Transference, Transference Interpretations, and Transference-Focused Psychotherapies

Kenneth N. Levy and J. Wesley Scala
Pennsylvania State University

The concept of transference and the use of transference interpretations in psychotherapy have been highly controversial topics garnering frequent attention both within psychoanalysis and across multiple orientations of psychotherapy. In this article, we review the empirical evidence as it bears on this controversy and discuss the implications of the evidence for psychoanalysis, psychodynamic psychotherapy, and therapy in general. We provide a brief historical and contextual overview, followed by a discussion of the development of the concept of transference. We then discuss the evidence for the concept of transference from basic psychological research and contend that these findings are not only consistent with a social–cognitive and information processing model, but that they may also indicate conflict and defensive processes suggestive of a dynamic transference process model. We continue with a discussion of the evidence for the concept of transference from psychotherapy research and examine process findings relating to the use of transference interpretations and transference-focused psychotherapies. Finally, we present the implications of this emerging evidence for clinical practice.

*Keywords:* psychodynamic, transference, psychotherapy outcome and process research

Although the concept of transference has been central within psychoanalysis and psychodynamic psychotherapy since Freud’s earliest writings, and it has been used broadly across multiple psychotherapy orientations (Gelso, this issue, pp. 384–390; Gilbert & Leahy, 2007), the concept of transference and the use of transference interpretation in psychotherapy have been highly controversial topics (Frances & Perry, 1983; Gabbard et al., 1994; Gunderson, Najavitz, Leonhard, Sullivan & Sabo, 1997). In this manuscript, we will review recent empirical evidence as it bears on this controversy and the implications of these findings for psychoanalysis, psychodynamic psychotherapy, and therapy in general. We begin with a brief historical and contextual overview, outlining some of Freud’s early ideas and developments contributing to the evolution of the concept of transference, as well as those contributions from some of Freud’s followers. We then examine the evidence for the concept of transference from both basic psychological research and psychotherapy research. A substantial amount of evidence has emerged from basic social psychological research, suggesting that transference is not a uniquely clinical phenomenon (Andersen & Saribay, 2005; Andersen & Chen, 2002; Andersen & Berk, 1998). We will examine this evidence, paying special attention to findings related to attachment representations (Brumbaugh & Fraley, 2006, 2007), projective mechanisms (Mikulincer & Horesh, 1999), and transference as it operates in daily life (Andersen & Berk, 1998), all of which have implications for understanding basic transference processes and for clinical practice. Within the realm of psychotherapy research, we examine evidence from process and experimental psychotherapy research (randomized controlled trials [RCTs]) related to the use of transference interpretations, and transference-focused psychotherapies as they relate to therapeutic outcome. Finally, we present the implications of this emerging evidence for clinical practice.

**Brief Historical and Contextual Overview**

Transference first appeared in Freud’s neurological writings in 1888 (Freud, 1888). However, the concept of transference was not a simple solitary discovery, but evolved over years of creative synthesis that was rooted in the discourses of his time. In this early writing, Freud used the concept of “displaceable energies” to indicate the transfer of strong feelings developed within a particular relationship to another person who was independent of the origin of those feelings. The concept of transference was further elaborated later in Case Studies in Hysteria in 1895 (Breuer & Freud, 1895). Here, Freud talked about a false connection where the patient transfers unconscious ideas about a figure from the past onto the person of the physician, and also noted that this compulsion or illusion melted away with the conclusion of the analysis. By 1900, transference was ready to stand as it does today as the core psychoanalytic theory.

In his discussion of the Dora case, Freud (1905/1963) elaborated on this concept and introduced an interesting complexity. He noted that although transferences were generally regarded as the similes that replaced some earlier person by the person of the physician, some were found to be more ingenuously constructed by the patient. The patient does this, “... by cleverly taking advantage of some real peculiarity in the physician’s person or circumstances and attaching themselves to that” (Freud, 1905/1963; p. 107).
Freud also noted that these “infantile” prototypes re-emerged and were experienced with a strong sensation of immediacy and that at the beginning of his treatment with Dora, “... it was clear that [Freud] was replacing [Dora’s] father in her imagination” (Freud, 1905/1963; p. 108), she was constantly comparing him with her father, even at a conscious level. Freud was recognizing that there may be real aspects of the figure that the transference is projected onto, the physician or the psychotherapist in this case, that might actually pull for or allow for the experience of transference, points later elaborated by Gill (1979, 1982) and Gelso (2010).

Evolution of the Transference Concept

In 1912, using a metaphor of his time, Freud wrote about transference as a “stereotype plate” (Freud, 1912). A stereotype plate was a method of printing, developed in 1789, using a solid plate, where the image was perpetuated without change from that plate. Freud related this to the concept of transference by noting that these stereotype plates, or prototypes of others, form through one’s interactions with others and that these prototypes are carried forward to future relationships. In particular, he noted that patients incorporate aspects of the physician into these preexisting stereotype plates. Freud (1912) suggested that this transference, or stereotype plate, determines a person’s later erotic or sexual interests. He also noted that transference was partly conscious and changeable, and partly unconscious and relatively impervious to development or change. At this point, he also talked about transference as a resistance in psychotherapy and that the resolution of transference was synonymous with the resolution of neurosis. This is similar to Freud’s (1905/1963) earlier idea about a compulsion or illusion that is melted away with the conclusion of the analysis. Freud (1915) also noted that patients can enact interpersonal patterns in therapy, noting that in doing so they were showing new additions of old conflicts and that the most serious difficulties the analyst must grapple with are related to the transference.

Freud elaborated on his earlier work when, just before his death, he talked about transference as the central mechanism of therapeutic change and saw the central task in psychoanalysis as the establishment, interpretation, and resolution of transference (Freud, 1937).

Kleinian contributions to the transference concept. The evolution of the concept of transference continued after Freud’s death with Melanie Klein, who noted that the patient can nudge the analyst through interpersonal pressure to take on characteristics similar to the original source from which the transference is derived and is now being projected (Klein, 1952). This was an important evolution because Freud initially said that not only is transference a distortion, but that it also represented some real aspects of the relationship with the therapist, noting peculiarities or aspects of the therapist that might relate to issues of transference. However, Klein suggested that it goes even one step further, in that the patient might act in ways to elicit behaviors from others that are consistent with their transference, and the amount of transference can vary as a function of the individual, the target, and the situation. Finally, we contend that an important feature of transference is that some aspects are unconscious and related to conflicts and defensive processes.

Evidence for the Concept of Transference

There are three essential areas that have produced evidence for the concept of transference. These areas are basic cognitive and social psychological research, psychotherapy process research, and neuroscience research, although the data from neuroscience research are thin at this point.

Basic Research

Much of the basic cognitive and social psychological research on transference has centered around the work of Andersen and colleagues (Andersen & Berk, 1998; Andersen, Glassman, Chen, & Cole, 1995; Andersen, Glassman, Gold, 1998; Glassman & Andersen, 1999), who have developed an elegant paradigm for assessing transference phenomena. This paradigm, explained more fully in Andersen & Przybylinksi (this issue, pp. 370–383), generally involves two sessions. In the first session, participants are asked to provide a description of significant others. In a later session, participants are exposed to others and asked to rate or describe them. For example, participants might read narratives about fictional characters and are then asked to describe them. Consistently, participants wrongly attribute traits used to describe significant others to describe the new person. For example, par-
Participants wrongly attributed traits that were not part of a fictional character’s description but that stemmed from their description of their significant others. This research has consistently shown that significant-other representations are activated and applied or transferred to a novel target (e.g., new person) in everyday social perception. That is, people tend to view others in ways that are consistent with preexisting significant-other representations. Further, the social–cognitive research on transference indicates that these significant-other representations are chronically accessible and in a continual state of readiness for use and do not require priming. Additionally, these representations are applied to new people even when there is no concrete similarity between them. The effect is also enhanced or amplified when primed and persists and is exacerbated over time.

However, in thinking about the social–cognitive research on transference, one must ask, whether this research truly indicates that there is a *dynamic* transference process as posited in psychoanalytic theory? That is, is there a motivated, conflict, or defensive component to transference, as Freud suggested, or is the process simply a universal cognitive bias in making judgments about other people that might have accorded some general evolutionary advantage and yet sometimes results in incorrect inferences? One important step in showing that transference is a dynamic process and not simply a cognitive bias is to look at whether individual differences in transference exist and whether these individual differences are related to dynamic conflicts, defensive processes, or unconscious motivations. In this respect, we ask whether the findings of Andersen and colleagues could be understood solely from an information processing or cognitive bias paradigm or if there is something uniquely dynamic about the transference process? Andersen and colleagues tend to focus on the social–cognitive and information processing aspects of transference, as articulated by Sullivan and Horney, as well as social–cognitive psychologists (e.g., Mischel; Higgins), and downplay the conflict or defensive aspects of transference phenomena, as articulated by Kohut and Kernberg. However, some findings from research on transference from an attachment theory perspective, including some work from Andersen’s laboratory (Andersen, Bartz, Berenson, & Keckemethy, 2006), have yielded results that suggest a dynamic process consistent with a psychoanalytic approach.

From an attachment theory perspective, Brumbaugh and Fraley (2006, 2007) examined how perceptions of past romantic partners affected perceptions of new people. They found that people applied their attachment representations of past partners both to a possible romantic partner who resembled their past partner and to one who did not resemble their past partner, suggesting a general transference process. Furthermore, the participants did so to a greater degree when the target resembled their past partner, indicating a specific transference process. Importantly, the investigators also found that people tended to feel more anxious and less avoidant, or less defensive, toward the target who resembled their past partner. Taken together, these findings suggest that transference is influenced by aspects of the person, that some of it is reality based, and that there are both general and specific transference processes operating. But most importantly, these findings indicate that individual differences in attachment-based defensive processes are related to transference phenomena. Andersen and colleagues (Andersen et al., 2006) also examined transference from an attachment perspective. They found that securely attached individuals displayed more positive mood in the transference activation condition than did dismissive and preoccupied individuals. In contrast, preoccupied individuals had more anxious mood than secure, dismissive, and fearful individuals when in the transference activation condition. The authors also found that dismissively attached individuals were more motivated to avoid the new person in the transference activation condition, than were secure and fearful individuals. Thus, when the mental representations of caregivers were triggered in the transference activation conditions, individual differences in attachment style predicted affective states in ways consistent with a dynamic transference process model.

Earlier attachment research (Mikulincer & Horesh, 1999) on projective processes is also consistent with a dynamic transference process model. This research has shown that people with secure attachment representations were relatively unbiased by projective mechanisms. In contrast, people with anxious-ambivalent attachment representations perceived others in the same negative ways they viewed themselves, whereas avoidant individuals perceived in others the negative traits in themselves that they tried to suppress. These findings suggest that the perceptions of others can be based on individual differences in projective mechanisms where one transfers aspects of his or her sense of self onto other people in ways consistent with his or her attachment style. Furthermore, as predicted by a dynamic transference model, less-defensive securely attached individuals see others as relatively more realistic. Although these findings concern projections of self-representations as opposed to those of caregivers, they have relevance for thinking about transference processes, in that they involve the transfer of representations, which are influenced by relationships with caregivers (i.e., attachment style) and associated defensive processes.

Taken together, these studies of transference and projection from an attachment context are consistent with a dynamic process underlying transference phenomena. Specifically, working models of attachment relationships are not only transferred to new relationships in both general and specific ways that are consistent with a social–cognitive or information processing model, but individual differences exist as a function of attachment anxiety, avoidance, and security, which may be indicative of conflict and defensive processes suggestive of a dynamic transference process model.

**Psychotherapy Process Research**

A number of studies have noted the relationship between the patient’s narratives of others and their narratives of therapists (Fried, Cris-Crystoph, & Luborsky, 1992; Crits-Christoph, Demorest, Muenz, & Baranackie, 1994; Connolly, Crits-Christoph, Barber, & Luborsky, 2000; Barber, Foltz, DeRubeis, & Landis, 2002; Waldinger et al., 2002; Tellides et al., 2008). These studies generally find similarity between these two sets of narratives and that this relationship becomes stronger over time. However, it is important to note that the effect is rather modest, that it is only present for a subset of patients, and that even when present, there is a high degree of variability in the amount of transference shown. Blatt, Stayner, Auerbach, and Behrends (1996) assessed representations of self, each parent, and the therapist every 6 months over a 2-year period in a group of severely disturbed adolescent inpatients. Their sample was a highly comorbid group with many of the patients meeting criteria for personality disorders, particularly...
borderline personality disorder (BPD), but also depression, dysthymia, attention-deficit/hyperactivity disorder, learning disabilities, conduct disorder, and oppositional defiant disorder. The parent and self-descriptions were generally seen as negative at the beginning of treatment, whereas the therapist was seen as positive but with reservations, particularly in terms of a specific trait or traits admired or aspired for by the patient. By the middle of treatment, there was greater convergence between the representations of parents and of the therapist, with both seen negatively. As the treatment progressed, the parents were seen more positively, but the therapist continued to be seen negatively. However, at the end of treatment, all the descriptions were judged to be positive, balanced, and realistic. Blatt and colleagues suggested that conflicts about parents are worked out in the psychotherapy process through one’s sense or representation of the therapist, in part because the therapist is the safest person to do so with. Furthermore, as these conflicts are worked through, patients can tolerate more accurate representations of their self and significant others and see people, including themselves, more accurately. This interpretation, although not definitive, may also suggest a dynamic transference process.

Analysis of Transference and Transference Interpretations

The focus on transference phenomena within the therapeutic relationship in psychoanalysis led to an emphasis on the analysis of transference and the use of transference interpretations. Transference interpretations focus on connecting the patient’s feelings and behaviors that are occurring in the here-and-now of the therapy with regard to the therapist with the patient’s preconceived representational models of significant others. Most prototypically, a transference interpretation is a tactful comment that clarifies and links the patient’s experience of others outside of therapy with that of the therapist in therapy and to the patient’s experience of past relationships with caregivers (Levy, 2009). For example a therapist might say:

Therapist. I am not completely sure of this but I think something just happened between us that might be useful for us to explore. I think it is worth our considering (patient nods agreement). I noticed that you became tense and that your voice sounded angry in response to the question I just asked you. Is that consistent with what you noticed?

Patient. (pause, patient appears to be thinking). Yeah, I guess (offered in a somewhat calmer manner but still a bit annoyed and possibly acquiescing).

Therapist. Well, from the way you just said that, I am wondering if a part of you agrees with what I just pointed out, a part that on reflecting on your experience might have noticed that you seemed tense and annoyed, but that another part of you also might be having a reaction to what I said and that part is not fully buying what I said?

Patient. No, no I realize I am tense and feeling angry. But not [angry] at you.

Therapist. Well, I think what is happening right now in discussing your reactions to my observation is related to the larger comment that I was going to make (patient now looks attentive and interested). Based on your reactions to what I said just before, that is, my earlier question, it seems to me that you experienced my asking it [the question] as an attack on you—just like you might have experienced my observation as a miniattack. It is as if in asking the question, I was trying to belittle you. Do you think that sounds accurate or am I off the mark?

Patient. No that sounds right. It is hard to not hear it that way.

Therapist. Well, that is what I think would be useful for us to explore. Does it have to be heard that way or is there another way of hearing it? It seems like it is difficult for you to imagine that my question was an indication of concern and that I asked it to better understand your experience. (Pause), I think your reaction would be justified if I meant my comment as an attack, but on the other hand, if I did not, and if I asked it as sign of concern—and I do not mean this as a criticism, but I think it is important that we understand your reaction—then hearing it this way, you might be robbing yourself of a moment in which you could feel as if someone was on your side and trying to be helpful. And, this dynamic reminds me of what you were telling me about your experience of your coworkers in which you feel that they do not have your back and instead are trying to provoke you and undermine you.

Patient. Well, I am sure that they are [being provocative and undermining].

Therapist. Maybe so, but the question arises if your perception of those events, which is consistent with what we have discussed as your experience with your father, who you saw as critical, harsh, and competitive with you, is influencing how you are experiencing me, right now, here in this room, so that a question offered out of concern is similarly experienced as an attack and belittling.

In this interpretation, the therapist makes a clear connection between the here-and-now of the patient’s relationship with the therapist and the there-and-then relationship of the patient’s relationship with his or her father. However, the connection to early experiences with caregivers is not always explicitly mentioned. Particularly when working with certain patients who find such links disorganizing (e.g., patients with personality disorders), the connection to early experiences with caregivers is not always explicitly mentioned. Particularly when working with certain patients who find such links disorganizing (e.g., patients with personality disorders).

Therapist. Maybe so, but the question arises if your perception of those events, which is consistent with what we have discussed as your experience with your father, who you saw as critical, harsh, and competitive with you, is influencing how you are experiencing me, right now, here in this room, so that a question offered out of concern is similarly experienced as an attack and belittling.

Initially, transference interpretations such as these were considered the sin qua non of the psychodynamic approach (Freud, 1912; Strachey, 1934; Racker, 1968). Strachey (1934) referred to trans-
ference interpretations as mutative interpretations. Analysis of transference was seen as curative, and transference interpretations were seen as one of the distinguishing features of dynamic therapy. Bibring (1954) importantly noted that there is a hierarchy of principles guiding therapy, and that transference interpretations were the supreme agent in this hierarchy. Laplanche and Pontalis (1973) thought the transference represents psychoanalysis par excellence.

However, over the years, technical developments within psychoanalysis increasingly stressed the importance of the therapeutic relationship, ‘corrective emotional experience’, and ‘implicit relational knowing’ as agents of change (Gill, 1982; Loewald, 1960; Wallerstein, 1986, Lyons-Ruth et al., 1998; Sharpless & Barber, 2012). Further, early correlational research suggested that the use of transference interpretations was related to poorer outcome, particularly for the most seriously disturbed patients with personality pathology (Connolly, Crites-Christoph, Shappell, et al., 1999; Høglend, 1996; Piper, Azim, Joyce, & McCallum, 1991; Piper, Joyce, McCallum, & Azim, 1993; Ogrodniczuk and Piper, 1999). Taken together, this led many psychodynamic theorists to de-emphasize their use and importance. Some clinical writers went even further, arguing that transference interpretations were often experienced by patients as hostile and attacking, again, particularly by patients with personality pathology and disturbed relatedness (Waldinger & Gunderson, 1989; Bateman & Fonagy, 2004; Henry, Strupp, Schacht, Gaston, 1994; Frances & Perry, 1983). Nonetheless, other clinicians continued to see transference interpretations as a valuable clinical tool.

A question that is relevant to these different views regarding transference interpretations concerns their exact nature and timing. For example, one issue is whether early or late transference interpretations are more beneficial. Strachey (1934) posits that premature interpretations might result in a negative outcome, and that negative reactions to early transference interpretations may lead to dropout. However, Gunderson (1984), Davanloo (1980), Malan (1976b), Masterson (1978), Sîneuș (1972), and Kernberg (1975), all argue that early transference interpretations are beneficial.

Other issues surrounding transference interpretations include the interpretation of positive versus negative transference; the need for a strong working alliance for making transference interpretations, particularly with more disturbed patients (Winston, McCullough, & Laikin, 1993; Høglend, 1996; Bond, Banon, & Grenier, 1998); and responses to transference in healthy versus disturbed individuals. Gabbard (2006) suggested that healthier people can tolerate transference interpretations better than disturbed individuals with a personality disorder or low quality of object relations (QOR; see Piper et al., 1991; Ogrodniczuk & Piper, 1999).

Transference Interpretations in Psychotherapy Research

A number of psychotherapy studies have directly and indirectly examined transference interpretations. These studies include correlational studies that have looked at narratives, process research, outcome research, as well as quasi-experimental studies and experimental studies using RCTs. One of the earliest studies to examine transference interpretations was the Menninger project (Wallerstein, Robbins, Sargent, Luborsky, 1986). In this study, psychodynamic psychotherapy was compared with psychoanalysis for characterologically disturbed patients. Although the findings suggested that supportive techniques were ubiquitous in the treatments (Wallerstein, 1986), it was found that personality-disordered patients did poorly in psychoanalysis, but did even worse in the psychodynamic psychotherapy or supportive psychotherapy, and did best in an expressive or interpretive psychodynamic psychotherapy (Kernberg et al., 1972). Kernberg noted the importance of transference interpretations as a mutative agent for those patients with severe personality pathology. He noted that the patients did not fail apart as many writers previously contended, and that it was important to interpret negative transference, in particular, but also positive transference early in treatment. According to Kernberg (1989), negative transference should be interpreted as fully as possible early on in treatment. An example of an interpretation of negative transference follows:

**Therapist.** When you say that everyone in the world is an idiot, I realize that I might be included in that statement. I wonder if you think that some of the things I say are idiotic too, or at the very least not helpful, and that part of you sees me as an idiot or is concerned with my capacity as your therapist?

When interpreting positive transference, it is important to distinguish between distorted positive transference (e.g., idealization of the therapist) and modulated positive transference that is appropriate given the circumstances of treatment (e.g., the patient seeing the therapist as a helpful interested person). Kernberg (1989) suggests that modulated or realistic transference should not be interpreted. However, if the patient treats the therapist in an idealized way (e.g., seeing the therapist as omnipotent or unable of making a mistake), an interpretation should be made. For example, the therapist could simply state, “You treat me as if I can do no wrong.”

Consistent with Kernberg’s interpretation of the Menninger findings, a number of early studies found a positive relation between transference interpretation and outcome. In two studies, Malan (1976a; 1976b) found that increased transference interpretations were related to better long-term outcome. However, both studies were based on process notes and suffered from a number of common confounds representative of such studies, including non-blind ratings. Marziali and Sullivan (1980) replicated Malan’s finding and improved on one of his previous methodological issues by blindly rerating Malan’s data, but found that only nine of the 22 treatments contained a transference interpretation, and of the nine treatments, transference interpretations were only being used a small proportion of the time. Although transference interpretations were related to good outcome, they were occurring at low rates in only a few treatments, a finding that is consistent with the notion that transference interpretations were being de-emphasized during this time. Additionally, given the level of improvement in the context of low rates of transference interpretations, other variables that were not assessed could have been contributing to improved outcome in these treatments. In a second study, Marziali (1984) attempted to replicate Malan’s findings, this time using audiotaped sessions, rather than process notes, for a group of 25 patients, and found that transference interpretations once again generally predicted good outcome. However, a number of studies conducted by Piper and colleagues using brief dynamic therapy found contra-
dictory findings. Piper, Debbane, Bienvenu, de Carufel, and Gar-
rant (1986), in a sample of 21 patients undergoing brief dynamic
therapy, found that the average number of interpretations were 10,
but that only one-third of those were transference interpretations,
and only 5% of those were linked to early childhood, or what are
called genetic interpretations. There were no specific findings
regarding differences between transference and extra-transference
interpretations (interpretations about relationships outside the
therapy). McCullough et al. (1991) conducted a study of brief
dynamic therapy, in which four sessions for 16 patients were rated.
They found that transference interpretations followed by an affect-
ive response were related to better outcome, and transference
interpretations followed by a defensive response were related to
poorer outcome. However, defensive responses were 5 times more
likely than the affective ones, suggesting that, based on these
findings, transference interpretations might be related to poor
outcome. Thus, a number of early correlational studies found that
although transference interpretations were related to good out-
come, they were relatively uncommon and frequently lead to
defensive responses, which were related to poor outcome.

A number of later correlational studies found that transference
interpretations were related to poor outcome in psychotherapy.
Hoglund (1993) conducted a study of 43 patients treated in brief
dynamic therapy with high and low levels of transference inter-
pretations and found that transference interpretations were related
to less favorable outcome both at 2-year and at 4-year follow-up.
The first major study of that kind (Piper et al., 1991) examined 64
diagnostically mixed outpatients suffering from anxiety, depres-
sion, and personality disorders. Over 20 sessions of a brief dy-
namic psychotherapy, there was an inverse relationship between
the frequency and proportion of transference interpretations in
both the therapeutic alliance and the therapy outcome (general
symptoms and dysfunction according to patient self-report). This
finding was due to a significant correlation for those with high
QOR, indicating that the frequency and proportion of transference
interpretations had a stronger effect for patients who were health-
lier. In the upper quartile, that is, 10 or more transference inter-
pretations, only 25% of individuals recovered, whereas in the lower
quartile, two or less transference interpretations, 100% of individ-
uals recovered. In a follow-up study of these patients, Piper et al.
(1993) examined the correspondence of the transference inter-
pretation to the initial case formulation based on a Sampson and
Weiss approach (Curtis, Silherschatz, Sampson, Weiss, 1994).
They found that low to moderate concentrations of accurate inter-
pretations for those with high QOR led to better outcome. Thus,
the accuracy of the transference interpretation was also important
and interacted with aspects of patient characteristics such as QOR
to predict outcome.

A later study by the same authors (Piper et al., 1999) found that
transference interpretations were related to dropout in a time-
limited interpretive dynamic psychotherapy and that there was a
typical sequence by which this occurred. The pattern was charac-
terized by nine features, which included 1) the patient voicing
thoughts about dropping out of psychotherapy early in the session,
2) the patient expressing frustration, 3) the therapist then focusing
on the transference and 4) linking the patient’s concerns with other
relationships, 5) the patient resisting this focus on the transference,
6) the therapist persisting, 7) the patient continuing to resist, 8) the
patient reluctantly agreeing to return, and 9) then the patient never
returning. This pattern held true unanimously for the seven drop-
outs who had a high focus on transference and was less charac-
teristic for those dropouts with a moderate or low focus on trans-
ference.

Connolly, Crits-Cristoph, Shelton, et al. (1999) examined trans-
ference interpretations, as they were related to outcomes, and
found that high levels of early transference interpretations were
related to poor outcome for those with more interpersonal diffi-
culties. Schut et al. (2005) examined 14 patients treated in a
supportive expressive short-term dynamic therapy, and found that
interpretations were related to subsequent disaffiliative process for
those with avoidant personality disorder. Ryum, Stiles, Svartberg,
and McCullough (2010) examined the effects of transference
work, the therapeutic alliance, and their interaction as they were
related to interpersonal problems in a sample of 49 patients with
cluster C personality disorders who were part of an RCT of
short-term dynamic psychotherapy and cognitive therapy (Svart-
berg, Stiles, & Seltzer, 2004). The therapeutic alliance did not
predict interpersonal problems at treatment termination when con-
trolling for transference work, whereas less transference work
predicted a greater reduction in interpersonal problems at treatment
termination when controlling for the effects of the therapeutic alli-
ance. Post hoc analyses indicated that the therapeutic alliance did
predict interpersonal problems when not controlling for the effects
of transference interpretations. The transference work by therapeutic
alliance interaction was significant at treatment termination such that,
in the context of a weaker alliance, more of an emphasis on transfer-
ence work predicted smaller reductions in interpersonal problems.
Taken together, the results of these studies indicate that there might be
a technique by patient characteristic interaction and the context of that
interaction (e.g., strong vs. weak alliance) may be important. For those
with more interpersonal difficulties or cluster C problems, high levels
of transference interpretations may result in poorer outcome or less
reduction in interpersonal problems, particularly for those with
weaker alliance.

However, these findings regarding the relationship between the
use of transference interpretations, alliance, and poorer outcome
may be even more complicated than initially thought. The context
in which the transference interpretations are executed may be more
nuanced than indicated by merely noting the valence of the alli-
ance. There also may be important distinctions among patient
characteristics that have relevance for the use of transference
interpretations. Silberschatz, Fretter, and Curtis (1986), using the
Sampson and Weiss plan formulation model, found that the accu-
racv of the transference interpretations was related to improved
outcome in three patients. A later study (Cris-Christoph, Cooper,
& Luborsky, 1988) using a sample of 43 patients treated in a
moderate-length psychodynamic psychotherapy found a signifi-
cant and moderately strong relationship between accuracy of in-
terpretation and treatment outcome. This relationship between
accuracy and treatment outcome was not accounted for by error or
technique in alliance, decoupling accuracy from other indicators of
good psychotherapeutic technique. In the Vanderbilt II Study,
Henry and Strupp (1994) found that the effectiveness of transfer-
cence interpretations was dependent on the quality of interpersonal
process between the therapist and the patient. Interpretations em-
bedded in the context of hostility or blame resulted in worsening of
symptoms in the patient, regardless of the correctness, immediacy,
or emotional depth of the interpretation. More recent findings by
Høglend and colleagues (Høglend et al., 2006, 2008, 2011) also suggest a more complicated picture. They found evidence that a low- to moderate-level transference interpretation-based treatment, as compared with a no- to low-level transference interpretation-based treatment, led to better outcomes for those with low QOR. This was especially true for those with weaker therapeutic alliances. These findings will be discussed in greater detail in the following section.

In summary, transference interpretations may result in substantial increases in the patient’s ability to collaborate and in positive outcomes, but they also may produce marked decreases in collaboration and lead to worse outcomes. For this reason, Gabbard et al. (1994) referred to transference interpretations as a high-risk high-yield intervention. Clinical theory suggests that transference interpretations are most effective when a series of preparatory interventions by the therapist create a climate in which the patient can accept the therapist’s observation. These interventions are thought to provide a buffer for the patient from the sharpness of a transference interpretation. Along these lines, Geller (2005) notes that therapeutic tact is the capacity to tell patients something they do not want to hear in a manner in which they can hear it. Consistent with these clinical observations, the data indicate that patients with high levels of QOR may benefit from low to moderate levels of transference interpretations but have difficulty with high levels of transference interpretations. Patients with low levels of QOR may tolerate only low to moderate levels of transference interpretations; however, as will be presented later, these low to moderate levels of transference interpretations may be important for the treatment of those with low QOR. The accuracy of the interpretation as well as the process (e.g., accepting/nonjudgmental, nonhostile) in which the interpretation is embedded are also important and related to outcome, and this may be true particularly when the alliance is weak. The accepting/nonjudgmental nonhostile attitude as well as the use of preparatory comments and the tactful delivery of an accurate interpretation may foster the alliance such that poor alliances become stronger alliances.

**Transference-Focused Psychotherapies**

In recent years, a series of experimental studies suggest the possibility that transference interpretations may be a highly useful treatment tool, especially for patients who were thought to be most negatively affected by them, those with poor QOR and personality disorders. For example, the first author (KNL) and colleagues have found that transference focused psychotherapy (TFP; Clarkin, Yeomans, & Kernberg, 2006) for BPD was efficacious (Clarkin, Levy, Lenzenweger, & Kernberg, 2007; Levy et al., 2006), a finding now confirmed in an independent study (Doering et al., 2010). These findings are consistent with later studies by Per Høglend and colleagues (Høglend et al., 2006, 2008), who, using parametric techniques, examined a transference-based intervention versus a non-transference-based intervention. In some of these studies, there were improvements observed in comparison groups that did not include the use of transference interpretations, indicating that transference interpretations are not necessary for overall improvement. However, as will be discussed later, changes in structural variables such as reflective function (RF) and coherence were unique to the transference-based treatments.

**Transference-Focused Psychotherapy**

TFP (Clarkin et al., 2006) is a modified psychodynamic psychotherapy designed specifically for individuals with personality disorders, particularly BPD. It is a structured, twice-weekly, outpatient treatment based on Otto Kernberg’s object relations model. The primary goal of TFP is to reduce symptomatology and self-destructive behavior through the integration of representations of both self and other, what Kernberg called the resolution of identity diffusion or the accomplishment of identity consolidation. During the first year of treatment, the TFP therapist focuses on a hierarchy of issues, beginning with a treatment contract that is designed to provide containment of suicidal and self-destructive behaviors, and to articulate various ways that the patient may interfere with the treatment, as well as setting a collaborative agreement for which to understand deviations from the treatment contract. In session, the therapist follows the dominant affect of the patient and identifies and explicates the recapitulation of object relation dyads, or object relational patterns, as they are experienced and expressed in the here-and-now of the relationship with the therapist, which is conceptualized as a transference relationship, although it contains real aspects as well. The focus is not on reconstructing childhood experiences, but rather understanding how patients relate to their therapists, with the idea that this has relevance for thinking about how they relate to other people too. Techniques used in TFP are common to psychodynamic treatments and include the use of clarification of the patients’ subjective experience, confrontation by tactfully pointing out discrepancies between what the patient is saying and doing, and transference interpretations in which the therapist provides a timely, clear, and tactful interpretation of the dominant affect-laden themes in the patients enactments, in the here-and-now of the transference. These transference interpretations are hypothesized to integrate polarized self and other representations.

In a recent RCT (Clarkin et al., 2007; Levy et al., 2006), 90 patients with BPD were randomized to one of three treatments: TFP, dialectical behavior therapy (DBT; Linehan, 1993), and psychodynamic supportive psychotherapy (SPT; Applebaum, 2005). Patients were clinically referred, highly comorbid, highly traumatized, and had a strong history of engaging in self-injurious behaviors, as indicated by the fact that about one-third had begun self-injuring by age 12 and two-thirds had engaged in self-injurious behavior, all of which are rates consistent with the epidemiological and clinical data on BPD. Primary outcome variables included suicidality, anger, impulsivity, and assault. Secondary outcome variables included depression, anxiety, social adjustment, and global functioning. There were no significant between-group differences for primary outcome variables. However, there were significant within-group improvements in suicidality in both the TFP and the DBT cells, but not in the SPT cell; significant improvements in anger and impulsivity in the TFP and the SPT cells, but not in the DBT cell; and significant improvements in assault and irritability in only the TFP cell. With regard to secondary outcome variables, there were no significant differences between groups on depression, anxiety, social adjustment, or global functioning, although there were improvements in these variables for all three groups. In light of these findings, it appears that irritability, assault, anger, and impulsivity are all variables that are addressed effectively by TFP. Further, Levy et al. (2006) found...
that compared with SPT and DBT, TFP uniquely led to increases in structural variables such as metacognitive and social–cognitive capacities, as assessed through Mary Main’s adult attachment interview (AAI; Main, Goldwyn, & Hesse, 2002), coherence scores, and Peter Fonagy’s reflective function (RF) scores (Fonagy, Steele, Steele, & Target, 1998).

Another study of TFP, the Munich-Vienna Transference-Focused Psychotherapy Study (Doering et al., 2010), included 104 women diagnosed with BPD and randomized to 1 year of either TFP or treatment by experienced therapist in the community. They found that TFP was superior in reducing suicidality, suicide attempts, BPD symptomatology, Structured Clinical Interview for DSM-IV Axis II (SCID-II; First, Gibbon, Spitzer, & Williams, 1997) data, inpatient hospitalizations, and dropouts, and increasing global functioning. Both groups improved on anxiety and depression and global psychopathology, and there were no improvements in the groups in terms of parasuicidality. Additionally, patients in the TFP group experienced significantly greater improvements on the Structured Interview for Personality Organization (STIPO; Clarkin, Caligor, Stern, & Kernberg, 2004), a measure of overall personality functions that are important in the regulation of self and other representations. This particular finding is consistent with the structural improvements that were unique to TFP in the study by Levy et al. (2006).

Thus, the evidence from RCTs for TFP indicates that there are both statistically and clinically significant changes in severely disturbed BPD patients over 1 year of treatment, the changes are in both symptoms and social cognition, and that social cognition changes appear to be unique to TFP. Further, the findings have been replicated in independent samples.

It is important to note that the comparison groups in these RCTs did not include the use of transference interpretations and that these groups also had significant improvements in some symptoms and functioning. In particular, the SPT group from the study by Clarkin et al. (2007) had large effect sizes for social and occupational functioning, as measured by the Social Adjustment Scale (Weissman & Bothwell, 1976), although they did not differ significantly from the other groups. In light of these findings, it appears that the use of transference interpretations may not be necessary for overall improvement. However, it is important to stress that, contrary to correlational findings that suggested transference interpretations were related to worse outcome and possibly harmful, findings from experimental studies indicate that transference interpretation-based treatments, such as TFP, are as efficacious as other nontransference interpretation-based treatments (e.g., DBT, SPT). Additionally, TFPs may uniquely result in structural change in terms of reflective function, coherence, and overall personality organization (as assessed with the STIPO; Clarkin et al., 2004). However, further research is needed to confirm whether these findings can be replicated and are especially true for those with BPD and low QOR. Furthermore, it will be important to determine whether the use of transference interpretations in session is directly related to these outcomes, and whether other processes, such as the therapeutic alliance, could be contributing to the outcomes as well.

One important question to ask is whether these structural variables (e.g., reflective function, coherence, personality organization) are related to outcomes in terms of symptoms and functional improvement? In fact, there is evidence that supports a link between structural and functional improvement. Blatt et al. (1996) have shown that structural improvements in the representations of self and other among adolescent inpatients were significantly related to overall clinical improvement, as measured by the Global Assessment Scale (Endicott, Spitzer, Fleiss, & Cohen, 1976). Other evidence has shown that reflective function is related to neurocognitive functioning in patients with borderline personality disorder (Levy et al., 2005), such that lower RF is related to higher levels of impulsivity and higher RF is related to fewer errors on the Wisconsin Card Sorting Task. There is also evidence that RF and coherence in adults are related to the strange situation behavior of their infants (Fonagy et al., 1991), indicating that, at the very least, improvements in structural variables such as RF and coherence might lead to secure attachment in one’s children. However, these structural variables have not been tied to functional outcomes in RCTs, with the exception of the findings discussed later by Höglend and colleagues (Johansson et al., 2010), who found that the relationship between a transference interpretation-based treatment and improved interpersonal outcome was mediated by improvements in insight. It will be important for future research to examine this relationship in an effort to determine the value of structural improvements.

First Experimental Study of Transference Interpretations

Another investigation of transference-focused psychotherapies occurred in the First Experimental Study of Transference Interpretations (FEST; Höglend et al., 2006) study, an RCT that compared a group of patients who received a dynamic treatment with low to moderate levels of transference interpretations (the transference group) and a group of patients who received a dynamic treatment with no transference interpretations (the comparison group). One hundred patients referred from general practitioners, outpatient departments, and independent practice and with mixed disorders, depression, anxiety, and personality disorders (46%) were randomized to one of these two dynamic therapies. Regardless of condition, the treatment was once weekly for 40 sessions, and one particular strength of this study is having the same therapists carry out the treatment in both conditions. Process ratings to check the treatment integrity in 447 sessions found that transference interpretations occurred 1.7 times ($SD = .7$) in the transference group and only .1 times in the comparison group. Importantly, extratransferential comments, supportive comments, and general skill of the therapists were equal in both conditions. No other differences were found in these rated variables, other than the number of transference interpretations. Primary outcome variables for this study included the Psychodynamic Functioning Scale (PFS; Höglend et al., 2000), Global Assessment of Functioning Scale (GAF; Spitzer, Williams, Gibbon, & First, 1990), the Inventory of Interpersonal Problems-Circumplex version (Alden, Wiggins, & Pincus, 1990), and the Global Severity Index from the Symptom Checklist-90-revised (SCL-90-R; Derogatis, 1983). The PFS includes subscales that assess the quality of friendships and family relations, romantic/sexual relations, tolerance for affects, insight, and problem-solving and adaptive capacity. There were no significant between-group differences at pretreatment, midtreatment, or post-treatment. Both groups had statistically significant change from pre- to post-treatment, with large effect sizes for all primary outcome variables.
However, consistent with previous findings with regard to TFP, improvements on two of the outcome scales indicate that a transference interpretation-based treatment was particularly useful for personality-disordered patients with poor QOR.

Moderator analyses examined changes in primary outcome variables for those with high and low levels of QOR (Høglend et al., 2006). There were no significant differences between treatment groups for those in the high QOR group on any of the primary outcome variables. However, for those in the low QOR group, the percentage of cases with clinically significant change on the PFS and GAF scales was twice as high for the transference group, compared with the comparison group. Patients with low QOR experienced more improvements on these primary outcome variables when transference interpretations were included as part of their treatment. Furthermore, these findings held up over follow-up periods of 1 and 3 years (Høglend et al., 2008).

Using the same sample, these findings were explored further in a more recent study by Høglend et al. (2011), which found that for patients in the low QOR group, being in the transference interpretation group had a more positive impact on psychodynamic functioning in the context of a weaker therapeutic alliance. For patients with a stronger therapeutic alliance and high QOR, being in the transference interpretation group had a more negative effect (although these patients did well too, just not as well as those with low QOR in the transference interpretation group). These disparate outcomes on the PFS between those with high and low QOR with regard to strong and weak alliance help to shed light on previous findings. For example, Ryum et al. (2010) found that the use of transference interpretations within the context of a weak therapeutic alliance was related to poorer outcome. This general finding, however, may not be valid for those with low QOR and is consistent with the findings of clinical theorists who suggest that transference interpretations are particularly useful in the treatment of those with severe personality disorder who tend to be characterized by low QOR (e.g., Kernberg, 1975).

Another way of understanding the Høglend et al. (2011) finding in the context of recent meta-analyses (Horvath, Del Re, Flückiger, Symonds, 2011; Flückiger, Del Re, Wampold, Symonds, Horvath, 2012), which have shown that the therapeutic alliance is related to a range of positive outcomes across large patient populations, is in relation to the concept of therapist responsiveness (Stiles, Honos-Webb, & Surko, 1998). Specifically, therapists may be addressing poor alliance through a number of techniques (e.g., treatment tactics such as verbal techniques and the selection of topics, or moment-to-moment responsiveness such as making adjustments in the wording of interpretations), and through this responsiveness, patients may be achieving a stronger alliance. However, as we indicated earlier, to make stronger statements about the role of transference interpretations and their specificity, it will be important to replicate and extend the findings from Høglend et al. (2011) to establish that transference interpretations are indeed directly related to positive outcomes when controlling for other reasonable alternative explanations such as the role of therapeutic alliance or some other technical intervention.

Interestingly, Høglend and Gabbard (2012) reported the correlations between the number of transference interpretations and outcome within the high and low QOR groups. The number of transference interpretations was significantly negatively correlated with outcome on the PFS ($r(24) = -0.40, p < 0.05$). Transference interpretations were also negatively correlated with the other outcome scales, although less strongly and not significantly. This finding is consistent with earlier correlational studies (Piper et al., 1991; Høglend, 1993) and suggests that more transference interpretations are not necessarily better. Instead, there seems to be an optimal range of low to moderate levels of transference interpretations (1–3 per session) that results in good outcome on the PFS.

Høglend and colleagues’ findings also point to the added value of using experimental designs in psychotherapy research. It is possible that these earlier studies (Piper et al., 1991; Høglend, 1993) would have found a favorable relationship between transference interpretations and outcome if the data were examined using parametric designs, quadratic regression, and/or multilevel modeling techniques, rather than with simple bivariate correlations. Doing so protects against complex relationships that can mask important findings. Thus, findings of Høglend et al. (2006), if replicated, would suggest that simple correlational analyses may have masked the value of low to moderate levels of transference interpretations. Perhaps no/low levels of transference interpretations ($M = 0.1$) are not useful for those with low QOR, whereas moderate levels ($M = 1.7$) result in better outcome, and high levels of transference interpretations are related to poorer outcome for those with either high or low QOR. It is also important to note that Høglend et al.’s component control FEST study is technically a dismantling design and not strictly a parametric design because it does not have three or more levels (Behar & Borkovec, 2003). However, the between-group findings of the FEST study, combined with the correlational findings from their study and previous research in which the number of transference interpretations were related to worse outcome, could be interpreted in line with a parametric design.

Høglend and colleagues (Johansson et al., 2010) have addressed the question of why or how transference interpretations work in subsequent analyses. They used a mediated moderation model to test whether the mechanism in a transference interpretation was actually a development of insight. They found that receiving a transference interpretation-based treatment versus receiving a non-transference interpretation-based treatment was related to interpersonal outcome at 3-year follow-up. Additionally, being in a transference-based treatment was related to increases in insight in patients, which in turn was related to increases in interpersonal functioning at a pre-3-year follow-up. Further, when controlling for insight, the relationship between transference-based treatment and interpersonal functioning disappears, indicating that the mechanism by which a transference-based treatment may be making an impact, and may lead to better outcome, is through the development of insight. Transference interpretations, if shown to be directly related to outcome in future studies, may help patients to think about themselves in relation to others with greater insight, and in turn will help them function better in the world.

**Conclusions/Summary**

Good experimental evidence has emerged from both social psychological findings and psychotherapy process and outcome research that supports the concept of transference. To reiterate our earlier definition of transference, we conceptualize it as a tendency in which representational aspects of important and formative relationships (such as with parents and siblings) can be both...
consciously experienced and/or unconsciously ascribed to other relationships (Levy, 2009). This conceptualization of transference is consistent with what is known about schemas and pattern matching, implicit memory processes, and other concepts from cognitive and neuroscience. However, there is also some burgeoning evidence that transference is not just a cognitive-information bias or process, but that it is also a dynamic process related to attachment and defensive processes. Nevertheless, these findings are preliminary and need to be confirmed in future research. Early correlational data from psychotherapy research found that transference interpretations were negatively correlated to outcome, particularly for those with low QOR. Some data suggest that the accuracy and competent delivery of interpretation as well as the correspondence of the interpretations to the therapist’s treatment plan predict good outcome (Crits-Christoph et al., 1988). This held true, even after controlling for alliance, and the hypothesis that accuracy would have its greatest impact in the context of good alliance was also not confirmed. Future research should not only be concerned with controlling for the alliance but also with controlling for other indicators of good psychotherapy skill.

For a number of years now, the prevailing view has been that transference interpretations are potentially harmful, that one needs to establish a strong working alliance first, and that they are not useful early in treatment. Most importantly, transference interpretations were seen as particularly problematic for those with low QOR or personality disorders. However, more recent experimental data using RCT designs suggest that low to moderate levels of transference interpretation and transference-based treatments are highly effective, and that these treatments can lead to structural change, or change in social–cognitive processing. In fact, this is especially true for those with low QOR and personality disorders. Still, some clinical theorists have been reluctant to endorse the use of transference interpretation for those with personality disorders (Applebaum, 2005; Gunderson, 2009; Bateman & Fonagy, 2007). Even so, these theorists advocate that therapists attend to, follow, and work within the transference even if no explicit mention of it is made. Thus, despite continued controversy, most dynamic theorists recognize the importance of transference and attending to it. Further, emerging empirical evidence from two RCT studies (Høglend et al., 2006; Clarkin et al., 2007) suggests that transference-based treatments that contain low to moderate doses of transference interpretations can be well tolerated by those with borderline personality disorder as well as those with personality disorders and/or low QOR, even in the context of a weak therapeutic alliance. These findings, using observer-rated measures such as the AAI (i.e., RF, coherence), STIPO (personality organization), and GAF (global symptoms and functioning), as well as clinician-rated measures such as the PFS (quality of romantic/sexual relationships and friendships, tolerance of affects, insight), suggest that, although not necessary, transference interpretations may be clinically useful for good outcome when working with these patients. Of course, as stated earlier, these findings need to be replicated and confirmed in future research, particularly research that links specific interventions to outcome when controlling for relevant constructs (such as alliance). Additionally, we tentatively suggest that transference interpretations may be specifically related to structural change (Levy et al., 2006; Doering et al., 2010; Høglend et al., 2006), as assessed by the AAI, STIPO, and PFS.

This hypothesis should be confirmed in process research that relates transference interpretations to structural change while controlling for reasonable alternative hypotheses. Finally, the findings reviewed also stress the importance of experimental investigations and advanced statistical modeling rather than relying solely on simple correlational designs (Crits-Christoph et al., in press). Future research is needed to confirm, disconfirm and extend these findings to best inform clinical practice.

References


Main, M., Goldwyn, R., & Hesse, E. (2002). *Adult attachment scoring and classification system*. Unpublished manuscript, Department of Psychology, University of California at Berkeley, Berkeley, California.


Received May 10, 2012
Revision received June 3, 2012
Accepted June 4, 2012