Association Between Insight and Outcome of Psychotherapy: Systematic Review and Meta-Analysis

Simone Jennissen, M.Sc., Julia Huber, M.Sc., Johannes C. Ehrenthal, Ph.D., Henning Schauenburg, M.D., Ulrike Dinger, D.Sc., M.D.

Objective: An increased understanding of repetitive dysfunctional patterns and their relationship to an individual’s life history is regarded as a key mechanism of change in insight-oriented therapies. At the same time, empirical research on the insight-outcome relationship is rare, and its generalizability is restricted by the use of a wide range of definitions and methods among studies. The authors conducted a meta-analysis to systematically examine the association between patient insight and psychotherapy outcome across a range of treatment modalities.

Method: Insight was defined as patients’ understanding of associations between past and present experiences, typical relationship patterns, and the relation between interpersonal challenges, emotional experience, and psychological symptoms. From 13,849 initially identified abstracts, the authors extracted 23 independent effect sizes. A random-effects meta-analysis was performed to assess the magnitude of the insight-outcome relationship. Risk of publication bias was assessed with funnel plot inspections, Egger’s regression test, and Duval and Tweedie’s trim-and-fill procedure as sensitivity analyses.

Results: A significant, moderate correlation (r=0.31) was observed between insight and treatment outcome. Sensitivity analyses demonstrated the robustness of the results.

Conclusions: The findings support the importance of insight for psychotherapy outcome. Insight may be a relevant mechanism of change across different treatment modalities.

Despite compelling evidence for the efficacy of psychotherapy for mental disorders, there is an ongoing debate about underlying mechanisms of change (1, 2). Hence, process-outcome research is a key area of contemporary psychotherapy research (3–6), and understanding the mechanisms that underlie psychotherapy effects is vital to increase the effectiveness of psychosocial treatments (3, 7).

One of the oldest proposed curative mechanisms of psychotherapy is insight. Freud and Breuer (8) assumed that neurotic symptoms would disappear if patients became aware of split-off and repressed memories. Later developments in ego psychology stressed that insight must be understood as a process that involves both cognitive and affective components (9). Contemporary psychodynamic theorists point toward the importance of relational aspects of seeking, constructing, and conversing about insight (10–14). Insight is closely related to psychodynamic interventions, such as clarification, confrontation, and interpretation (15). According to psychodynamic theory, higher levels of self-understanding result in fewer negative automatic reactions to stress and challenges, more positive emotional experiences, and greater freedom to choose adaptive interpersonal and health-related behaviors (16–20).

A fundamental challenge to the study of insight is that the term has no clear definition (21). In psychiatry, it often denotes insight into illness as a prerequisite for treatment motivation (22). Experiential models emphasize attaining a new perspective through experiencing (23). In cognitive-behavioral therapy (CBT), insight relates to becoming aware of automatic negative thoughts (24). Within these definitions, insight refers to a process, a state, or both. As a trait characteristic, it is often referred to as insightfulness (25). Our definition of insight is based on the psychodynamic conceptualizations outlined above. Throughout this article, the terms insight and self-understanding are used interchangeably to refer to patients’ understanding of associations between past and present experiences, typical relationship patterns, and the relation between interpersonal challenges, emotional experience, and psychological symptoms (2, 26).

Although a focus on insight as a curative factor is most prevalent within psychodynamic treatments, other types of psychotherapy, such as CBT, enhance patients’ self-understanding, for example, by targeting automatic thought processes. To gain a comprehensive image of the association between insight and outcome, we use the term psychotherapy.

See related features: Editorial by Dr. Høglend (p. 923), CME course (p. 1047), and AJP Audio (online).
to refer to those primarily verbal psychosocial treatments in which self-understanding is encouraged. An important, although nonsystematic, first overview of the literature by Crits-Christoph et al. (2) suggests that self-understanding is related to psychotherapy outcome and that this relationship may be specific to psychodynamic treatments. A systematic review of studies on the insight-outcome association would allow a comprehensive evaluation of the current evidence base, an estimation of the population effect size, and identification of topics for future research. We therefore conducted the present systematic review with two overall aims. First, we comprehensively reviewed the literature to collect empirical evidence on the insight-outcome association and assess the validity of the findings. Second, using meta-analysis, we quantitatively combined the results of previous studies to estimate the mean population effect size for the insight-outcome association.

**METHOD**

This meta-analysis was conducted using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The review protocol was registered in advance at the international prospective register of systematic reviews (PROSPERO), registration number CRD42016043104.

**Eligibility Criteria**

**Population.** Studies had to be based on adult patients (age 18 or older) who were seeking treatment for a psychological condition. One of the studies otherwise identified as eligible employed a subsample (15%) of adolescents (27). Since the proportion of adolescents was small, we deemed the threats to validity negligible and included the study. Excluding this study from the analyses did not change the pattern of results. We had no requirements regarding diagnoses, but we excluded studies that were based on volunteer samples without any psychological burden.

**Interventions.** Participants had to receive psychological treatments such as psychotherapy or counseling. Treatments had to be delivered by trained therapists, such as psychiatrists, clinical psychologists, or counselors. We included all bona fide psychotherapies (28) that were conducted in a dyadic or group therapy setting. We did not include interventions such as dream interpretation, hypnosis, bibliotherapy, and music or art therapy unless they were integrated into a broader psychological treatment.

**Predictor and outcome variables.** As a predictor variable, insight had to be measured by a quantitative instrument that fit with our definition. Assessment had to take place during treatment or at termination. We did not include studies employing only a pretreatment measure of trait insightfulness. As outcome variables, we included any quantitative outcome measure such as symptom improvement, quality of life, or psychological functioning. We accepted session and overall treatment outcomes but required that outcome be measured simultaneously with or, preferably, subsequent to the assessment of insight.

**Study designs and publications.** We included empirical studies that reported quantitative group statistics. Studies could be published as journal articles, books, or dissertations.

**Search Strategy**

Publications were identified by a search of PubMed, PsycINFO, PsycARTICLES, and PSYNDEx using the search terms (insight OR insightfulness OR self-understanding) AND (psychotherapy OR counseling). Results were restricted to publications in English or German. The last search was conducted on April 6, 2018. In addition, reference lists of relevant publications were examined by hand search, and experts were consulted to identify missing studies.

**Study Selection**

A flow diagram of the study selection process is presented in Figure 1. Electronic database and hand search yielded a total of 13,849 records. After duplicate removal, the titles and abstracts of 12,780 publications were screened. Next, full texts were inspected. Unpublished dissertations were requested from authors. A total of 261 full-text publications were assessed for eligibility. Publications with unclear inclusion status were discussed between two authors (S.J., U.D.), and ambiguity was resolved by consensus. If two publications were based on the same data, we chose the study with the insight measure that fit our definition of insight best. When insight measures were identical, we selected the study that reported the most comprehensive statistics. We selected Andreoli et al. (29) over Andreoli et al. (30) because the latter only reported associations between insight and outcome for subsamples. Hoffart et al. (31) was chosen over Hoffart and Sexton (32) because only the former reported a zero-order correlation between insight and outcome. We selected Kolden (33) over Kolden (34) because the former reported the purer association between insight and outcome. We chose Luborsky et al. (35) over Grenyer and Luborsky (36) because the insight operationalization in the former fit our definition more closely. In cases of ambiguity about the primary data source, authors were contacted for clarification. This procedure resulted in 22 unique studies (21 peer-reviewed journal articles and one dissertation) representing a k of 23 effect sizes from independent samples that were included in the meta-analysis.

**Data Extraction Procedure**

We developed a data extraction template according to the preregistered study protocol, pretested it on five studies, and subsequently refined it. One author (S.J.) extracted the following data: study characteristics, sample characteristics, information on predictor and outcome measures, and effect sizes for the association between insight and outcome. If zero-order correlations could not be extracted or computed from the reported data, authors were contacted for
additional information. One randomized controlled trial reported data for two independent treatment subgroups (dynamic therapy and CBT) (37). In this case, each subgroup was treated as a separate study (38).

When correlations of several insight scores with outcome (39, 40), of insight with several outcomes (41), or all subscales of an outcome measure (42) were reported, we computed the mean correlation coefficient for that study. For studies reporting only partial information on the insight-outcome association (e.g., only for significant subscales), we imputed missing nonsignificant correlations with a correlation coefficient of zero before calculating a mean score (42, 43).

**FIGURE 1. Flow Diagram of the Study Selection Process in a Systematic Review and Meta-Analysis of Insight and Outcome of Psychotherapy**

Records identified through database searching (N=13,751)  
Records identified through other sources (N=98)  
Records screened on abstract level after duplicates removed (N=12,780)  
Records excluded (N=12,497)  
Full text not received (N=22)  
Full-text articles assessed for eligibility (N=261)  
Full-text articles excluded (N=236)  
• No quantitative study (N=116)  
• No clinical sample (N=11)  
• No adequate measurement (N=81)  
• No association reported (N=27)  
• Multiple publications based on the same study (N=4)  
Studies included in meta-analysis (N=22)

**Risk of Bias Assessment**

Two authors (S.J., J.H.) independently assessed the risk of bias in each primary study. For this assessment, we adapted the Systematic Assessment of Quality in Observational Research Scale (44) and included additional items to assess common artifacts in meta-analyses of correlations (45). The modified 16-item risk of bias scale covers four categories: sample, quality of measurements, quality of statistical analyses, and reporting of data (see the online supplement). Each category includes three to five items, which are evaluated with ratings of 0 (no risk of bias), 1 (definite or unclear risk of bias), or not applicable. A mean score of applicable items yields a ratio between 0 and 1. Higher scores represent more risk of bias. Interrater reliability for the risk of bias assessment was good (intraclass correlation coefficient [ICC] of 0.80 using the formula ICC[3,1]; see reference 46), and scores were averaged across raters.

**Data Analysis**

Effect sizes were measured as correlation coefficients. All effects were coded such that positive correlations indicate more insight to be associated with better outcome. All Pearson correlations were transformed with Fisher's r-to-z transformation prior to any computations and back-transformed for interpretation purposes. Analyses were performed using the R package metafor, version 1.9-9 (47). Alpha was set at 0.05 for all analyses. First, we estimated the average true population effect size of the insight-outcome association in a random-effects model. This model was chosen a priori to allow for unconditional inferences about the effect size distribution in the population (48). Residual heterogeneity was estimated with the restricted maximum likelihood estimator (49). Study weight was calculated as the inverse of the sum of the study’s standard deviation and the estimated heterogeneity. The amount of heterogeneity was assessed using the Q-test for heterogeneity, the estimated amount of variability in the true effect sizes (tau), and the proportion of observed variability that can be attributed to true heterogeneity (I²) (38). Values for I² of 25%, 50%, and 75% indicate a low, moderate, and high degree of heterogeneity, respectively (50). We performed prospective power analyses to detect different population effect sizes for conditions of low, moderate, and high heterogeneity in SAS Studio, version 3.6 (51), according to the method described by Hedges and Pigott (52). The analyses indicated adequate power to detect small (power=0.70–0.84), medium (power=0.85–1) and large (power=1) effects for all heterogeneity conditions.

Finally, we assessed the risk of publication bias by visually examining the funnel plot and conducting Egger's regression test for funnel plot asymmetry (38). We applied Duval and Tweedie's trim-and-fill procedure (53) as a sensitivity analysis to estimate the unbiased average population effect.

**RESULTS**

**Characteristics of Included Studies**

The 22 studies (with a k of 23 independent effect sizes) subsumed a total of 1,112 individuals. The mean age of
**TABLE 1. Study and Measurement Characteristics in a Systematic Review and Meta-Analysis of Insight and Outcome of Psychotherapy**

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Type of Treatment</th>
<th>Instrument or Measure</th>
<th>Type of Outcome Measure</th>
<th>Type(s) of Outcome(s)</th>
<th>Type of Association</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aafjes-van Doorn et al. (54)</td>
<td>31</td>
<td>Dynamic</td>
<td>ATOS</td>
<td>CR</td>
<td>CORE-OM, BSI, IIP-32</td>
<td>Symptoms, interpersonal functioning, other</td>
</tr>
<tr>
<td>Ambühl et al. (55)</td>
<td>18</td>
<td>Client-centered/</td>
<td>HRS</td>
<td>CR</td>
<td>Self-developed</td>
<td>Other</td>
</tr>
<tr>
<td>Andreaoli et al. (29)</td>
<td>31</td>
<td>Supportive psychotherapy</td>
<td>PSAI</td>
<td>CR</td>
<td>session outcome</td>
<td>Other</td>
</tr>
<tr>
<td>Castonguay et al. (56)</td>
<td>30</td>
<td>CBT</td>
<td>EXP</td>
<td>CR</td>
<td>BDI, HAM-D, GAS</td>
<td>Symptoms, other</td>
</tr>
<tr>
<td>Cogan and Porcerelli (57)</td>
<td>25</td>
<td>Dynamic</td>
<td>SWAP</td>
<td>CR</td>
<td>GAF</td>
<td>Other</td>
</tr>
<tr>
<td>Connolly et al. (58)</td>
<td>29</td>
<td>Dynamic</td>
<td>SUIP</td>
<td>SR</td>
<td>BAI</td>
<td>Symptoms</td>
</tr>
<tr>
<td>Connolly Gibbons et al. (27)</td>
<td>124</td>
<td>Interpersonal/</td>
<td>SUIP-R</td>
<td>SR</td>
<td>HAM-A, HAM-D, BAI, BDI, QOLI</td>
<td>Symptoms, quality of life</td>
</tr>
<tr>
<td>Cromer and Hilsenroth (41)</td>
<td>71</td>
<td>Dynamic</td>
<td>CDPS</td>
<td>CR</td>
<td>BSI, PEI, SOS, GAF, GARF, SOFAS</td>
<td>Symptoms, interpersonal functioning, other</td>
</tr>
<tr>
<td>Diemer et al. (59)</td>
<td>25</td>
<td>Counseling</td>
<td>Event Insight Scale</td>
<td>CR</td>
<td>GSI (SCL-90)</td>
<td>Symptoms</td>
</tr>
<tr>
<td>Gelso et al. (60)</td>
<td>33</td>
<td>Counseling</td>
<td>RO</td>
<td>CR</td>
<td>COM</td>
<td>Other</td>
</tr>
<tr>
<td>Grande et al. (39)</td>
<td>39</td>
<td>Dynamic</td>
<td>HSCS</td>
<td>CR</td>
<td>Dynamic interview</td>
<td>Other</td>
</tr>
<tr>
<td>Hoffart et al. (31)</td>
<td>35</td>
<td>CBT</td>
<td>PPIQ: self-understanding</td>
<td>SR</td>
<td>PPIQ: postsessional distress</td>
<td>Other</td>
</tr>
<tr>
<td>Høglend et al. (61)</td>
<td>43</td>
<td>Dynamic</td>
<td>Dynamic interview</td>
<td>SR</td>
<td>EDI-3</td>
<td>Symptoms</td>
</tr>
<tr>
<td>Ivanova (62)</td>
<td>16</td>
<td>Emotion focused</td>
<td>CTSC</td>
<td>SR</td>
<td>Other</td>
<td>Interpersonal</td>
</tr>
<tr>
<td>Johansson et al. (63)</td>
<td>100</td>
<td>Dynamic</td>
<td>Self-developed scale</td>
<td>CR</td>
<td>PFS</td>
<td>functioning</td>
</tr>
<tr>
<td>Kallestad et al. (37)</td>
<td>24</td>
<td>Dynamic</td>
<td>ATOS</td>
<td>CR</td>
<td>GSI (SCL-90)</td>
<td>Symptoms</td>
</tr>
<tr>
<td>Kallestad et al. (37)</td>
<td>25</td>
<td>CBT</td>
<td>ATOS</td>
<td>CR</td>
<td>GSI (SCL-90)</td>
<td>Symptoms</td>
</tr>
<tr>
<td>Kivlighan et al. (43)</td>
<td>12</td>
<td>Dynamic</td>
<td>IEQ, IRS</td>
<td>CR</td>
<td>TCS</td>
<td>Other</td>
</tr>
<tr>
<td>Kolden (33)</td>
<td>106</td>
<td>Dynamic</td>
<td>TSR: therapeutic</td>
<td>SR</td>
<td>TSR: session</td>
<td>Other</td>
</tr>
<tr>
<td>Levy et al. (64)</td>
<td>76</td>
<td>Dynamic</td>
<td>SWAP</td>
<td>CR</td>
<td>GSI (SCL-90)</td>
<td>Symptoms</td>
</tr>
<tr>
<td>Luborsky et al. (35)</td>
<td>43</td>
<td>Dynamic</td>
<td>CCRT</td>
<td>CR</td>
<td>Self-developed</td>
<td>Other</td>
</tr>
<tr>
<td>Mohr et al. (65)</td>
<td>90</td>
<td>Counseling</td>
<td>RO</td>
<td>CR</td>
<td>COM</td>
<td>Other</td>
</tr>
<tr>
<td>Nyklicek et al. (42)</td>
<td>86</td>
<td>CBT</td>
<td>BIPM</td>
<td>SR</td>
<td>SCL-90</td>
<td>Symptoms</td>
</tr>
</tbody>
</table>

*a ATOS=Achievement of Therapeutic Objectives Scale; BAI=Beck Anxiety Inventory; BDI=Beck Depression Inventory; BIPM=Balanced Index of Psychological Mindedness; BSI=Brief Symptom Inventory; CBT=behavioral therapy; CCRT=core conflictual relationship themes; CDPS=Capacity for Dynamic Processes Scale; COM=Counseling Outcome Measure; CORE-OM=Clinical Outcomes in Routine Evaluation Outcome Measure; CTSC=Client Task Specific Change Measure; EDI-3=Eating Disorder Inventory-3; EXP=Experiencing Scale; GAF=Global Assessment of Functioning; GARF=Global Assessment of Relational Functioning; GAS=Global Assessment Scale; GSI=Global Severity Index; HAM-A=Hamilton Anxiety Rating Scale; HAM-D=Hamilton Depression Rating Scale; HRS=Heuristik-Rat1ng-Skalen; HSCS=Heidelberg Structural Change Scale; HSRS=Health-Sickness Rating Scale; IIP-32=Inventory of Interpersonal Problems-32 item; IRS=Insight Rating Scale; PBI=Patient Estimate of Improvement; PFS=Psychodynamic Functioning Scales; PPIQ=pre- and postsession impact questionnaire; PSAI=Psychotherapeutic Attainment Index; QOLI=Quality of Life Inventory; RO=Relationship Questionnaire, insight scale; SCL-90=Symptom Checklist-90; SOFAS=Social and Occupational Functioning Assessment Scale; SSR=Schwartz Outcome Scale; SWAP=Shedler-Westen Assessment Procedure, insight scale; TCS=Target Complaint Scale; TSR=Therapy Session Report. CR=clinician rated; SR=self-report.

This column reports which indicators of insight and outcome were correlated:

1=insight and same-session outcome; 2=insight and subsequent outcome; 3=insight and change in outcome; 4=change in insight and change in outcome.

Participants was 35.27 years (SD=6.01, range=23.70-46.40), and 63.84% of the total sample was female. Participants attended a median of 20 sessions (mean=73.27, SD=239.63, range=1-1118). There was considerable variety in psychopathology. Eighteen of the studies did not have major restrictions regarding diagnostic categories and used convenience samples of patients seeking treatment. In these studies, patients were mostly diagnosed as having axis I disorders according to DSM-IV-TR criteria (or criteria from earlier versions of DSM), among them mood, anxiety, somatoform, and eating disorders.
There was considerable comorbidity with axis II disorders. The remaining four studies comprised samples of patients with mood, anxiety, eating, or cluster C personality disorders. Eighteen studies provided individual outpatient treatment. Three studies were based on inpatients employed an outpatient group of individual and group who received a combination of psychotherapies. One study other treatment types, such as psychodynamic and 11 as focused, and mixed forms of humanistic, client-centered, and interpersonal psychotherapy. The 22 studies used 18 different insight measures and 27 different outcome scales. Six studies examined patients' self-reports of insight, and the remaining 16 studies used clinician-rated insight scales. As outcome measures, 12 studies employed measures of psychopathological symptoms, three studies assessed interpersonal functioning, and one study evaluated quality of life. For 13 studies, other categories, such as occupational functioning and target complaints, were used as outcome measures. Studies employed various designs to associate insight and outcome. Five effect sizes were based on cross-sectional correlations between insight and session outcome, three were correlations of insight and subsequent outcome, five were correlations of insight with change in outcome, and nine represented associations between change in insight and change in outcome.

**Overall Meta-Analysis and Heterogeneity**

The random-effects model for the association between insight and outcome estimated an average true population effect size (r) of 0.31 (95% CI=0.22–0.40, p<0.05). Effect sizes and 95% confidence intervals of the studies (27, 29, 31, 33, 35, 37, 39, 41–43, 54–65) are presented in a forest plot in Figure 2. Eleven of the 23 reported individual effect sizes were nonsignificant within the primary study. There was significant heterogeneity in the effect sizes (Q=58.71, p<0.05). The estimated standard deviation of true effects (tau) was 0.19 (95% CI=0.10–0.28), and the ratio of true heterogeneity to total variation in the observed effects (I²) was 60.30% (95% CI=29.08–78.02). The amount of heterogeneity between studies suggested the examination of potential moderating variables. Moderator analyses were attempted for type of treatment (dynamic versus other treatments), patient diagnosis (percentage of personality disorders within study sample), and source of information on insight (observer rating versus self-report). However, the presumed statistical power for the moderator analyses was low (<0.20) (38, 66). Within treatments classified as dynamic, the magnitude of effects (r) ranged from 0.01 to 0.59. Effect sizes (r) for the studies on CBT ranged from −0.08 to 0.34, and those for the unspecified counseling studies varied from 0.24 to 0.68. The associations in the remaining studies on emotion-focused, interpersonal, client-centered, supportive, and mixed types of therapy ranged from 0.19 to 0.53.

**Study Quality and Risk of Bias**

The assessment of study quality with the risk of bias scale demonstrated moderate risks across individual studies and risk of bias categories (see Table 2). Risk of bias scores were highest for sample risks, meaning that most studies lacked sufficient sample size or representativeness or did not adequately clarify the source of the given sample. There was a high risk of measurement bias in about a third of the studies, where insight or outcome measures were not operationalized precisely or did not have adequate psychometric properties. In about half the studies, there was a notable risk of biased

![Forest Plot of Effect Sizes Measured as Correlations Between Insight and Outcome in a Meta-Analysis of Insight and Outcome of Psychotherapy](image-url)
TABLE 2. Risks of Biases in Individual Studies in a Meta-Analysis of Insight and Outcome of Psychotherapy

<table>
<thead>
<tr>
<th>Study</th>
<th>Risk of Sample Bias</th>
<th>Risk of Measurement Bias</th>
<th>Risk of Biased Statistical Analyses</th>
<th>Risk of Reporting Bias</th>
<th>Total Risk of Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aafjes-van Doorn et al. (54)</td>
<td>0.67</td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.20</td>
</tr>
<tr>
<td>Ambühl (55)</td>
<td>0.67</td>
<td>0.63</td>
<td>0.20</td>
<td>0.00</td>
<td>0.39</td>
</tr>
<tr>
<td>Andreoli et al. (29)</td>
<td>0.33</td>
<td>0.38</td>
<td>0.33</td>
<td>0.33</td>
<td>0.31</td>
</tr>
<tr>
<td>Castonguay et al. (56)</td>
<td>0.67</td>
<td>0.00</td>
<td>0.60</td>
<td>0.00</td>
<td>0.31</td>
</tr>
<tr>
<td>Cogan and Porcerelli (57)</td>
<td>0.50</td>
<td>0.00</td>
<td>0.35</td>
<td>0.17</td>
<td>0.32</td>
</tr>
<tr>
<td>Connolly et al. (58)</td>
<td>0.33</td>
<td>0.00</td>
<td>0.42</td>
<td>0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>Connolly Gibbons et al. (27)</td>
<td>0.17</td>
<td>0.00</td>
<td>0.42</td>
<td>0.00</td>
<td>0.17</td>
</tr>
<tr>
<td>Cromer and Hilsenroth (41)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.42</td>
<td>0.00</td>
<td>0.17</td>
</tr>
<tr>
<td>Diemer et al. (59)</td>
<td>0.50</td>
<td>0.13</td>
<td>0.42</td>
<td>0.00</td>
<td>0.28</td>
</tr>
<tr>
<td>Gelso et al. (60)</td>
<td>0.67</td>
<td>0.63</td>
<td>0.50</td>
<td>0.67</td>
<td>0.59</td>
</tr>
<tr>
<td>Grande et al. (39)</td>
<td>0.50</td>
<td>0.38</td>
<td>0.25</td>
<td>0.17</td>
<td>0.31</td>
</tr>
<tr>
<td>Hoffart et al. (31)</td>
<td>0.50</td>
<td>0.63</td>
<td>0.42</td>
<td>0.00</td>
<td>0.41</td>
</tr>
<tr>
<td>Haglend (61)</td>
<td>0.67</td>
<td>0.63</td>
<td>0.42</td>
<td>0.00</td>
<td>0.44</td>
</tr>
<tr>
<td>Ivanova (62)</td>
<td>0.50</td>
<td>0.00</td>
<td>0.50</td>
<td>0.33</td>
<td>0.34</td>
</tr>
<tr>
<td>Johansson et al. (63)</td>
<td>0.17</td>
<td>0.00</td>
<td>0.50</td>
<td>0.00</td>
<td>0.23</td>
</tr>
<tr>
<td>Kallestad et al. (37)</td>
<td>0.33</td>
<td>0.25</td>
<td>0.08</td>
<td>0.00</td>
<td>0.16</td>
</tr>
<tr>
<td>Kivlighan et al. (43)</td>
<td>0.50</td>
<td>0.25</td>
<td>0.25</td>
<td>0.67</td>
<td>0.38</td>
</tr>
<tr>
<td>Kolden (33)</td>
<td>0.00</td>
<td>0.13</td>
<td>0.42</td>
<td>0.50</td>
<td>0.28</td>
</tr>
<tr>
<td>Levy et al. (64)</td>
<td>0.00</td>
<td>0.38</td>
<td>0.00</td>
<td>0.00</td>
<td>0.10</td>
</tr>
<tr>
<td>Luborsky et al. (35)</td>
<td>0.50</td>
<td>0.38</td>
<td>0.40</td>
<td>0.50</td>
<td>0.43</td>
</tr>
<tr>
<td>Mohr et al. (65)</td>
<td>0.17</td>
<td>0.00</td>
<td>0.47</td>
<td>0.00</td>
<td>0.21</td>
</tr>
<tr>
<td>Nylíkéck et al. (42)</td>
<td>0.00</td>
<td>0.25</td>
<td>0.50</td>
<td>0.33</td>
<td>0.31</td>
</tr>
</tbody>
</table>

a Risks of biases were assessed with the adapted Systematic Assessment of Quality in Observational Research Scale (see the online supplement for the complete scale). Columns list mean scores of risk of bias categories averaged across two independent raters. Scores are ratios of prevalent risks standardized by applicable risk items and yield numbers between 0 and 1; higher scores represent greater risk of bias.

calculations because of attrition rates ≥5%, insight and outcome ratings provided by the same person, unclear times of measurement, same-session assessment of insight and outcome, or no direct availability of a zero-order correlation. The least concerns were present for risk of reporting bias, as most studies gave explanations for missing data, presented data clearly and accurately, and showed no indicators of selective reporting. Only three studies reported therapist effects, and they controlled for them using a multilevel modeling approach (54, 64, 65). None of the studies had a preregistered study protocol, and none explicitly stated taking preventive actions against possible researcher allegiance effects, such as blind ratings of insight or blind data analysis.

Risk of Publication Bias and Sensitivity Analysis

Finally, we examined whether publication bias was a reason for concern. Visual inspection of the funnel plot for the random-effects model suggested minor asymmetry, and Egger’s regression test for funnel plot asymmetry was nonsignificant (z = −0.77, p > 0.05). We applied Duval and Tweedie’s trim-and-fill procedure to assess the potential impact of a bias. The method imputed two studies to achieve funnel plot symmetry, leading to an estimated unbiased effect size (r) of 0.35 (95% CI: 0.25–0.45, p < 0.05). The original and adjusted effect size estimates therefore differed only in the second decimal. Figure 3 presents the funnel plot including the data filled in by the trim-and-fill method.

DISCUSSION

In this study, we reviewed the association between insight and psychotherapy outcome. A systematic literature search yielded 23 eligible independent effects. The random-effects meta-analysis estimated a moderate association between insight and psychotherapy outcome. Its magnitude is comparable with effect sizes of established treatment factors, such as the therapeutic alliance (67, 68), positive regard (69), and empathy (70). In other words, our findings indicate that insight may be a similarly relevant factor in psychotherapy. This is in line with clinical models about how self-understanding influences outcome. When gaining insight, patients may become aware of connections between present psychological problems and past experiences (25). They may also realize how they themselves contribute to the recurrence of these experiences (17). In terms of defense mechanisms, an increased awareness of patients’ pathological compromise formations, which contribute to the development and maintenance of the disorder, is a prerequisite to developing more adaptive ways of dealing with threatening or unpleasant experiences. This may lead to a reduction of distorted perceptions of oneself and others and to a better integration of unpleasant experiences into the conscious part of the self (21, 22). Therefore, self-understanding may lead to symptom reduction in two ways. First, understanding in itself increases patients’ sense of control and mastery. Although sometimes painful at first, knowing about one’s own contributions to the development and maintenance of symptoms is an inevitable first step in confronting difficulties, acting on, and eventually mastering
them. Second, self-understanding enables patients to find new solutions and more adaptive ways of behaving (11, 16, 18).

Limitations and Further Research
Although the results of our meta-analysis are encouraging for the relevance of insight as a curative factor in psychotherapy, limitations must be noted. First and most important is that the analyses are based exclusively on correlational data and thus do not allow causal inferences. Not only could insight lead to improved well-being, but it is also possible that symptom reduction enables patients to reflect about themselves more openly and gain new insights (reverse causation). Furthermore, the association between insight and psychotherapy outcome could be caused by a third variable, such as enhanced quality of object relations. The studies discussed in this review vary in the study designs used and the conclusions that can be drawn. Cross-sectional correlations between insight and session outcome have the highest probability of confounding factors or reverse causation. For associations between insight and subsequent outcome or changes in outcome, confounding factors may still be present, but the risk of reverse causation is diminished. About half the studies in this review correlated change in insight with change in outcome and thus tested the insight-outcome association in a more rigorous way. To establish insight as a mediator, that is, a mechanism that brings about change in psychotherapy, several conditions must be met. First, the treatment must be associated with change in outcome. Second, the treatment must be associated with change in insight. Third, change in insight must predict change in outcome. Fourth, when insight is included in the statistical model, the effect of treatment on outcome must be diminished or eliminated. Fifth, it should be demonstrated that change in insight temporally precedes change in outcome (71–73). In order to establish insight as a specific mechanism (as opposed to a common factor), the treatment rationale should target insight, for example, via interventions such as confrontation or interpretation for dynamic therapies. Ulberg and colleagues (71) demonstrated that increased insight during treatment mediated the effect of transference interpretations on improvements in interpersonal functioning. The data available for the present correlational meta-analysis only permitted tests of the association between insight (or change in insight) and outcome (or change in outcome). Further longitudinal research that facilitates all required steps for mediation is necessary to provide evidence for insight as a putative mediator. Nevertheless, demonstrating a significant association between self-understanding and psychotherapy outcome is an important first step to establish insight as mechanism of change (3).

Second, although the topic is relevant for a variety of therapeutic methods and inclusion criteria were defined broadly, relatively few studies were eligible for this meta-analysis. This not only reduced the precision of the population effect size estimate, but it also prevented moderator analyses because of low statistical power. Moderators are pretreatment variables that alter treatment response and thus would have answered the question of which patient characteristics or treatment conditions influence the magnitude of the insight-outcome association (71–73). Further research is necessary to examine potential moderators, such as type of treatment, patient diagnosis, and the source of information on insight. Testing the type of treatment as a moderator will address whether insight should be considered a common factor or whether it is specific to psychodynamic therapy. Patient diagnosis is of interest, because different mechanisms of change may operate for patients with and without personality disorders. For example, patients with personality disorders seem to benefit more from transference interpretations that are assumed to enhance their interpersonal functioning (74, 75). Whether insight is of different importance for various groups of patients is an important topic for future research. Furthermore, the source of information on insight may be influential. Because lower insightfulness can limit patients' ability to report on their degree of insight, studies with self-reports may yield different results compared with observer ratings.

Finally, the large number of different insight measures may be viewed as a threat to validity. Although only studies using scales that fit our preregistered definition of insight were included, the variety of measures raises the concern that different aspects of the construct were assessed. The different measures may have caused part of the heterogeneity between study effect sizes. Nonetheless, the amount of true heterogeneity in our study was within the usual range of meta-analyses on change mechanisms (such as the alliance [67, 68]) and still led to a considerably precise population effect size estimate that allows conclusions on the insight-outcome association.
ASSOCIATION BETWEEN INSIGHT AND OUTCOME OF PSYCHOTHERAPY

CONCLUSIONS

To our knowledge, this study is the first systematic review and meta-analysis of the association between insight and outcome of psychotherapy. Across studies, more insight is moderately associated with better psychotherapy outcome. Although further research is needed to understand determinants of the association, this contributes to our understanding of how and why the investigated psychotherapies work.

AUTHOR AND ARTICLE INFORMATION

From the Department of General Internal Medicine and Psychosomatics and the Institute of Medical Psychology, Heidelberg University, Heidelberg, Germany.

Address correspondence to Dr. Dinger (ulrike.dinger@med.uni-heidelberg.de) and Ms. Jennissen (simone.jennissen@med.uni-heidelberg.de).

The authors report no financial relationships with commercial interests.

Received August 2, 2017; revisions received January 22 and April 6, 2018, accepted April 19, 2018; published online Aug. 2, 2018.

REFERENCES

6. Norcross JC, Lambert MJ: Psychotherapy relationships that work, II. Psychotherapy (Chic) 2011; 48:4-8
Insight Into Insight in Psychotherapy

Per Høglend, M.D., Ph.D.

There have been more than 5,000 randomized controlled trials of psychotherapy. Most meta-analyses conclude that psychotherapy is effective for a range of disorders, and different modalities of therapy are equally effective. However, studies with the strongest research designs indicate that psychotherapy may be considerably less effective than we tend to believe (1). In fact, different psychotherapies may be equally ineffective. Too many patients fail to respond or continue to experience residual symptoms after treatment termination. All treatments for mental disorders have a wide margin for improvement.

Mental health professionals still believe that psychotherapy can be improved by theoretical discussion, logic, and expert opinion (2). This may be one of the reasons for the existence of several hundred “brand name” psychotherapies. However, in order to improve clinical theories, treatment interventions, and training, we need to identify how psychotherapy works—the mechanisms of change underlying the treatment effects (3). It is only with the specification of mediators of change, whether psychological (3) or biological (4), that the field will move forward. Many mechanisms may have an impact in the complex process of psychotherapy, but insight is regarded as one of the most important mechanisms of change, not only in dynamic psychotherapy (5) but also, increasingly, in cognitive behavioral therapy (6) and experiential psychotherapy (7).

In this regard, the study reported in this issue of the Journal by Jennissen and colleagues (8) is timely. This is the first meta-analysis of studies with data on the insight-outcome correlation. The meta-analysis is generally well done, with all of the elements that one would want to see—checks for heterogeneity, study quality, weighting, risk of publication bias. The authors screened more than 12,000 abstracts, assessed 261 full-text papers, and ended up with 22 studies (with 23 effect sizes). Their definition of insight is broad, but it is mainstream and reasonably accurate. They define it as “patients’ understanding of associations between past and present experiences, typical relationship patterns, and the relation between interpersonal challenges, emotional experience, and psychological symptoms.” The eligibility criteria used are sensible. Although determining which studies to include in a meta-analysis is a notoriously difficult gray area, study selection seems carefully implemented. Data extraction also seems reasonable, given the design limitations in many (most?) of these studies. Two authors independently assessed study quality (risk of bias), and interrater reliability on these assessments was high. Risk of bias varied from high to low among the studies. Patients were mostly diagnosed with mood, anxiety, somatoform, eating, and cluster C personality disorders. Twelve treatments were psychodynamic and 11 were other treatments, such as cognitive-behavioral, emotion-focused, client-centered, and interpersonal psychotherapy. The studies used 27 different outcome measures and 18 different insight measures. Most measures of insight were clinician rated. Tests of heterogeneity indicated significant variation in effect sizes and a moderate to high degree of “true” heterogeneity.

This meta-analysis has limitations, largely because of the nature of the studies included. As outlined above, there was considerable variety in psychopathology, and treatments varied in theoretical orientation. The studies employed various designs to associate insight and outcome. None of the studies had a preregistered study protocol, and researcher allegiance effects could not be evaluated. When a meta-analysis includes different treatments for a range of psychiatric disorders, a number of different outcome measures, and different measurement designs, the otherwise sensible approach of lumping data together to address the problem of small samples may develop into a progressively messy endeavor, confounding the clinical relevance of the findings. To offset some of these validity threats, the authors commendably addressed many design issues and described relevant details of what was actually tested in each study. In this way, readers may be helped to draw their own conclusions regarding the validity of the findings.

Jennissen et al. report a consistent pattern of associations between insight and outcome across disorders and therapy types. There was a significant moderate relationship ($r=0.31$) between insight and treatment outcome. The study was not powered to detect moderators of the insight-outcomes associations. Overall, the findings constitute preliminary support for the importance of insight in different forms of psychotherapy.

Jennissen and colleagues’ choice to include many different diagnoses and different treatments in their meta-analysis may reflect current developments in our field. Given the shared brain circuitry, genes found across disorders, medications used across disorders, and the reality
of considerable comorbidity, there is growing recognition that only a few domains of dysfunction may extend across the many DSM diagnostic categories. Caspi et al. (9) have argued that psychopathology could be explained with one general psychopathology dimension. The Research Domain Criteria initiative of the National Institute of Mental Health focuses on only a few domains of dysfunction, social processes being one of them. Transdiagnostic approaches and unified trans-theoretical treatment protocols may improve future research.

Jennissen and colleagues' meta-analysis estimated a moderate association between insight and psychotherapy outcome across diagnostic categories and different schools of psychotherapy. Its magnitude is comparable to effect sizes of established treatment factors, such as therapeutic alliance. A large number of studies have documented a moderate but consistent relationship between therapeutic alliance and outcome, across different treatments and diagnoses (10). Brain changes, too, mainly in the prefrontal cortex, cingulate gyrus, and amygdala, are associated with symptom improvement across different diagnoses and treatments (11). However, an association with outcome is necessary but not sufficient. Just like insight and alliance, brain changes may be a consequence of symptom change or a proxy for some other therapeutic change processes. Further longitudinal research with study designs that include most or all steps in mediation research is necessary to establish insight, or alliance, or brain changes as mechanisms for treatment outcome. In our own research into the effect of insight on the outcome of psychotherapy, my colleagues and I found that use of a specific technique in dynamic psychotherapy (transference work) increased gain of insight during therapy substantially, with a subsequent improvement in outcome (12).

The crucial questions are: Which therapist interventions and treatment processes increase insight, or strengthen the alliance, or influence the desired brain changes? Nonetheless, this study is a first important step to establishing insight as a possible mechanism of change in psychotherapy and may serve to challenge investigators to conduct more rigorous evaluations of the insight-outcome relationship to advance knowledge.

REFERENCES