At its foundation, attachment theory (AT) (Bowlby, 1969) is a theory of developmental psychology that uses evolutionary and ethological frameworks to describe how the caregiver-child relationship emerges and how it influences subsequent social, emotional, and cognitive development. And while AT emerged out of observations of child-caregiver dynamics, it was quickly and readily generalised to address similar psychosocial phenomena within adult romantic relationships (e.g., Hazan & Shaver, 1987, 1994). Betrayal trauma theory (BTT) (Freyd, 1994, 1996), building on the most central concepts of AT, has focused very specifically on understanding psychological responses to trauma. Like AT, BTT proposes that trauma occurring within the context of an attachment relationship is qualitatively different than trauma that takes place outside of one. Also as with AT, BTT was first developed with the child-caregiver relationship in mind but has since been applied to other adult relationships, including not only romantic relationships but hierarchical relationships (such as that between an employer and an employee, or an institution and its member) as well (e.g., Freyd, 1996; Smith & Freyd, 2013).

Regarding the specific circumstance of maltreatment or traumatisation by an attachment figure, both AT and BTT make specific predictions about how humans adaptively respond. The following paper aims to describe where and how these predictions overlap, and where they differ. More specifically, we will argue that the significant theoretical concordances include:

1. A central assumption that humans have evolved a strong motivation to maintain affectional bonds with close others.
2. The rationale that it is adaptive to defensively exclude knowledge of and/or selectively process experiences of maltreatment by a caregiver, as complete

*The use of the term “caregiver” in this article refers to the child’s primary caregiver or attachment figure.
-awarement entails acknowledgement of information that may threaten established attachment representations.
3. The idea that adaptive short-term responses to relational trauma can become problematic later in life.

Differences between the two theories include:

1. A focus within BTT on what is known as the “cheater-detection system”, and its interaction with the attachment system.
2. Slightly divergent emphases on the roles of explicit vs. implicit caregiver threats of abandonment in motivating defensive processing or what is termed “betrayal blindness”.
3. Different conceptualisations of the role of fear in promoting dissociation in the context of abuse by a caregiver or depended-upon other.

For these latter two issues, we suggest future studies to explore these differences collaboratively.

Theoretical convergences

Biologically based human motivation to form and maintain attachment relationships

AT and BTT both argue that forming and maintaining close affectional bonds is an essential component of the human condition. A core tenet of AT is that humans are born biologically pre-programmed to form attachments with others and to seek close contact (proximity seeking) with protective others when threatened or upset (Bowlby, 1969; Hazan & Shaver, 1987, 1994; Simpson & Rholes, 1994) as cited in Hesse and Main, (2000). In Bowlby’s view, the attachment behavioural system is more significant to human life than are feeding and sexual behaviours, and peoples’ attachments to others are the hubs around which their lives revolve (Bowlby, 1980).

Like AT, BTT shares the assumption that humans are intrinsically and highly motivated to form and preserve attachments with those upon whom they depend—not only for their psychological and emotional well-being, but for their very survival. Indeed, Freyd (1996) has asserted that to ensure their own survival, children need to and must align themselves with their caregivers. Out of this basic assumption, Freyd’s BTT (1996) lays out an evolutionary rationale for why and how humans—who are usually excellent detectors of betrayal (Cosmides & Tooby, 1992)—may fail to detect betrayal under circumstances in which detection could be counterproductive to survival. She has written, “if betrayal by a trusted caregiver is the key to predicting amnesia for abuse, attachment is the key to understanding why amnesia is adaptive in instances of such betrayals” (Freyd, 1996, p. 69).
Importantly, because both AT’s and BTT’s conceptualisation of the human response to trauma within an attachment relationship focuses on describing evolutionarily adaptive responses to negative relational experiences, both theories depathologise the behavioural and psychological responses to trauma. Further depathologising these responses are the theories’ shared emphasis on relational trauma that places the origin of the traumatic response in the traumatic event rather than as maladaptive symptoms residing within the individual.

**Adaptive responses to betrayal/child abuse by a caregiver**

**Adaptive cognitive responses: betrayal blindness/defensive processing and amnesia**

Central to BTT is the idea that the degree to which a trauma involves betrayal by another person significantly influences the traumatized individual’s cognitive encoding of the experience of trauma, the accessibility of the event to awareness, and the psychological as well as behavioral responses. (Freyd, 1996, pp. 9–10)

The theory holds that the closer and more necessary one’s relationship is to the perpetrator(s), the greater the degree of betrayal involved. Under conditions where betrayal is strongest, victims may experience “betrayal blindness” in which the betrayed person does not have conscious awareness, or memory, of the betrayal. Within this model, betrayal blindness serves the important and adaptive function of allowing individuals to maintain needed attachment relationships with their perpetrator(s). Consistent with this proposition, research has shown that even after controlling for age of abuse onset and abuse duration, the caregiver status of the perpetrator predicts survivors’ self-reported memory impairment (i.e., “I now remember basically what happened but I didn’t always”) for physical and sexual abuse experiences (Freyd, DePrince, & Zurbriggen, 2001).

According to BTT, “the cognitive mechanisms that underlie the blockage of information (from consciousness) are dissociations between normally connected, or integrated, aspects of processing and memory” (Freyd, 1996, p. 115). Possible forms of dissociation underlying this phenomenon include those which occur early in information processing that prevent information from inclusion and storage in episodic or narrative memory, and those which occur later in the process after initial event processing, encoding, and storage in memory. This latter form of betrayal “amnesia” can range widely in severity, with pervasive and profound amnesia residing on one end of this continuum, partial amnesia falling in the moderate range, and relatively complete and continuous recall coupled with shifting interpretations (delayed understanding) of these experiences representing milder forms of amnesia. Furthermore, Freyd proposes that even when amnesia is absent, at least some degree of information blockage is necessary (Freyd, 1996, p. 76). This may take the form of shame, where blame for the
maltreatment is internalised. Though shame is much more subtle than amnesia, it, too acts to block from awareness the reality of the perpetrator’s culpability (e.g., Platt & Freyd, 2012).

BTT’s basic propositions of betrayal blindness and amnesia are wholly consistent with Bowlby’s proposal of “defensive processing”—a subject about which Bowlby wrote at length in his book, Attachment and Loss (1980). More specifically, Bowlby posited that information that threatens the representational model of our attachment relationship can be “defensively excluded” from awareness (i.e., excluded from further processing for relatively long periods or even permanently). Like Freyd did decades later, Bowlby hypothesised that defensive exclusion can take place at a number of different points in the course of information processing (Bowlby, 1980, pp. 44–45), and when information is already stored in long-term memory, “defensive exclusion results in some degree of amnesia” (Bowlby, 1980, p. 46). Placed in context, Bowlby theorised that the mechanism by which early attachment experiences influence subsequent development involves the internal working models (IWMs) children construct of their social world and of themselves as agents within it. Because IWMs are constructed with input from multiple sources, data has the potential to be incompatible (to illustrate this possibility, Bowlby (1980) gives the example of a child whose parent insists that they love him, while his firsthand experience of his parent’s behaviour suggests the reverse). According to Bowlby (1980), children who regularly or persistently experience such incompatibilities are faced with a dilemma, and may respond by: (1) giving some credence to both sets of information and oscillate uneasily between models, or (2) excluding the incompatible information from processing—an outcome that Bowlby thought to be quite common and expectable. In fact, Bowlby argued that when new information is incompatible with established models:

it is the models which win the day—in the short run almost always in the long run very often . . . To dismantle a model which has played and is still playing a major part in our daily life and to replace it by a new one is a slow and arduous task, even when the new situation is in principle welcome. When the new situation is by contrast unwelcome the task is not only arduous but painful and perhaps frightening as well . . . Certain situations that are both new and unwelcome may, indeed, appear at first sight so appalling that we dread to recognize their very existence. As a result we postpone evaluating them in their true proportions and fail to frame plans to meet them . . . For, should it prove so, we should be faced with the task of replacing existing models with new ones in circumstances in which the change is wholly unwelcome. (Bowlby, 1980, pp. 230–231)

This model of defensive processing, when applied to the context of child abuse by a trusted caregiver, is analogous to the phenomenon of betrayal blindness proposed by BTT—that is, we are very likely to ignore betrayal trauma (which is likely both unwelcome and very different from the other kinds of interactions we have with a
close depended upon other) so as to preserve our existing representational model of our loved one, and continue behaving in ways that elicit care and nurturance.

In addition to the theoretical agreement upon the basic existence of betrayal blindness/defensive exclusion as an adaptive cognitive phenomenon, both BTT and AT propose that being a young age at the time of the abuse makes its occurrence more likely.\(^3\) Indeed, Bowlby (1980) wrote that:

there is reason to suspect the vulnerability to conditions initiating defensive exclusion is at a maximum during the early years (roughly the first three) of life . . . (and) diminishes during later childhood and early adolescence. (p. 72)

Nevertheless, he predicted that this diminishment occurs only slowly and remains comparatively high throughout most of those years. There is probably no age at which human beings cease to be vulnerable to factors that maintain or increase any defensive exclusion already established. (p. 72)

Freyd (1996) too has asserted that while early betrayal trauma may be a particularly powerful precipitant of betrayal blindness, those who first experienced betrayal trauma in adulthood remain susceptible. For example, in the case of adult survivors of domestic abuse by an intimate partner, women may be reluctant to leave because “batterers often become more violent when women leave the relationship, the woman may be financially dependent on the batterer, and the woman may believe she has no choice” (Freyd, 1996, p. 191). In this situation where escape is not an option, women might have to remain blind to the betrayal in order to survive. Indeed, research has shown that women who are abused by their romantic partners often idealise their abuser (Douglas, 1987; Dutton & Painter, 1981, 1993), which may reflect more moderate forms of betrayal amnesia characterised by shifting interpretations of the abuse experiences (an example of which is provided in Note 2).

What is implicit in the above description of adaptive cognitive responses to betrayal is the assumption that betrayal blindness or defensive processing is advantageous because awareness or recognition of betrayal might change an individual’s behaviour towards a depended-upon perpetrator in a way that could threaten their attachment relationship.

**Behavioural responses: maintaining attachment behaviours as adaptive**

Attachment theorists identify secure attachment behaviours as those that help individuals achieve proximity to and care from their attachment figure under stress and aid in his or her protected exploration of the environment. In infancy and childhood, these behaviours include reaching, smiling, laughing, vocalising, clinging, and crying. Bowlby (1969) hypothesised that adults are hard-wired to respond to these bids, and, as research has confirmed,\(^4\) find it rewarding to do so. Thus, when babies smile at their caregivers, or when children come and sit in

---

Rosemary E. Bernstein and Jennifer J. Freyd
their laps, for example, most caregivers respond with positive feelings of love and
tenderness towards the child that makes them more likely to provide the child
with the care and protection they need for survival and healthy development.
Freyd (1996) summarised the utility of this transactional interplay of attachment
and caregiving behaviours in her assertion that,

because attachment is of overwhelming significance, a complex system of emotional,
cognitive, and behavioral components is operative during the child’s development
(to) ensure attachment: children love their caregivers, and that love motivates the
children to display affection towards their caregivers, which in turn elicits love,
nurturing, and protection from the caregivers. (p. 71)

In the case of childhood betrayal trauma, where children are dependent on their
perpetrating caregiver(s) and escape is not a (perceived if not real) option, a
child’s detection of betrayal would be likely to cause him or her to distrust their
caregiver(s), and may prompt the child to inhibit proximity-seeking behaviours.
This inhibition of attachment behaviour, however, may very well make the situa-
tion worse by further alienating the caregiver and thus placing the child in danger
of more abuse and less love or care. It is thus essential to the child’s survival that
they not stop behaving in a way that inspires attachment in the face of caregiver
abuse. Indeed, incest survivors, for example, have been found to be more attuned
to their offending parents’ needs (Levang, 1989), presumably so as to appease
them and maintain the attachment relationship. Attachment theorists Hesse and
Main (2000) explain this phenomenon in evolutionary terms, arguing that since
children have no haven of safety beyond their attachment figure(s), those whose
attachment figures are abusive must nonetheless continue to approach and solicit
proximity from them. More specifically, they have argued:

(W)hile at first glance an inherent propensity to approach the location of (threat)
appears irrational, it may be helpful to recall an observation made by Darwin
regarding the Galapagos sea lizard, an animal able to move about with ease on both
land and sea, but which exhibited the peculiar behavioral feature that, when fright-
ened, it would not enter the water. To further his understanding of this phenome-
on, Darwin repeatedly threw one of these lizards into the water. Strikingly,
although possessed “of perfect powers of diving and swimming”, the sea lizard
invariably returned in a direct line to the spot where its attacker stood. Darwin spec-
ulatively solved this conundrum by consideration of the animal’s evolutionary
history: “Perhaps this singular piece of apparent stupidity may be accounted for by
the circumstance, that this reptile has no [natural] enemy whatever on shore,
whereas at sea it must often fall a prey to the numerous sharks. Hence, probably,
urged by a fixed and hereditary instinct that the shore is its place of safety, what-
ever the emergency may be, it there takes refuge” (Darwin, 1839) . . . Turning our
attention now to human phylogeny, we can see that like the sea lizard an infant
(threatened) by its attachment figure has no inherent (i.e. instinctively organized)
means for separating the location of its attacker from the location of its haven of safety. Thus, confronted with circumstances unanticipated within its evolutionary history, it should experience strong propensities to approach the place of threat. (Hesse & Main, 2000, p. 1104, my italics)

Although this prediction has some support, the human response is more complicated. The human attachment literature has demonstrated that while maltreated children continue to be attached to and seek proximity to their abusive caregivers, such children often display atypical behaviours toward their perpetrating parent not seen in most non-maltreated children. The childhood attachment behaviours that have been most frequently associated with parental abuse are the various organisational anomalies characterising infant-caregiver attachment disorganisation. Though it was not original to Bowlby’s writings (1969, 1980) or Ainsworth and Bell’s (1970) early classification models of attachment (which included secure, avoidant (which Bowlby called “assertion of independence of affectional ties”) and anxious ambivalent variants), the disorganised/disoriented (D) attachment pattern observed in a child towards his or her primary caregiver was introduced in the late 1980s as the fourth major category of infant attachment to describe attachment behaviours seen in maltreated children (Carlson, Cicchetti, Barnett, & Braunwald, 1989). As with the other classifications, disorganised attachment patterns are not related to infant temperament, and often appear in respect to only one attachment figure (Hesse & Main, 2000). The specific anomalous behaviours seen in children in this group upon reunion with their primary attachment figure in the Strange Situation Procedure are idiosyncratic and diverse, and most commonly include:

1. The sequential or simultaneous display of contradictory (e.g. approach and avoidance) behaviour.
2. Undirected, misdirected, incomplete, and interrupted movements and expressions.
3. Stereotyped, asymmetrical, or mistimed movements or anomalous postures.
4. Freezing, stilling, and slowed movements and expressions.
5. Direct indices of apprehension regarding the parent.
6. Direct indices of disorganisation, disorientation, and confusion (Blizard, 2006; Liotti, 2000; Main & Solomon, 1990).

At first glance, it may seem as though this type of anomalous behaviour is incompatible with what we would expect based on the core tenet of BTT—that is, that children remain blind to or forget betrayal so that they can continue to behave normally toward their parent and avoid acting in ways that might put the child at further risk. These behaviours no longer seem incongruent with BTT when we consider the fact that for the majority of the time infants categorised as disorganised in relation to their primary attachment figure show behaviour which is organised in predictable ways intended to maintain proximity to the caregiver. Indeed, manifestations of disorganised infant behaviour during the strange situation (Ainsworth & Bell,
1970) that are sufficient for assignment to the disorganised/disoriented attachment category are often subtle, difficult to observe, and quite brief—often limited to a single episode lasting between ten and thirty seconds (Hesse & Main, 2000; Thompson, 2008). Children's behaviour outside of these episodes is categorised into a best-fitting alternate secondary attachment pattern (i.e., secure, avoidant, or anxious–ambivalent). The fact that these episodes are often called strategy “lapses” or “breakdowns”, drives home the reality that they are a temporary departure from a particular child's typical behaviour. Thus, it appears that in the face of betrayal, children are, for the most part, able to continue displaying the behaviours that they have learned are most successful for maintaining proximity to their primary attachment figure (i.e., secure attachment behaviours for children whose primary caregiver is responsive, avoidant behaviours for children whose primary caregiver is rejecting, and anxious–ambivalent behaviours for children whose primary caregiver is inconsistently responsive).

That said, even the temporary and anomalous lapses of strategy seen in children with disorganised attachment relationships are consistent with BTT. According to Freyd (1996), dissociation does not necessitate that traumatic information be entirely blocked from entering the nervous system—just that the information is blocked from entering systems that control attachment behaviour. Because peoples' control of social interactions typically involves consciousness, it is often adaptive for the traumatic information to be blocked from consciousness. Importantly, however, while selective attention strategies may effectively block traumatic information from entering conscious awareness and/or declarative memory stores, it likely allows for some degree of unconscious sensory and/or procedural processing. Thus, even in cases where amnesia for betrayal is strong, it remains possible—and perhaps likely—that at an unconscious level and without understanding the basis for it, the abused person comes to fear the abuser and/or the context of the abuse. This fear may, in turn, motivate certain behaviour, such as avoidance of certain people or situations. Seen from this conceptual framework, the anomalous disorganised behaviours seen in young children who have been abused by a primary caregiver could be seen as brief behavioural manifestations of unintegrated and implicit intact sensory and affective memories of betrayal.

**Developmental predictions**

A third major convergence between AT’s and BTT's conceptualisations of trauma within the context of attachment relationships (i.e., betrayal trauma) is their shared view that when persistent, the very dissociative strategies that are adaptive while abuse is ongoing (i.e., betrayal blindness, defensive processing) are likely to be problematic when they emerge in the context of non-abusive relationships, or as a more general pattern of mental disconnection or dissociation. To convey this cost benefit tension of dissociative coping, Freyd has referenced childhood sexual abuse survivor Lillian Green (1992), who wrote,
as children in abusive families, we tried to minimize the abuse and earn the care-taking we needed by complying with our parent’s demands, both spoken and non-verbal. To satisfy them, we assumed the characteristics which became our roles. To avoid abuse and neglect, we suppressed or distinguished what didn’t fit these roles. Parts of us went underground, disconnecting from our external selves, and remaining undeveloped. Splitting ourselves in hidden pieces enabled us to survive, but it cost us dearly. (Green, 1992, p. 131, as quoted by Freyd, 1996, p. 77)

Indeed, in the twenty years that have passed since Green’s book was published, research has demonstrated that these costs are wide ranging, and include both inter- and intrapersonal difficulties.

With regards to adolescent and adult interpersonal functioning, both BT and ATT contend that early betrayal trauma can influence quality of future romantic partnerships. More specifically, BTT posits that the betrayal blindness that allows those who have been betrayed to survive may place them at risk for future victimisation (Gobin & Freyd, in press; Zurbriggen & Freyd, 2004). Gobin (2011) provided at least partial support for this hypothesis when she found that betrayal traumatisation influences romantic partner preferences such that young adults who experienced high betrayal trauma in childhood rated loyalty as a less desirable trait in a potential romantic partner than those who did not, and those who experienced high betrayal trauma in both childhood and adulthood reported a higher tolerance for verbal aggression in a potential mate. Though this study did not examine the separate influences of betrayal blindness and betrayal trauma itself, BTT proposes that betrayal blindness mediates the associations she observed. Research in the attachment field has identified related results, demonstrating, for example, that adults with a disorganised/unresolved state of mind with regard to abuse have the most acrimonious romantic relationships (Crowell, Treboux, & Waters, 2002).

It has been established that among the long term intrapersonal ramifications of betrayal traumatisation are pathological dissociation, inexplicable somatic symptoms (e.g., intermittent paralysis), shame, depression, and substance abuse—all of which are at least marginally related to the concept of blindness or unawareness (Goldsmith, Freyd, & DePrince, 2012; Martin, Cromer, DePrince, & Freyd, 2013; McNally, Perlman, Ristuccia, & Clancy, 2006). Attachment theorist Liotti (1999, 2004) has similarly theorised that infant attachment disorganisation (which, as described above, can be understood as a dissociative phenomenon) predisposes individuals to pathological dissociation later in development. In support of this idea, we know that infant–mother attachment disorganisation, alone (Ogawa, Sroufe, Weinfield, Carlson, & Egeland, 1997) or in combination with substantial intervening trauma (Carlson, 1998) is highly predictive of dissociative behaviour in middle childhood and late adolescence. Beyond predicting severe dissociation, infant attachment disorganisation is the attachment pattern most strongly related
to later mental health difficulties (Kobak, Cassidy, Lyons-Ruth, & Ziv, 2006) and some researchers have gone so far as to consider it one of the earliest measurable risk factors for a maladaptive developmental trajectory (Sroufe, Carlson, Levy, & Egeland, 1999; van IJzendoorn, Schuengel, & Bakermans-Kranenburg, 1999). For example, infant–caregiver disorganisation has been linked to later anxiety, depression, peer aggression, and disruptive behaviours (Lyons-Ruth, Bronfman, & Parsons, 1993; Lyons-Ruth, Zoll, Connell, & Grunebaum, 1989; Munson, McMahon, & Spieler, 2001; Shaw, Keenan, Vondra, Delliquadri, & Giovannelli, 1997; Shaw, Owens, Vondra, & Keenan, 1996; Sroufe, Carlson, Levy, & Egeland, 1999). Likewise, the disorganised/unresolved attachment pattern during adolescence and adulthood has also been linked to concurrent so-called mood and personality disorders, substance use disorders, eating disorders, schizoaffective disorder, and schizophrenia (Allen, Hauser, & Borman-Spurrell, 1996; Fonagy et al., 1996; Nakashi-Eisikovits, Dutra, & Westen, 2002; Tyrell & Dozier, 1997). In support of the idea that long-term emotional difficulties are a product of the dissociative response to betrayal trauma rather than the mere trauma itself, many of the studies that have considered attachment disorganisation and maltreatment or trauma experience within the same model find that attachment disorganisation partially, if not fully, mediates the relationship between early abuse by a caregiver and later psychosocial difficulties (Dutton, Saunders, Starzomski, & Bartholomew, 1994; Magdol, Moffitt, Caspi, & Silva, 1998; Styron & Janoff-Bulman, 1997).

It should be noted in this discussion of developmental predictions that both theories postulate that the relationship between betrayal blindness/pervasive defensive exclusion and later emotional difficulties is moderated to some degree by the quality of concurrent or subsequent close relationships. Freyd has noted, for example, that the isolation and psychological orphaning unwittingly imposed on survivors of betrayal trauma by non-comprehending observers can exacerbate dissociative symptoms (Summit, Miller, & Veltcamp, 1998, as cited in Freyd, 1996). In the other direction, individuals who have a validating and supportive attachment figure outside the abusive relationship have been shown to have better long-term psychosocial outcomes than those whose non-perpetrating caregivers instead deny or discredit the abuse experience (e.g., Fromuth, 1986; Johnson & Kenkel, 1991; Tremblay, Hebert, & Piche, 1999). Though Bowlby did not theorise about this subject directly, he proposed similar dynamics regarding the relational moderators of the long-term psychosocial impact of the death of an attachment figure. More specifically, Bowlby posited that:

not infrequently after a person has been bereaved the situation with which he has to deal is unique, for the death entails the loss of the very person in whom he has been accustomed to confide. Thus, not only is the death itself an appalling blow but the very person towards whom it is natural to turn in calamity is no longer there. For
that reason, if his mourning is to follow a favorable course, it becomes essential that
the bereaved be able to turn for comfort elsewhere. (Bowlby, 1980, pp. 231–232) 10

While Bowlby’s rationale was focused on caregiver loss, not caregiver abuse, is
not difficult to imagine how it could be extended to describe both.

Theoretical differences
Attachment vs. cheater detection
As we have argued above, BTT hinges upon AT’s assumption that humans have
a strong innate drive to create and then preserve needed attachment relationships.
However, BTT highlights the importance of the interplay between two evolved
systems: attachment and cheater-detection (Cosmides & Tooby, 1992). Indeed,
almost as important as its assumptions regarding attachment is BTT’s underlying
proposition that as social animals, humans have evolved the capacity to readily
learn to be superb detectors of betrayal—an ability that protects individuals
against further betrayal, and is, under many circumstances, essential to survival.
However, when a person is being betrayed by someone upon whom they depend
for physical and emotional survival, this normally protective capability becomes
highly problematic, as the withdrawal and confrontation behaviours invoked by
betrayal detection are antithetical to the approach and engagement behaviours
needed to maintain attachment relationships (Freyd & Birrell, 2013).

In cases where the two systems come into conflict, BTT proposes that attach-
ment is so important to our survival that it trumps the cheater-detector mecha-
nism. In other words, when preserving an attachment relationship with a betrayer
is more essential to survival than detecting the attachment figure’s wrongdoings,
betrayal blindness becomes adaptive, as it serves to resolve the conflict. In this
way, within BTT is a central dialectic (i.e., the conflict between needing to know
and needing not to know about betrayal) that is not explicitly articulated in AT.

Divergent motivational foci for betrayal blindness/defensive exclusion
As we have previously outlined, Bowlby (1980) argued that defensive exclusion
of traumatic information occurs when new information is incompatible with
established IWMs. According to Bowlby, the task of dismantling or altering ones’
representational model—indeed, of whether the change is positive or nega-
tive—is by nature arduous and undesirable. Change in a negative direction (as
would be the case when an existing model of a benevolent caregiver is replaced
by a new model of an abusive caregiver) is not only difficult and unwelcome, but
painful and frightening as well. As such, humans are very likely to ignore or
exclude this information from processing. Bowlby was hardly alone in this
position, as many psychologists from within and outside psychodynamic schools

Rosemary E. Bernstein and Jennifer J. Freyd
of thought have postulated that dissociation is a defence against the psychological and emotional pain that results from traumatic experiences (Freud, 1900a; Goleman, 1985; Green, 1992; Greenson, 1967; Janoff-Bulman, 1992).

Freyd has actively distinguished BTT from this pain-aversion-focused explanatory model, for example by drawing contrasts between BTT and Janoff-Bulman's (1992) Shattered assumptions theory, which proposes that “trauma damages people by shattering fundamental assumptions of a benevolent world and meaningful self. Trauma leads to denial and repression . . . in order to protect the self from this overwhelming new view of the world and the self” (Freyd, 1996, p. 23). BTT, in contrast, proposes that betrayal blindness happens not because people want to avoid pain or preserve current representational models, but because there is a basic incompatibility between “knowing about a betrayal in the external world and maintaining a necessary system of belief in order to guide adaptive behavior” (Freyd, 1996, p. 25). Remaining blind to betrayal solves this incompatibility and allows victims to maintain an attachment to a figure vital to their survival and healthy development.

While, as we have seen, Bowlby endorsed the pain-aversion hypothesis at a general level, regarding the more specific motivations that underlie the defensive processing of maltreatment by a caregiver, he offered three non-mutually exclusive possibilities. The first explanation, prominent within traditional psychoanalytic schools of thought but least compelling to Bowlby, is that children are simply unable to accommodate within a single image the parent’s kind and unfavourable treatment of him or her. A second explanation proposed that “the young child, being totally dependent on his parent’s care, is strongly biased to see them in a favorable light and so to exclude contrary information” (Bowlby, 1980, p. 71). As we have explored above, it is this conceptualisation that is most central in BTT. While Bowlby considered this explanation possible, he favoured a third option, which, in his view, had the most empirical support at the time. This third possibility highlights some parents’ explicit insistence that their children regard them in a favourable light, and the pressure children put upon themselves to comply. Regarding this view, Bowlby wrote,

> on threat of not being loved or even of being abandoned a child is led to understand that he is not supposed to notice his parent’s adverse treatment of him or, if he does, that he should regard it as being no more than the justifiable reaction of a wronged parent to his bad behavior. (Bowlby, 1980, p. 71)

In support of this third idea, Bowlby drew from the writings of psychiatrist Emanuel Peterfreund (1971), who proposed that the information most likely to be defensively excluded is of a kind that, if accepted, might result in serious conflict with parents, which, in turn, would bring on acute distress. Indeed, Bowlby speculated that most instances of defensive exclusion could be accounted for by a particular kind of parent–child conflict, wherein “a child is in course of observing
features of a parent’s behavior that the parent wishes strongly he should not know about” (Bowlby, 1980, p. 70). In contrast, Freyd (1996) has emphasised the intrinsic motivations for remaining blind to betrayal. While she hypothesises that explicit threats and demands for silence from the abuser should increase the likelihood of amnesia\textsuperscript{11} (i.e., that it serves as a potential intensifying \textit{moderator} of the influence of betrayal trauma on amnesia), she holds that this is not the \textit{mechanism} itself.

\textbf{Future studies}

While betrayal amnesia is a notoriously difficult process to study, a preliminary study that could act as a first step in clarifying this theoretical discrepancy could more explicitly examine the association between betrayer’s explicit threats and amnesia in adult survivors of documented child abuse. In this design (similar to that used by Williams, 1994), presence of parental threats as documented in original Child Protective Service (CPS) reports may be examined in conjunction with degree of amnesia for the abuse (operationalised as self-reported recollection of abuse during an assessment interview). Instances of amnesia with no documented parental threats may indicate that, as predicted by BTT, explicit threats are not a necessary part of the amnestic process. Of course, this design is not without serious shortcomings, as abuse documentation may not include everything that occurred, and self-reports may be obscured by amnesia. Moreover, the results may not be generalisable, as survivors who have had their abuse documented are likely to differ in meaningful ways from those whose abuse was never reported to the authorities. For example, in the former case, one can imagine that the survivor more likely had a supportive non-abusing adult who identified and took action against the abuse, which is unlike the high level of secrecy within the homes of many incest survivors. Thus, a more complete understanding of the importance of explicit threats on producing abuse amnesia would likely require a series of studies using a number of different populations (i.e., abused children, and adults who report having at one time partially or wholly forgotten childhood abuse but who now remember it) as well as a number of different measures to assess actual abuse experience (i.e., observer/sibling report, perpetrator report, and self report).

\textbf{The role of fear in amnesia for traumatic events}

Earlier, we discussed the way in which attachment theorists account for abused children’s continued engagement in proximity-seeking with their abusing caregiver, including Hesse and Main’s (2000) analogy to Darwin’s (1839) observations of the Galapagos sea lizard. What was not included in the above review was Hesse and Main’s name for this situation: “fright without solution” (Hesse & Main, 2000, p. 1105), and a more elaborated explanation of the way they conceptualise the role of fear in promoting dissociation. In short, Hesse and Main

\textit{ATTACHMENT}
(2000) argue that disorganised/disoriented attachment behaviour should be expected whenever an individual is *markedly frightened* by its primary haven(s) of safety (i.e., the attachment figure). It is here that BTT and AT again diverge, as Freyd (1996) has actively pushed against the field of trauma studies' overwhelming focus on fear as the sole motivating reaction in responses to trauma (i.e., the “fear paradigm”; DePrince & Freyd, 2002; see Janoff-Bulman, 1992; Roth & Newman, 1991 and Shay, 1994 for exceptions). Instead, BTT shifts the focus from the impact of trauma on affective systems to a focus on its impact on cognitive systems, and conceptualises betrayal—not frightfulness—as the feature of traumatic experience most important in predicting post-traumatic dissociation. To illustrate, Freyd (1996) has pointed out that not all forms of betrayal trauma that can result in dissociation are inherently frightening. For example, while sexual abuse certainly has a strong potential to be fear provoking when force is involved, it may well not be when it involves the slow progression from once-appropriately loving gestures (Freyd, 1996). BTT proposes that “the traumas that are most likely to be forgotten are not necessarily the most painful, terrifying, or overwhelming ones (although they may have those qualities), but (rather) the traumas in which betrayal is a fundamental component” (Freyd, 1996, pp. 62–63).

To some extent, Hesse and Main acknowledge the possibility that non-frightening caregiver behaviour can also result in child–caregiver attachment disorganisation, as in their updated (2006) coding system for the “frightened/frightening” (FR) caregiver behaviours associated with both adult unresolved states of mind and infant attachment disorganisation, they identify FR behaviours that are and are not expected to directly evoke infant alarm. Those in the former group include:

1. Direct indices of entrance into a dissociative state (e.g., complete freezing; altered voice).
2. Threatening behaviour inexplicable in origin or anomalous in form (e.g., stalking or growling at the infant in the absence of meta-signals of play).
3. Frightened behaviours inexplicable in origin or anomalous in form (e.g., sudden frightened look without any environmental threat or frightened retreat from/apprehensive approach toward the infant).

Those in the latter, non-frightening group include:

1. Timid/deferential/role-inverting behaviour (e.g., postural submission to infant aggression or turning to the child as a haven of safety when alarmed).
2. Sexualised behaviour toward the child (e.g. deep kissing or sexualised caressing).
3. Disorganised/disoriented behaviours (e.g., mistimed movements, anomalous postures, or any observable collapse in caregiving strategy) (Hesse & Main, 2006, p. 320).
While this inclusion of non-frightening caregiver behaviours (especially non-frightening sexual advances) may suggest movement toward theoretical convergence between BTT and AT, Hesse and Main (2006) continue to emphasise the affective role of fear, and diminish the stand-alone/direct effects of non-frightening FR behaviours by asserting that:

The final three (“secondary”) subcategories within the FR coding system most likely would not in themselves lead directly to an approach–flight paradox for the infant. Nonetheless, they each imply alterations in normal consciousness. This should, in principle, increase the likelihood that the anomalous behaviors capable of directly producing disorganization will appear at other times. (p. 324)

Future studies
Preliminary research that may help clarify this theoretical difference includes that which more explicitly compares the dissociative/disorganised symptoms of individuals who have experienced frightening vs. non-frightening forms of betrayal trauma (e.g., violent vs. non-violent forms of sexual abuse by a caregiver). For example, a study might recruit a sample of adults with documented child abuse histories to assess the level of association between betrayal amnesia and the degree to which the abuse was frightening (as reported in the CPS abuse report). For those with documented abuse histories, failure to report the abuse, partial reporting of the abuse, and reporting of abuse experiences without labelling them as abusive could be considered varying degrees of betrayal amnesia. The degree with which betrayal amnesia relates to coder-rated frightfulness of the abuse as reported in the CPS case file may provide support for one theory over the other (i.e., a high association may provide support for the predictions made by AT, while a weak association may provide support for those made by BTT).

In addition, researchers might examine the dissociative behaviours and autonomic responsivity (i.e., fight or flight system activation) of children in response to frightening vs. non-frightening forms of caregiver conflicting signals (which have been shown to experimentally induce child disorganised behaviours; Heinicke & Westheimer, 1966; Hesse, 1999; Solomon & George, 1999). If non-frightening and frightening caregiver behaviours both resulted in strong sympathetic nervous system (SNS) activation (suggesting higher levels of alarm or fright), or if strong SNS activation was associated with more severe forms of child dissociation, then we might conclude that (as predicted by AT) fear is an essential component of dissociation. If not, we might conclude that for dissociation to occur, it is not necessary that individuals experience fear (as predicted by BTT).
Summary and conclusions

In this paper, we have discussed three points of convergence and three points of
differentiation between AT’s and BTT’s conceptualisations of the human response
to traumatisation by an attachment figure. Areas of theoretical agreement
included:

1. The shared assumption that individuals are highly and innately motivated to
   maintain affectional bonds with depended others.
2. Common predictions about what cognitive and behavioural strategies are
   adaptive in response betrayal, including the defensive and selective processing
   of information that may threaten established attachment representations.
3. The similarities in the developmental predictions each theory makes about the
   ways in which betrayal influences subsequent relational and emotional health.

Differences between the theories included:

1. BTT’s focus on not only the attachment system but the cheater-detection
   system.
2. The slightly divergent emphases each theory places on the role of explicit vs.
   implicit caregiver threats of abandonment in defensive processing/betrayal
   blindness.
3. The different weight ascribed to the role of child fear in promoting dissocia-
   tion in the context of betrayal.

Regarding the second and third areas of divergence, specific study sketches were
provided that may help identify the theory which provides the most helpful
explanatory fit.

In exploring these points of convergence and divergence, it became clear that
BTT is highly compatible with what AT proposes about adaptive and expectable
interpersonal phenomenon following abuse by a depended upon other. Even in
some areas where the theories diverge, the theories remain essentially com-
patible. For example, while it is true that Bowlby expressed some degree of
preference towards an explanation of defensive processing which focused on
extrinsic (i.e., a parent’s explicit threats of withdrawal or punishment highlighted
in the second view) over intrinsic ones (i.e., the internal drive for the child to
maintain the attachment relationship emphasised in the third proposal), it
remains the case that BTT’s conceptualisation of betrayal blindness and the inter-
nal motivations underlying it are consistent with AT and with Bowlby’s work.

In conclusion, our aim in this paper was to describe BTT and the way it pro-
vides an important extension of AT, in particular a deepening understanding of
D attachment. It is also very useful in providing creative and effective additional
resources for clients/survivors and therapists to empower recovery and healing.
Notes

1. In this case, betrayal “amnesia” most likely occurs via repeated processing through various feedback loops. Freyd (1996) has proposed that betrayal amnesia might be most adaptive in cases where abuse becomes identifiable as a betrayal only sometime after it occurs.

2. As an example of this type of mild amnesia, Freyd (1996) provides the case of childhood incest survivor Lee Davidson, who had always remembered the events of her abuse, but had not interpreted them as abusive until she was in therapy years later. Instead, she grew up thinking that these experiences happened to all children, and then later assumed they happened because she was bad.

3. According to BTT (Freyd, 1996, p. 140), there are seven factors that should positively predict the likelihood of amnesia for abuse. In addition to the central factors of:
   1. Perpetration by a caregiver.
   2. Young age at time of abuse.

   These also include:
   3. Explicit threats demanding silence (see section “Divergent motivational foci for betrayal blindness/defensive exclusion”).
   4. Separation between abuse and non-abuse contexts (e.g., night and day).
   5. Isolation during abuse.
   6. The perpetrating caregiver’s declarations of alternative reality-defining statements.
   7. Lack of discussion of the abuse.

4. For example, neuroimaging research has demonstrated activation in the reward and approach regions of the maternal brain (at least those with a secure attachment organisation) in response to infant smiles and cries; for example, Strathearn, Fonagy, Amico, and Montague, 2009.

5. Studies have demonstrated that between 50% (Thompson, 2008) and 80% (Carlson, Cicchetti, D., Barnett, D., & Braunwald, 1989; Lyons-Ruth & Block, 1996) of maltreated infants in have disorganised attachment relationships with their caregivers, as compared to about 15% of infants in middle class community samples (see Thompson, 2008 for review).

6. Notably, these same behaviours have been observed in non-disorganised infant–caregiver’s dyads in response to major separations, experimentally induced situations of inescapable shame, and experimentally induced conflicting signals (see, for example, Heinicke & Westheimer, 1966; Hesse, 1999; Solomon & George, 1999).

7. Adult unresolved state of mind with regard to trauma or loss is analogously characterised by atypical lapses in monitoring or reasoning during the adult attachment interview (AAI; George, Kaplan, & Main, 1984; Main & Goldwyn, 1998; Main, Goldwyn, & Hesse, 2002). Hesse and Main (2006) have conceptualised these lapses as fitting well into a dissociative model.

8. Regarding this point, Freyd references a sizeable body of evidence that conscious recall (i.e., explicit memory) and unconscious influences (i.e., implicit memory) are...
Indeed dissociated at the cognitive level (see, e.g., Jacoby & Kelly, 1992). Similarly, Bowlby (1980), cited Tulving's (1972) claim that storing information according to personal experiences (i.e., autobiographically) is distinct and separate from storing it according to its meaning (semantically).

9. According to BTT, symptoms of traumatic response should depend on two independent features of the trauma: (1) the degree to which it induced terror or fear, and (2) the degree to which it induced betrayal blindness (which, as we have discussed, occurs when the trauma represents a social betrayal). More specifically, BTT proposes that intensely fear-inducing traumas should lead to hyperarousal and anxiety; those causing betrayal blindness should lead to numbing, dissociation, amnesia, and shame; and those that induce both (which many traumas do) should lead to both kinds of symptoms. While research has indeed linked betrayal (and not fear) with dissociation and PTSD withdrawal symptoms (e.g., DePrince, 2001), interestingly, other research has shown that betrayal also predicts outcomes previously associated with the fear dimension, such as anxiety, anxious romantic attachment, and hypervigilance (e.g., Bernstein, Hailey, Knight, & Freyd, in preparation; Klest & Freyd, 2007; Owen, Quirk, & Manthos, 2011). At present, it is unclear whether these outcomes are a function of the betrayal itself, of the blindness reaction, or perhaps an interaction between betrayal blindness and peritraumatic fear. Thus, future research will be necessary to clarify the predictive roles of each.

10. Indeed, quality of care following parental loss has been found to be more predictive of later adaptation than the occurrence of loss per se (Harris, Brown, & Bifulco, 1986; as referenced by Lyons-Ruth, Yellin, Melnick, & Atwood, 2003).

11. Drawing from the work of Lister (1982) on forced silence, Freyd has acknowledged that

   explicit threats and demands for silence from the abuser (statements such as “if you tell I’ll kill you” or “I’ll kill your mother”) would hypothetically increase the advantages for the abused child in forgetting the betrayal in order to maintain critical attachment bonds and would thus increase the probability of amnesia.
   (Freyd, 1996, p. 137)

12. It could be argued that this divergence represents one within AT rather than one between theories, as Bowlby’s original writings, like Freyd’s, emphasise cognitive systems as the primary mechanisms underlying the human response to negative caregiving experiences. This is evidenced in his lengthy writings on defensive exclusion and internal working models of self and of other. Thus, while BTT’s de-emphasis on the causal role of fear for post-traumatic dissociation/disorganisation is not consistent with AT as furthered by Hesse and Main (2000), it does fall in line with AT as it was originally presented by Bowlby.

13. See Zurbriggen and Freyd (2004) for further discussion of several cognitive mechanisms proposed to underlie the associations between childhood sexual abuse and subsequent risky behaviour, including low self-esteem (akin to deflated internal working model of self), compromised reality-detection mechanism (i.e., learned distrust of one’s own perceptions of reality stemming from chronic exposure to...
authority figures’ denial of the betrayal), dissociation and divided attention, and non-functioning or damaged consensual sex decision mechanisms (CSDMs), which may include inaccurate beliefs, unhelpful cognitions about the self (e.g., an internalised belief that one is not deserving of respect), alexithymia (i.e., a lack of access to one’s internal affective state), and risk-seeking sexual decision rules.

References


Bernstein, R. E., Hailey, B. C., Knight, J., & Freyd, J. J. (in preparation). Posttraumatic hypervigilance and dissociation: comorbid or contradictory?


Main, M., Goldwyn, R., & Hesse, E. (2002). Adult Attachment Scoring and Classification System. Unpublished manuscript, University of California at Berkeley, Department of Psychology.


Rosemary E. Bernstein and Jennifer J. Freyd