta-analysis. Fifteen studies also reported gender differences for n Power scores and 13 studies reported gender differences for n Achievement scores. Studies covered the publication period from 1964 until 2009. The total sample included 5962 individuals (3439 women; 58%). Table 1 gives an overview of all studies included in the meta-analysis. The mean of the uncorrected and unweighted affiliation effect sizes, \(d = 0.48, SD = 0.30\), differed significantly from zero, \(t(32) = 9.03, p < 0.001\).

3.3. Overall effect sizes and statistical significance

Gender differences in affiliation motivation measures were of medium size (\(d = 0.45\)) and highly significant (\(p < 0.001\); see Table 2), with the positive sign indicating higher scores for women, compared to men. After correction for attenuation due to measurement error, the average effect size increased to \(d = 0.48\) (\(SE = 0.04\)). In contrast to affiliation motivation, the average effect size for measures of power motivation (\(d = -0.19\)) and achievement motivation (\(d = 0.14\)) were small and non-significant (\(p > 0.05\)). Effect sizes for power and achievement both differed significantly from affiliation, \(z_{\text{diff}} = 4.82, p < 0.001\) and \(z_{\text{diff}} = 3.51, p < 0.001\), respectively.

3.4. Analysis of heterogeneity

Affiliation study effect sizes ranged from \(-0.06\) to 1.35. The Q-test produced a significant result (Table 2), with 41% true heterogeneity that was not caused by sampling error (Huedo-Medina, Sanchez-Meca, Martin-Martinez, & Botella, 2006).

3.5. Moderator analyses

In most studies (\(n = 27\)), coding systems that measured both facets of the affiliation motive were used. Nine studies reported measures of n Intimacy. Three studies reported measures for both, affiliation and intimacy motivation. In this case we only considered intimacy-related measures, as there were fewer data for this motive and the usage of both values would have led to the problem of dependent samples. As shown in Table 3, there was a significant effect for both subgroups, without any evidence that they were significantly different from each other. Thus, coding systems that captured hope and fear components of affiliation motivation did not differ in terms of gender differences from n Intimacy, that is, a coding system targeting primarily the hope aspect.

Secondly, we investigated whether the use of gender-congruent picture cues versus unitary picture cues had an effect on gender differences in affiliation motivation. Because only three studies reported the use of gender-congruent cues, the calculation of the between-study variance (\(T^2\)) for the subgroup would have been very imprecise. We therefore performed the moderator analysis with a pooled \(T^2\) across both categories (see Borenstein et al., 2009). Studies with unitary picture cues produced a slightly, but not significantly larger effect size than those with gender-congruent picture cues (see Table 3). Hence, the gender-specificity of picture sets had no discernible effect on gender differences in affiliation motivation measures.

A moderator analysis of affiliation motivation scores corrected for protocol length versus uncorrected scores showed that both categories produced significant effect sizes, but the difference between them was not significant (Table 3). Thus, correction for PSE protocol length does not attenuate the gender difference in affiliation motivation measures.

3.6. Publication bias

Meta-analyses can be affected by the file-drawer problem (Rosenthal, 1979), with published studies reporting higher effect
sizes than unpublished findings. Due to selective inclusion of the former in meta-analyses, this may cause an upward bias in the overall effect size estimate (Borenstein et al., 2009). In order to address this problem, we examined the presence of a potential publication bias with a funnel plot, comparing effect sizes and their standard errors (Fig. 2). An asymmetric scattering of data points

Table 1
Sample size and uncorrected, unweighted effect size information of all studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>No.</th>
<th>Reference</th>
<th>N</th>
<th>dAch (SE)</th>
<th>dPow (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Carroll (1987)</td>
<td>65</td>
<td>0.53 (0.26)</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>Chusmir (1983)a</td>
<td>124</td>
<td>0.06 (0.18)</td>
<td>0.73 (0.19)</td>
</tr>
<tr>
<td>3A</td>
<td>Craig (1996)b</td>
<td>48</td>
<td>0.60 (0.31)</td>
<td>-</td>
</tr>
<tr>
<td>3B</td>
<td>Craig (1996)b</td>
<td>162</td>
<td>0.36 (0.16)</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>Daugherty, Kurtz, and Phebus (2009)</td>
<td>120</td>
<td>0.00 (0.19)</td>
<td>0.26 (0.18)</td>
</tr>
<tr>
<td>5</td>
<td>Dember (1964)</td>
<td>44</td>
<td>0.70 (0.31)</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>Hien, Haas, and Cook (1998)</td>
<td>34</td>
<td>0.06 (0.35)</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>Jacob (1977)</td>
<td>97</td>
<td>0.71 (0.25)</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>James, Lewkowicz, Llibaber, and Lachman (1995)</td>
<td>150</td>
<td>0.35 (0.17)</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>Karabenick (1977)a</td>
<td>131</td>
<td>0.36 (0.21)</td>
<td>-0.03 (0.30)</td>
</tr>
<tr>
<td>10</td>
<td>King (1995)b</td>
<td>101</td>
<td>0.56 (0.23)</td>
<td>-0.21 (0.22)</td>
</tr>
<tr>
<td>11A</td>
<td>Langan-Fox and Grant (2006)</td>
<td>334</td>
<td>0.44 (0.13)</td>
<td>0.32 (0.13)</td>
</tr>
<tr>
<td>11B</td>
<td>Langan-Fox and Grant (2006)</td>
<td>213</td>
<td>0.15 (0.14)</td>
<td>-0.13 (0.14)</td>
</tr>
<tr>
<td>12</td>
<td>Mazur (1989)b</td>
<td>68</td>
<td>0.83 (0.25)</td>
<td>-</td>
</tr>
<tr>
<td>13A</td>
<td>McAdams et al. (1988)</td>
<td>153</td>
<td>0.60 (0.17)</td>
<td>-</td>
</tr>
<tr>
<td>13B</td>
<td>McAdams et al. (1988)</td>
<td>1317</td>
<td>0.41 (0.06)</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>McAdams, Rothman, and Lichter (1982)</td>
<td>160</td>
<td>0.22 (0.19)</td>
<td>0.54 (0.19)</td>
</tr>
<tr>
<td>15</td>
<td>McAuley (2002)b</td>
<td>85</td>
<td>0.20 (0.22)</td>
<td>0.27 (0.22)</td>
</tr>
<tr>
<td>16A</td>
<td>McKay (1987)</td>
<td>31</td>
<td>0.52 (0.36)</td>
<td>-</td>
</tr>
<tr>
<td>16B</td>
<td>McKay (1987)</td>
<td>32</td>
<td>-0.06 (0.36)</td>
<td>-</td>
</tr>
<tr>
<td>16C</td>
<td>McKay (1987)</td>
<td>71</td>
<td>0.70 (0.25)</td>
<td>-</td>
</tr>
<tr>
<td>17</td>
<td>Pang and Schultheiss (2005)b</td>
<td>320</td>
<td>0.57 (0.11)</td>
<td>0.27 (0.11)</td>
</tr>
<tr>
<td>18</td>
<td>Schroth (1979)</td>
<td>80</td>
<td>0.63 (0.23)</td>
<td>-</td>
</tr>
<tr>
<td>19</td>
<td>Schroth (1985)</td>
<td>90</td>
<td>1.29 (0.23)</td>
<td>-0.51 (0.21)</td>
</tr>
<tr>
<td>20</td>
<td>Schultheiss and Brunstein (2001)b</td>
<td>428</td>
<td>0.51 (0.10)</td>
<td>0.09 (0.10)</td>
</tr>
<tr>
<td>21</td>
<td>Schultheiss, Dargel, and Rohde (2003b)b</td>
<td>54</td>
<td>0.39 (0.29)</td>
<td>-</td>
</tr>
<tr>
<td>22</td>
<td>Schultheiss, Pang, Torges, Wirth, and Treynor (2005)b</td>
<td>216</td>
<td>0.38 (0.14)</td>
<td>-</td>
</tr>
<tr>
<td>23</td>
<td>Sorrentino, Ye, Szeto, et al. (n.d.)</td>
<td>679</td>
<td>0.34 (0.08)</td>
<td>-</td>
</tr>
<tr>
<td>24</td>
<td>Sorrentino, Ye, Wilson, et al. (n.d.)</td>
<td>181</td>
<td>0.36 (0.17)</td>
<td>-</td>
</tr>
<tr>
<td>25</td>
<td>Touliatos and Lindholm (1975)</td>
<td>64</td>
<td>1.35 (0.28)</td>
<td>-0.09 (0.26)</td>
</tr>
<tr>
<td>26</td>
<td>Winter and Wiecking (1971)</td>
<td>65</td>
<td>0.00 (0.25)</td>
<td>0.06 (0.25)</td>
</tr>
<tr>
<td>27</td>
<td>Wirth and Schultheiss (2000)b</td>
<td>87</td>
<td>0.50 (0.22)</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>Zurbriggen (2000)b</td>
<td>158</td>
<td>0.39 (0.16)</td>
<td>-</td>
</tr>
<tr>
<td>29</td>
<td>Average</td>
<td>181</td>
<td>0.48 (0.21)</td>
<td>0.12 (0.19)</td>
</tr>
</tbody>
</table>

a Usage of gender-congruent cues.
b Correction for verbal fluency.
c Effect size is based on a measure of need for intimacy.
d Effect size represents the average of two separate study effect sizes.

Table 2
Overall meta-analytic results.

<table>
<thead>
<tr>
<th>Motive</th>
<th>N</th>
<th>k</th>
<th>d</th>
<th>SD</th>
<th>SE</th>
<th>CI 95%</th>
<th>z</th>
<th>Q</th>
<th>I² (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affiliation</td>
<td>5962</td>
<td>33</td>
<td>0.45</td>
<td>0.13</td>
<td>0.04</td>
<td>[0.37; 0.53]</td>
<td>11.25***</td>
<td>54.32***</td>
<td>41</td>
</tr>
<tr>
<td>Power</td>
<td>2493</td>
<td>15</td>
<td>-0.19</td>
<td>0.45</td>
<td>0.13</td>
<td>[-0.44; 0.05]</td>
<td>1.53</td>
<td>115.12***</td>
<td>88</td>
</tr>
<tr>
<td>Achievement</td>
<td>2235</td>
<td>13</td>
<td>0.14</td>
<td>0.23</td>
<td>0.08</td>
<td>[-0.03; 0.30]</td>
<td>1.63</td>
<td>35.84***</td>
<td>67</td>
</tr>
</tbody>
</table>

Note. Asterisks signify that weighting was applied.

Table 3
Results of moderator analyses.

<table>
<thead>
<tr>
<th>Moderator</th>
<th>k</th>
<th>d</th>
<th>SD</th>
<th>SE</th>
<th>CI 95%</th>
<th>zog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coding system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dual-facet affiliation</td>
<td>24</td>
<td>0.49</td>
<td>0.18</td>
<td>0.05</td>
<td>[0.38; 0.60]</td>
<td>1.15</td>
</tr>
<tr>
<td>n Intimacy</td>
<td>9</td>
<td>0.41</td>
<td>0.14</td>
<td>0.04</td>
<td>[0.32; 0.49]</td>
<td></td>
</tr>
<tr>
<td>Cues</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender-congruent</td>
<td>3</td>
<td>0.35</td>
<td>0.14</td>
<td>0.15</td>
<td>[0.06; 0.64]</td>
<td>0.71</td>
</tr>
<tr>
<td>Unitary</td>
<td>30</td>
<td>0.46</td>
<td>0.14</td>
<td>0.04</td>
<td>[0.37; 0.55]</td>
<td></td>
</tr>
<tr>
<td>PSE protocol length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected</td>
<td>10</td>
<td>0.46</td>
<td>0.18</td>
<td>0.05</td>
<td>[0.37; 0.56]</td>
<td>0.05</td>
</tr>
<tr>
<td>Raw scores</td>
<td>23</td>
<td>0.46</td>
<td>0.18</td>
<td>0.06</td>
<td>[0.35; 0.57]</td>
<td></td>
</tr>
</tbody>
</table>

Note. Asterisks signify that weighting was applied.
would indicate the presence of a bias (Borenstein et al., 2009; Light & Pillemer, 1984). Although the interpretation of a funnel plot is rather subjective, we suggest that the pattern we obtained is symmetric: large studies appeared towards the top of the graph and smaller studies towards the bottom of the diagram, spreading out across a broader range of values around the mean effect size. The funnel plot clearly shows two outliers on the right-hand side, but no conspicuous systematic absence of effect sizes on the lower left-hand side.

4. Discussion

In this meta-analysis we investigated the assumption that women score higher than men in measures of the implicit affiliation motive. In order to exclude alternative explanations for a potential gender difference, we performed several moderator analyses, focusing on methodological aspects of the primary studies.

4.1. Summary of the results

As predicted, the meta-analysis supports the presence of a gender difference in measures of implicit affiliation motivation. The effect, which differed significantly from zero, is of medium size according to Cohen (1992) and suggests that women have a greater need for affiliation than men. As expected, the gender difference was limited to measures of affiliation motivation – men and women did not differ significantly on measures of n Power or n Achievement. These effects remained essentially unchanged when we Winsorized outliers.

To explore to what extent the gender difference in measures of affiliation motivation was due to methodological differences of the primary studies, we examined three potential moderators concerning the assessment of affiliation motivation. First, we differentiated between coding instruments that assess both hope and fear facets of motivation and personality, and cues with rather clear-cut and prescriptive content are usually avoided, because this reduces the variance and validity of test scores (see Gieser & Stein, 1999; Schultheiss & Pang, 2007). As a consequence, the majority of the cues used in the studies included in this meta-analysis were also ambiguous with regard to the gender roles they presented. For instance, one widely used picture shows a woman and a man performing a trapeze act. If anything, this picture shows both in non-traditional roles. Another frequently used picture shows a woman and a man from the back on a park bench near a bridge. Apart from the mere fact that representatives of both genders sitting side by side are depicted, it is not clear in what sense this picture shows people in traditional gender roles. Yet both pictures frequently yield higher scores on measures of affiliation motivation for women than for men (Pang & Schultheiss, 2005; Schultheiss & Brunstein, 2001).

Finally, PSE protocol length did not turn out to be a significant moderator of the affiliation-related gender difference, either. This is particularly notable, because we consistently write longer stories on the PSE (Pang & Schultheiss, 2005; Schultheiss & Brunstein, 2001), and overall protocol length is frequently positively associated with motive imagery raw scores (Schultheiss & Pang, 2007). The observation that the gender difference in measures of affiliation motivation persisted more or less unchanged in studies that used length-corrected measures suggests that the gender difference in narrative fluency may have a different source than the gender difference in affiliation motivation.

4.2. Possible explanations

We were able to establish that a gender difference in affiliation motivation exists and that it cannot be attributed to some of the methodological factors most closely associated with the measurement of affiliation motivation. So what might account for our Duncan and Peterson's (2010) concern about picture set differences being to blame for observed gender differences in affiliation motive scores is not supported by the empirical record. However, this conclusion is only tentative at this point, because only three of the included studies reported using same-sex stimuli. The data basis may therefore be too small to arrive at a clear-cut conclusion about the role of gender-congruent versus general picture sets.\footnote{Across the studies used in this meta-analysis, researchers have used a rather diverse array of story-eliciting cues, ranging from original TAT pictures (e.g., Touliatos & Lindholm, 1975) to picture sets custom-tailored to the assessment of implicit affiliation motivation (e.g., Wirth & Schultheiss, 2006) to picture sets aimed at assessing several motives at once (e.g., Zurbriggen, 2000). In addition to the lack of a moderating effect of gender-congruence of the picture sets, this heterogeneity makes it rather unlikely that gender differences are based on a certain picture or set of pictures.}

Although our findings suggest that the use of gender-congruent picture sets cannot be held accountable for the gender difference in affiliation motivation scores, one could argue that Duncan and Peterson's (2010) concern is still valid at a more general level. Perhaps some stimulus sets just happen to feature cues that elicit more affiliation-related imagery in women than in men, possibly due to gender-specific learning processes (see Schultheiss & Schultheiss, 2014, for a possible mechanism). If other pictures had been used, this difference might not emerge or might even be reversed. However, we think this explanation is rather unlikely for two reasons. First, the studies included in this meta-analysis employed many different sets of cues. If only some picture sets or verbal-cue sets had elicited the gender difference, but not others, we would not have observed the gender difference with such consistency (see Table 1). Second, the cues typically used in implicit motive assessment are intentionally rather ambiguous in terms of their content to allow test takers free expression of their motivational needs and personality, and cues with rather clear-cut and prescriptive content are usually avoided, because this reduces the variance and validity of test scores (see Gieser & Stein, 1999; Schultheiss & Pang, 2007). As a consequence, the majority of the cues used in the studies included in this meta-analysis were also ambiguous with regard to the gender roles they presented. For instance, one widely used picture shows a woman and a man performing a trapeze act. If anything, this picture shows both in non-traditional roles. Another frequently used picture shows a woman and a man from the back on a park bench near a bridge. Apart from the mere fact that representatives of both genders sitting side by side are depicted, it is not clear in what sense this picture shows people in traditional gender roles. Yet both pictures frequently yield higher scores on measures of affiliation motivation for women than for men (Pang & Schultheiss, 2005; Schultheiss & Brunstein, 2001).
observation of women having higher scores on implicit measures of affiliation motivation than men? In the following, we focus on two possible explanations, one social-psychological and one psychoendocrinological.

From a social-psychological and feminist perspective, one could interpret our findings as reflecting the impact of traditional gender role stereotypes, with women being expected to be more socially attentive and affiliative than men (e.g., North & Fiske, 2014; Ridgeway & Correll, 2004). Focusing on the flip side of gender differences in narrative measures of affiliation motivation, Gilligan (1982) suggested that men score lower because they fear intimacy (for a related argument regarding the meaning of low motive scores, see Schultheiss, 2008). Despite critiques of these social-psychological perspectives on gender differences in affiliative orientation (e.g., Colby & Damon, 1982; McAdams et al., 1988), gender roles are frequently portrayed as important determinants of human behavior (Eagly, 1987; McClelland, 1975; North & Fiske, 2014) and so it might be possible that women are better socialized to express affiliative needs than men, be it in their imaginative stories or in their daily lives (Duncan & Peterson, 2010). Furthermore in many societies women often have a lower social status than men. Hence, for their own safety and survival they need to be socially attentive and avoid interpersonal conflicts (Duncan & Peterson, 2010; North & Fiske, 2014). These diverse, but conceptually related accounts may explain why women do not just portray themselves as more affiliation-oriented explicitly (Feingold, 1994), but are also more affiliation-motivated at the implicit level.

A second possible explanation for the affiliation-related gender difference is based on research on the psychoendocrinology of affiliation motivation. Examining the association between sex hormones and affiliation motivation, Schultheiss, Dargel, and Rohde (2003a) observed that in normally cycling women, progesterone levels and affiliation motivation were positively correlated, with increasing progesterone preceding increased affiliation motivation levels across the menstrual cycle. Similarly, Wirth and Schultheiss (2006) reported increases in progesterone to be associated with increases in affiliation motivation in a mixed-gender sample. Schultheiss et al. (2003a) also made the surprising observation, recently replicated by Schultheiss and Zimni (2015), that women using oral contraceptives, which typically contain progesterone-like substances, had higher affiliation motive scores than women who did not use oral contraceptives or men, with the latter two showing little difference. Two conclusions can be drawn from these findings. One, high levels of progesterone are associated with, and may perhaps even be a cause of, heightened affiliation motivation. Two, the gender difference in affiliation motivation measures we observed in our present meta-analysis may therefore be due to (a) women taking oral contraceptives and (b) women tested in the high-progesterone phase of their menstrual cycles, when their progesterone levels are higher than those of men. Future studies and meta-analyses could explore whether gender- or medication-related differences progesterone are indeed behind the gender difference we observed in our data.

To summarize, we think that the observed gender difference in measures of implicit affiliation motivation is unlikely to arise for purely methodological reasons. Instead, we see more promise in explaining them based on gender differences in social roles and socialization and in natural (i.e., menstrual-cycle related) and elective (i.e., oral contraceptive use) gender differences in hormone levels. Of course, the latter two perspectives are not mutually exclusive, and there may be other theories that may also account for our findings, such as evolutionary theories (e.g., Taylor, 2006). However, we think our results carry particular weight in informing whatever substantial theory of gender differences in behavior one endorses, because unlike meta-analytic findings that are based on questionnaire-based need measures (e.g., Feingold, 1994) and thus susceptible to whatever gender stereotypes people may believe in, the measures of affiliation motivation we meta-analyzed tap into individuals’ nonconsciously operating affective preferences and cannot be equated with findings based on self-report (see Kölner & Schultheiss, 2014). Precisely because implicit motivational needs may be more difficult to control or change than salient explicit beliefs (see McClelland et al., 1989), we think future research needs to carefully dissect the causes and consequences of the sizeable gender difference in implicit affiliation motivation we observed.

### 4.3. Limitations

Although the medium-sized effect revealed in this meta-analysis helps to clarify whether men and women differ in their implicit need for affiliation, we note some limitations of our study. As only few studies have focused on gender differences in affiliation motivation so far, we had to extend our literature search to all studies dealing with the implicit need for affiliation. For economic reasons it was not possible to examine every study harvested in this manner in detail, and we may therefore have missed some relevant papers. However, we see no compelling reason to assume that there was any specific bias in which studies we ended up missing, and we therefore suggest that the findings we report here are representative for the entire pool of studies reporting gender differences in measures of implicit affiliation motivation.

Furthermore, many studies that fulfilled our eligibility criteria failed to report sufficient statistical data (such as SDs) and had to be excluded from the data pool as a consequence. Additionally, most doctoral theses identified in our search could not be procured. It is therefore not surprising that only a relatively small number of studies could be integrated into our meta-analysis. However, in our opinion the data set nevertheless had a sufficient size in terms of studies and samples included as well as overall N, and conducting a meta-analysis was therefore fully justified.

Finally, one potential danger for any kind of meta-analysis is that it refines an effect that only exists due to the selective reporting of (false) positives (see, for instance, Carter & McCullough, 2014). So may our core meta-analytic finding simply reflect the selective reporting of gender differences in affiliation motive measures? Although some of the articles that we included in this meta-analysis directly focused on gender differences in affiliation motivation (i.e., Chusmir, 1983; Hien et al., 1998; James et al., 1995; Karabenick, 1977; Mazur, 1989; McAdams et al., 1988), the effect sizes we found for these studies covered the entire spectrum of our overall pool of study effects, not just the upper range. Moreover, the majority of studies tested other hypotheses and the reporting of gender differences was incidental. In other words, we were not in pursuit of a hot hypothesis that has galvanized the field and produced dozens of studies trying to demonstrate a gender difference in implicit affiliation motivation. The two studies that produced the largest effect sizes and were identified as outliers were not among those whose specific hypothesis concerned gender differences in affiliation motivation. We therefore suggest that the core finding of our meta-analysis is not an artifact of selective reporting of gender differences in the published or unpublished literature.

### 4.4. Conclusion

To conclude, using meta-analytic techniques, we were able to document a robust, medium-sized gender difference in measures of implicit affiliation motivation. This effect could not be observed for measures of other motivational needs and was unaffected by the gender congruence of story-eliciting cues, coding systems applied, or protocol length. We were thus able to confirm
Duncan and Peterson’s (2010) hypothesis that women generally score higher on measures of implicit affiliation motivation than men.

Acknowledgment

We gratefully acknowledge the helpful input of Abigail J. Stewart to this manuscript.

References

References marked with an asterisk indicate studies included in the meta-analysis


