Concepts of Causality in Psychopathology: Applications in Clinical Assessment, Clinical Case Formulation and Functional Analysis

Stephen N. Haynes, Ph.D.
University of Hawaii at Manoa

Joseph Keawe'aimoku
Kaholokula, Ph.D.
University of Hawaii at Manoa

William H. O'Brien, Ph.D.
Bowling Green State University

Cilia Witteman, Ph.D.
*Radboud University
Nijmegen, Netherlands*

ABSTRACT

This paper discusses and integrates concepts of causality in psychopathology, clinical assessment, clinical case formulation and the functional analysis. We propose that identifying causal variables, relations and mechanisms in psychopathology and clinical assessment can lead to more powerful and efficient interventions. Four criteria must be met in order to infer a causal relation between two variables: (a) covariation between the variables, (b) temporal precedence of the causal variable over the effect, (c) the exclusion of plausible alternative explanations, and (d) a logical basis for inferring a causal relation. These are challenging criteria to meet in clinical assessment. They require multivariate time-series assessment or manipulation strategies with measures that are sensitive-to-change. The bidirectional and dynamic nature of causal relations and the role of the clinician in deriving causal judgments present further challenges to causal inference in clinical assessment. We emphasize the functional analysis as a model of clinical case formulation. The functional analysis emphasizes the specificity of constructs and requires that we identify the mechanisms that explain the causal relations—the means through which a causal effect operates. In sum, concepts of causality provide a scientific foundation and methodological guidance for clinical assessment, clinical case formulation, and the functional analysis.

[1]Correspondence concerning this article should be addressed to: Stephen N. Haynes, Ph.D., Department of Psychology, University of Hawaii, 2530 Dole St., Sakamaki, D-400, Honolulu, HI 96822. E-mail: sneil@hawaii.edu

Concepts of causality in psychopathology are important elements in psychological assessment, clinical case formulation, and the functional analysis. They help the clinician estimate what variables might be influencing a person's behavior problems, the mechanisms through which a causal variable exerts its effects on a behavior problem, and how other variables might be influencing the relations between the causal variable and behavior problem. Concepts of causality in psychopathology also inform the clinician about the best assessment strategies, methods, instruments, and measures to use for a particular client, and are important in all psychotherapy paradigms (cf. Allen, 1993; Magnavita, 2006). Finally, concepts of causality are central to clinical judgments because many interventions attempt to modify causal variables that are hypothesized to affect a client's behavior problem or positive treatment goals. Many cognitive-focused therapies, for example, attempt to modify a person's dysfunctional thoughts that affect the duration and intensity of depression symptoms because these thoughts are presumed to be one of the causal mechanism that explain the relation between life stressors and depressive symptoms (Beck & Brad, 2009).

In this paper we discuss concepts of causality, review several attributes of causal variables and causal relations in psychopathology, and discuss how concepts of causality inform clinical assessment strategies and clinical judgments. More broadly focused discussions of causality and methods of causal analysis can be found in Bunge (2009), James, Mulaik and Brett (1982), Pearl (2000), and Shadish, Cook, and Campbell (2001).

Diverse Concepts of Causality

Causality has been discussed for centuries among scholars and scientists. These discussions have primarily focused on the relevance of causal inference, the types of causal relations, and the necessary conditions for inferring a causal relation. In clinical assessment, the discussions have focused on assessment strategies for identifying and estimating causal relations related to a client's behavior problems and the role of causal inferences in clinical case formulation.

Most behavioral scientists assume that causal relations in psychopathology exist and are important foci in clinical assessment and intervention. Among philosophers of science, the existence of causality apart from our perceptions of it is still debated (see Danks, 2005 and Newsome, 2003). Regardless of the philosophical complexities associated with the concept of causality, thousands of studies and clinical interventions are based on the implicit or explicit assumption that changes in certain variables (hypothesized causal variables) often lead to important changes in other variables (client behavior problems).

Concepts of causality differ across disciplines. Some types and conditions of causality that have been proposed are more relevant than others to clinical assessment and judgment. For example, assume that "Y" is a dimension of a behavior problem (e.g. *frequency* of binge eating or *intensity* of social anxiety) and that "X" is a potential causal variable (e.g., a life stressor, discriminative stimulus, or social reinforcement). The causal relation between these two variables can have a number of forms. If Y *always* occurs after X occurs it can be labeled as a *sufficient cause*. It is important to note however that when there is sufficient cause, Y can still occur without the prior occurrence of X. A practical example of this type of cause is observed in the relation between blood glucose levels and food consumption among persons with diabetes. Consumption of carbohydrates (X) can be sufficient to cause

a subsequent rise of blood glucose levels (Y). However, activities other than food consumption can also cause blood glucose levels to rise (e.g., failure to take an insulin injection, chronic daily stress).

A *necessary cause* is identified when Y *never* occurs without the prior occurrence of X. Note that under this condition, X can still occur without Y occurring. A practical example of this type of relationship is sexual contact and a sexually transmitted disease. Specifically, a sexually transmitted disease (Y) never occurs without the prior occurrence of sexual contact (X). However, sexual contact can occur without the subsequent occurrence of a sexually transmitted disease.

An *insufficient cause* is observed when Y occurs only after X occurs in combination with another variable, Z. However, Y does not occur when X occurs alone. An example of this type of relation is observed when a person may develop a particular condition (Y, e.g., schizophrenia) only when they carry a genetic susceptibility (Z) and are concurrently exposed to a specific environmental event (X, e.g., life stressors).

An *immediate (proximal) cause* is one wherein the causal variable exerts its influence without any intervening events (i.e., there is a *temporal contiguity* between Y and X). In the behavioral sciences, this type of cause-effect relation is very difficult to establish because one can frequently identify myriad intervening events occurring between a particular causal variable and behavior problem. Other types of causal relations are outlined in Haynes et al., (2011a).

Many of the causal variables encountered in psychopathology and clinical assessment are *insufficient*, in that they affect a person's behavior problems only in combination with other causal variables. As noted in the earlier example, a life stressor (e.g., increased conflict within the family) might trigger schizophrenic behaviors for a patient only when it occurs in combination with a genetic susceptibility to schizophrenia or in combination with other major stressors in his or her life. Additionally, clinical case formulations, such as the *functional analysis[1]* (Haynes et al., 2011a) often include *distal* as well as *proximal* causes. For example, an immediate cause of a psychiatric patient's relapse while at home might be a series of critical comments by family members. However, this proximal cause is more likely to adversely influence the patient's behavior when more distal causes such as medication noncompliance, fatigue from sleeplessness or recently experienced social stressors are also co-occurring.

What Evidence is Necessary to Infer a Causal Relation In Clinical Assessment?

In this section we briefly review several concepts of causality, methods of estimating causal relations, and limitations of causal inferences that are relevant to clinical assessment, clinical case formulation and functional analysis. First, we present four necessary conditions for inferring a causal relation: (a) covariation, (b temporal precedence, (c) negation of an alternative explanation for an inferred causal relation, and (d) a logical connection.

Covariation

Covariation is the most widely accepted, but deceptively simple, condition for inferring a causal relation between two variables. Evidence of covariation can include meaningful correlation or elevated conditional probability of two variables. Absence of a functional relation, when controlling for confounding effects (see discussion of exceptions below), means that two variables cannot have a causal relation. For example, in order for "escape from an aversive emotional state, i.e., experiential avoidance (Giorgio et al., 2010)" to function as a causal variable for a client's alcohol use, alcohol use must be more likely to

occur when the client is experiencing an aversive emotional state relative to a non-aversive emotional state.

There are several challenges to the detection of covariation between variables in clinical assessment.

1. Due to measurement error (e.g., using an invalid measure or poorly timed measurements), a true functional relation may not be identified although one exists.

2. Two measures of presumably different constructs can have overlapping elements, which gives the false impression that the constructs covary. An example of this is the high correlation often observed between self-report measures of depression and of anxiety—two psychological constructs presumed to be qualitatively different but which share some similar elements (Stulz & Crits-Christoph, 2010)[2].

3. Two variables can be causally related but only within particular domains or contexts. For example, response contingencies may be an important causal variable for a child's oppositional behavior with one but not the other parent. For another example, conflict with other patients on a psychiatric unit could trigger a patient's aggressive behavior but only when he has recently refused medication or recently returned to the unit following a stressful home visit. If the clinician measured the covariation between conflict and aggression in other contexts, significant covariation might not be detected.

4. Some causal relations may operate within some but not other *values of variables*. For example, there can be a significant causal relation between the severity of a client's nighttime worry about her life stressors and her delayed sleep onset, but only when those life stressors exceed a certain level. Measurement of the client's nighttime worry and sleeping problems when her life stressors vary within mild-to-moderate ranges could suggest that sleep is unaffected by nighttime worry.

5. Two variables can appear to covary because both are concurrently affected by the same causal variable, which we illustrate in Figure 1 with examples A and B. For example, a person's binge drinking and tension headaches could covary due to the effects of a recent job loss.

Temporal Precedence
Another condition for inferring a causal relation in clinical assessment is *temporal precedence*. Temporal precedence specifies that the hypothesized causal variable must precede the behavior problem in time. Without establishing temporal precedence between X (an hypothesized causal variable) and Y (a behavior problem), it is difficult to rule out alternative functional relations, for example, that X is a result, rather than a cause, of Y, or that a third variable effects the apparent covariation between X and Y, as noted in "5" above.

The temporal relation between variables can be difficult to detect in clinical assessment. First, precedence is a necessary but insufficient condition for excluding the possibility of a "third variable effect." As illustrated in diagram A and B in Figure 1, a causal variable could affect two behavior problems, but with different causal latencies. Consequently, in this case the two behavior problems would demonstrate covariation and one would reliably precede the other, even though they were not causally related.
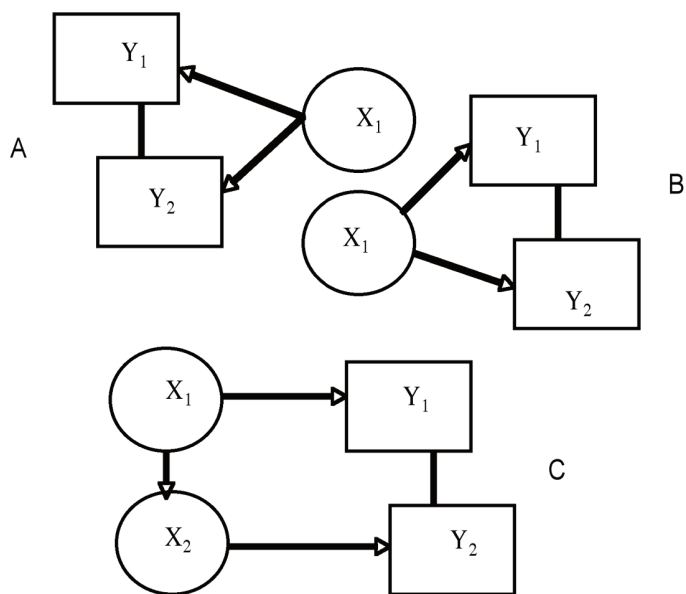
Figure 1. Three ways in which two behavior problems (Y$_1$ and Y$_2$) can demonstrate reliable covariation and temporal precedence in the absence of a causal relation. In all cases, the latency of causal effect is different for the two behavior problems, which results in temporal precedence of one over the other (O = causal variable, □ = behavior problem, —- = correlated noncausal functional relation, -→ causal relation (from Haynes et al., 2011a). Note that in all cases, it would still be possible for Y$_1$ and Y$_2$ to have a causal relation.

The requirement of temporal precedence in causal judgments makes it difficult to interpret many published studies in psychopathology. For example, causal relations sometimes have *restricted temporal boundaries*. That is, the *latency of a causal effect* and *the duration of a causal effect* can differ across types of causes, for different behavior problems, for different dimensions (e.g., duration vs. severity) of a behavior problem, and for different persons. Therefore, measurement of a hypothesized causal relation between two variables outside of the temporal domain of causal effects (for example, before or after the causal effects) would suggest no covariation between two events when there is in fact a causal relation

between them. Recall that we are only *inferring* causality, what we measure is only covariation.

Longitudinal research often includes assessment strategies that involve measurement outside the likely temporal domain of a causal relation. For example, many longitudinal studies of the functional relations between depression and marital distress (see discussions in Beach, 1990 and Rehman et al., 2008) measure these two constructs two or three times, perhaps a month, 3 months, or a year apart. The goal has been to examine the degree to which depression (or marital distress) at time 1 predicts marital distress (or depression) at time 2. This is a *convenience time-sampling strategy* (Minke & Haynes, 2011) but it is likely that the temporal parameters of these time-sampling strategies are not congruent with the latency and duration of causal effects. For example, given the normal variability in depressed mood and its effects on interpersonal behaviors and cognitive processes, its strongest effects on marital distress might occur within hours, days, or weeks, rather than after a month or a year[3]. Further complicating the task of causal inference, the latency of causal effects might differ (i.e., might be *conditional*), depending on the level of relationship satisfaction and the valence of recent dyadic interactions. Consider that if there are multiple causal variables that affect both depressed mood and marital satisfaction, influential events are more likely to occur as the time between measurement increases. In sum, longitudinal measurement strategies outside of the temporal domain of causal relations could lead to a misestimate of the functional relations between the variables and a misjudgment about their causal relation.

It is particularly difficult to draw inferences about causal relations in clinical assessment by simply administering self-report question-

naires at a single point in time (e.g., at an intake assessment session), unless there are temporal and conditional elements contained in the questionnaires (e.g., "I often feel anxious" (an unconditional query) vs. "I often feel anxious before meeting people for the first time" (a conditional query)). In many clinical assessment contexts, the clinician is required to estimate causal relations based only on apparent covariation between measures of variables that are hypothesized to be causally related. As we illustrated in the previous example, a clinician might erroneously infer a causal relation between a client's life stress and alcohol use when measures of these two constructs are both elevated and prior research has suggested the likelihood of a causal relation between them.

Further complicating the identification of causal relations in clinical assessment is that temporal precedence of a causal variable can hold for one dimension of a behavior problem and not another. For example, a client's inability to resolve family conflicts could precede and affect the severity or duration of his depressive episodes but not their onset (see discussion in Nezu et al., 2004). Similarly, a client's heightened sensitivity to physical sensations and likely misattributions about them, such as believing that a rapid heart beat indicates the start of a heart attack, may affect the duration or severity of the client's panic episode but not their onset (see review of panic disorder in Smits et al., 2006).

Because causal relations can differ across dimensions of a behavior problem, it is important for the clinician to identify the most important dimension (e.g., latency to onset, rate, severity, or duration) of a behavior problem during clinical assessment, to emphasize that dimension in the functional analyses, and to target that dimension during follow-up assessments. For example, failure to specify whether the functional analysis explains the onset, rate, likelihood, magnitude, or duration of self-injury, manic episodes, or tantrums could lead to erroneous inferences about the role of particular causal variables, less-than-optimal treatment foci, and the use of measures that are insensitive to the most important changes in the targeted problems.

The Exclusion of Alternative Explanations for the Functional Relation

The exclusion of alternative explanations for the functional relation is the third requirement for causal inference. We illustrated in prior examples, and in Figure 1, that two variables can have no causal relation even though they demonstrate covariation and one variable reliably precedes the other. This state of affairs can occur for several reasons but one of the most important and challenging in clinical assessment is that both variables can be influenced by an unmeasured third variable. In each causal diagram in Figure 1, Y_1 covaries with and precedes Y_2 but note the different causal latencies between X and Y_1 and Y_2. The clinician could erroneously infer that Y_1 influenced Y_2 if X was not also measured.[4]

An erroneous inference that there is a causal relation between two variables is especially likely under three conditions: (a) when the latency of effects of X are shorter for Y_1 than they are for Y_2 (which results in a reliable temporal precedence of the first variable over the second), (b) when the researcher or clinician fails to measure X (an issue of the "*content validity*" of the assessment strategy and functional analysis; Haynes, Smith, and Hunsley, 2011b), and (c) when the functional relation observed between Y_1 and Y_2 is consistent with the investigator's causal beliefs. Because there are endless possibilities, it is difficult within the clinical assessment context to exclude alterna-

tive explanations for an apparent causal relation. Consequently, whenever possible, the clinician should measure multiple variables in a carefully defined temporal sequence.

The strongest, but not infallible, test of a causal relation relies on observing the effects of *systematic manipulation* of the hypothesized causal variable (see Kazdin, 2003, for a discussion of within-subject designs useful in clinical assessment) that results from either pretreatment assessment (e.g., use of analog behavioral observation) or from the effects of intervention. If X_1 is a causal variable for Y_1, systematic manipulation of a dimension of X_1 (systematically changing its occurrence, magnitude, duration) must be followed by changes in Y_1 in a manner congruent with the hypothesized causal relation.

Even with demonstrated effects from the manipulation, the clinician must be cautious about inferring a causal relation. Systematic effects on Y_1 associated with the manipulation of X_1 are indicative of a causal relation only to the degree that alternative explanations for the apparent effects of the manipulation can be ruled out. Parent-child conflicts, for example, have often been presumed to result from how parents manage their child's positive and negative behaviors (e.g., inconsistent responses to challenging and positive child behaviors). However, parents' responses to their child's behavior can also be affected by the behavior problems of another family member or from the parents' stressful experiences outside of the home (see discussions in Doherty et al., 1998).

Logical Connections Among Variables

Causal inference in clinical assessment also requires that there be *a logical connection among variables*: The causal variable must have a logical causal connection with its effect. The clinician must ask, "Is there a logical connection between these variables?" or "In what way would X influence Y?" (see discussion of *causal mechanisms* below and in Haynes et al., 2011a). Logical connections are especially important when higher-order, more abstract, and less specific causal variables are used in a clinical case formulation. To illustrate the importance and clinical utility of considering logical connections stemming from some commonly proposed causal relations, how would "frustration" cause aggression, tantrums, or self-injury? How would "low self-esteem" lead to social avoidance? Specifying the causal pathways from these heterogeneous and molar variables could help the clinician develop a more clinically useful assessment strategy, functional analysis, and treatment plan. The treatment foci become more apparent when "frustration" is broken down into more specific and measurable components such as "difficulty expressing feelings and goals to a partner" and when "low-self esteem" is broken down into components such as "expectation of negative evaluation by others."

Identifying Causal Mechanisms in Clinical Assessment

In drawing causal inferences in clinical assessment, the exclusion of an alternative explanation for the apparent causal relation and the mandate for a logical connection between a cause and its effect, require that we identify the *mechanism* of the causal relation—the means through which a causal effect operates. When we consider the mechanism through which one variable affects another, we are asking the question, "*How* does X influence Y?"

Sometimes a causal mechanism between variables is unknown but it is assumed that one will eventually be identified. For example, some psychotropic medications have been shown to benefit some psychiatric patients

but the mechanisms of their actions have yet to be identified. There are obvious treatment benefits if the causal mechanisms underlying an apparent causal relation can be identified because new interventions could be developed that more powerfully or more efficiently produce the same benefit.

Baker et al. (2004) provided an example of the clinical utility of identifying causal mechanisms. The authors noted that depressive states and "craving" often preceded a relapse after substance abuse treatment. The authors proposed that "negative reinforcement" was the causal mechanism common to both. That is, the relation between depressive symptoms, craving, and relapse could be due to the fact that the use of alcohol or drugs is an effective way for some persons to reduce negative affective states. The treatment implications are clear: The clinician must help the client acquire a less harmful method for reducing negative emotional or physiological states (i.e., help the client to establish a behavior that is part of the same *functional response* class as substance use).

The clinical utility of identifying causal mechanisms applies equally to the well-documented finding that there can be *multivariate causal relations* for behavior problems (see Haynes, 1992, for an overview). That is, behavior problems can be affected by apparently dissimilar causal variables. However, the *mechanisms* underlying the common effects of these disparate variables have less frequently been identified. How can both caffeine ingestion and the cessation of a life stressor trigger a migraine headache? How can anxiety sensitivity and fear of negative evaluation both increase the chance of panic attacks in some settings? The answer again resides with the mechanisms that underlie the causal relations. Although a behavior problem can be a function of multiple causal mechanisms, and

a causal variable can affect a behavior problem through multiple causal mechanisms, we will consider now the manner in which different causal variables can operate through a common causal mechanism.

To illustrate the idea of a *shared causal mechanism* among multiple causal variables, we borrow an example from Haynes et al. (2011a) and consider the problem of child abuse. Most comprehensive causal models of child abuse are multivariate. They include deficient parenting skills, low education level of parents, poor parent anger management skills, parent alcohol and drug use, low levels of social support for the parents, insufficiently developed sense of empathy, challenging and aversive child behavior/misbehavior, a high frequency and magnitude of daily social/environmental stressors in the parents' lives, stressful family financial condition, marital distress and conflict, the learning experiences of the parents with their parents, deficient communications skills between parents, and overuse of drugs and alcohol (see discussion of child abuse in Wise, 2006).

Let's consider one causal variable for child abuse, "a high level of social/environmental stressors acting on the parent," to illustrate how multiple causal variables can operate through a common mechanism. Because a stressful environment can increase the chance of child abuse for some parents, any stressful event for the parent can function as a causal variable for child abuse. Stressful events could include a tenaciously noncompliant child, marital conflict, or withdrawal from chronic use of a psychoactive substance. These different causal variables can have similar effects on a parent because they operate through a common causal mechanism—increased social-environmental stressors. The case formulation and treatment implications are again evident: Interventions

that target the parent's strategies for mitigating life stressors, or assist the parent in developing alternative ways of managing his or her reaction to those stressors, would reduce the chance of abusive behavior toward a child.

Note that we could focus the discussion at a more specific level of causal analysis. For example, we could consider social stress in terms of causal mechanisms such as biological variables involving adrenocortical and hypothalamic activation, cognitive variables involving attention disruption and excessive worry, and learning variables in terms of the parent's history of reinforcement for using aggression to affect the behavior of others in their environment.

In the preceding example, we considered how different causal variables could affect a behavior problem through a common causal mechanism. Additionally, a causal variable could exert its influences through *multiple causal paths*. Consider the multiple ways in which excessive alcohol use by a parent could increase the risk of child abuse, as illustrated in Figure 2.
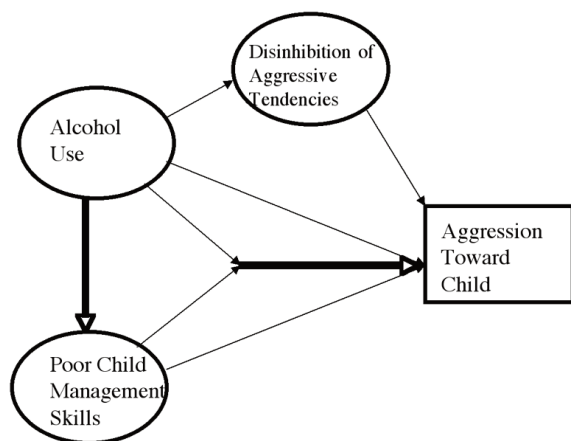
Figure 2: A causal diagram illustrating multiple paths and mechanisms through which alcohol use can increase the chance of aggression toward a child. Note interaction effects (with Poor Child Management Skills), direct effects, and mediated effects of alcohol use (e.g., through Disinhibition of Aggressive Tendencies; from Haynes et al., 2011a).

Multiple causal paths could also be involved in the adverse effects of child abuse on a person's interpersonal functioning as an adult. Sexually abusive experiences as a child could have long-term effects on a client's expectancies regarding the consequences of attentive behaviors from others and beliefs about the chance of being harmed in intimate interpersonal situations. These early experiences could also lead to conditioned fear responses to physically intimate situations and a tendency to avoid some social situations that resemble the abusive situation. A child could also develop negative self-labels (e.g., "I'm bad") and become nonresponsive to the positive approaches of others. These are all *causal mechanisms* (i.e., *mediating variables*) in that they explain *how* a history of sexual abuse (a *distal causal variable*) can lead to interpersonal difficulties as an adult through more *proximal causal variables*.

There are additional assessment and treatment implications of these complex aspects of causal relations in psychopathology. Perhaps foremost, because of *individual differences in causal relations* for a behavior problem, standardized empirically supported intervention programs are likely to be effective but unlikely to produce the maximum benefit for each client. Although any intervention program that effectively intervenes with any of the multiple causal variables relevant to a client's behavior problem is likely to be beneficial, benefits for a client would be enhanced to the degree that the intervention targeted the causal variables and mechanisms in proportion to their importance for a client.

Dynamic Aspects of Causal Relations

Hundreds of published studies have documented that causal variables, causal relations, and behavior problems are *dynamic*—they can change across time *(*see discussion in Haynes,

Blaine, & Meyer, l995). One consequence of the dynamic attributes of causal relations is that the validity of the functional analysis for a client's behavior problems is time-limited. To illustrate, Gerald Patterson and colleagues (2002) noted that in younger children aggression is often maintained by tangible reinforcement (e.g., acquiring objects). As a child ages, aggression is more often maintained by social reinforcement (e.g., peer approval, social dominance). Also, the causal variables for a person's alcohol use can change over time. With continued alcohol use from adolescence to adulthood, causal variables can change from social approval and peer bonding, to the anticipation of the pleasurable effects of alcohol, and eventually, to the avoidance of aversive biological states associated with not drinking (see overview of alcohol-related disorders in Wagner, Sobell, & Sobell, 2007).

The magnitude and direction of a causal variable's effects can also depend on its duration of action. A physical, environmental, or psychosocial stressor can have beneficial or detrimental effects on a client's physical and psychological health, depending partly on its duration of action (see reviews of acute and chronic stressors by Asterita, 1985, and a review of PTSD in Bryant, 2006). Acknowledging the importance of other attributes of a causal variable (e.g., intensity, controllability, type, time-clustering) and important individual differences in responses, brief stressors are more likely to result in increased resilience to future stressors while chronic stressors are more likely to decrease resilience to future stressors.

In addition to chronicity, the *time-cluster* of a causal variable—the pattern of occurrence of a causal variable across time (e.g., its rate or cyclicity)—can also influence its effects. For example, Hyland (1987) proposed that mismatches between a client's goals and actual achievement can lead to a depressed state, but most likely when mismatches cluster together (e.g., during a mismatch "burst"). He suggested that unclustered or infrequent mismatches have little impact on most persons' moods. It is possible that a burst of transient stressors could have different effects on a client as a function of the duration of the burst and that the effects of a burst could differ from that of a single chronic stressor (e.g., physical disability, health problems, death of a spouse) of the same duration and intensity. Some tightly clustered causal variables also might be associated with enhanced effects on the client because the effects are more likely to be additive. Unclustered causal variables may have fewer or weaker effects because there is a greater opportunity for the operation of moderator variables (e.g., positive social support).

The dynamic characteristics of causal relations strengthen the utility of time-series and time-focused clinical assessment strategies and the use of measures that are sensitive to change. Haynes et al. (2011b) reviewed several aspects of assessment instruments and measures that affect their sensitivity to change. Several of the most important are (a) the time frame included in items in self-report methods (e.g., queries requiring recall of events the previous day or week vs. the previous month), (b) the level of specificity of the variable being measured (e.g., a measure of a heterogeneous construct such as "depression" vs. a measure of more specific elements of the depression construct such as "fatigue" or "sleep disturbance"), (c) the response format (e.g., yes-no dichotomous response format versus a 5-pt continuous scale), (d) the content validity of the assessment instrument for a particular client (e.g., the degree to which a measure of "social anxiety" for a client captures the unique aspects and omits irrelevant aspects of that phenomenon for the client; see discussion of idiographic assessment in Haynes, et al.,

2009), and (e) floor and ceiling effects of a measure (e.g., the degree to which a measure of obsessive thinking or life stressors can capture changes at very high or very low levels of those phenomena).

The dynamic attributes of causal relations also draw attention to the importance of temporal factors in clinical case formulations. Attention to only the *state* of a variable when measured, without attending to its *dynamic phase* (e.g., increasing or decreasing slope across time, patterns of occurrence across time, and latency of causal effects) at the time of measurement, can lead to a misidentification of causal variables or the misestimation of the strength of functional relations relevant to a client's behavior problems or positive treatment goals. It can be useful and time-efficient for a clinician during clinical assessment to collect data on the momentary state of a variable, such as a measure of the momentary state of pain, social support, or life stressors. However, it may be helpful in developing a functional analysis for a client to measure the same variable at multiple points in time to capture not only the *state* (at point of measurement) but also the *phase* (changes in state across measurement points over time relative to the point of measurement). Data on a variable at one measurement point in the context of its time-course is an example of a *state-phase function* (for a more detailed discussion of *state-phase function* see Haynes et al., l995).

Data on state-phase functions of a client's behavior problems and causal variables are difficult to acquire in clinical assessment. Estimates of the state-phase function of a variable, for example acquiring estimates of changes in the rate of a client's compulsive behaviors, level of conflict with his or her partner, or frequency of daily life stressors, preceding and following a particular clinical assessment occasion, can

be obtained in only two ways: (a) through the client's or another's retrospective reports of the time course of behavior disorders and causal variables (perhaps using time-line follow-back methods), or (b) by a time-series measurement strategy involving frequent measurements of targeted variables, perhaps using hand-held computers or ambulatory biosensors (Haynes & Yoshioka, 2007). Both assessment strategies provide information about the phase-state of clinically important variables for a measurement occasion but also involve many potential sources of error. For example, reports of previous amounts of alcohol intake, marital satisfaction, depression, and stressful life events prior to an assessment occasion can be influenced by a client's cognitive abilities, biases, the client's current status on those variables, as well as errors in the assessment methods. In sum, data on state-phase functions derived from retrospective reports can be useful but should be cautiously interpreted.
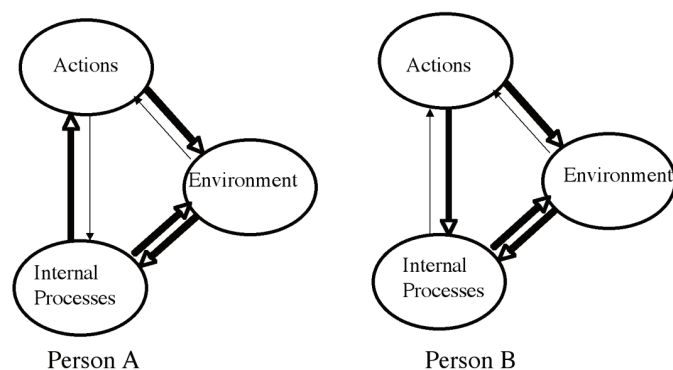
Bidirectional Causal Relations

Bidirectional causality—the idea that two variables can affect each other—is an important concept in causal models of psychopathology. Bidirectional causality can be readily observed in social interactions in which the behavior of one person can affect the behavior of another person, which, in turn, can affect the behavior of the first person. Bidirectional causal relations can also involve interactions between behavioral, cognitive, emotional, and psychophysiological response modes. For example, a person's angry emotions can affect the likelihood that he or she will behave aggressively. The behavioral expression of anger can, in turn, affect the behavior of others toward that person, which then can affect the person's degree of anger, and so on.

The concept of bidirectional causal relations, sometimes referred to as *reciprocal or circular*

causation, has been prominently featured in social learning theories of behavior (e.g., Bandura, 1986). Within social learning paradigms a set of overlapping tenets has been proposed: (a) a person's behavior is influenced by his or her environment, (b) the environment can be influenced by a person's behavior, (c) a person can influence some aspects of his or her environment, to some degree, and (d) internal processes (such as expectancies, biases, attention, attitudes, emotions, biologically based predispositions) affect and are affected by a person's behavior and environment.

As illustrated in Figure 3, more complex bidirectional causal relations, which are closer approximations to the bidirectional causal interactions involved in psychopathology, have also been proposed in social learning paradigms. One such causal model is *triadic reciprocal determinism* (Bandura,1986): A causal model that reflects the fact that there can be causal interactions among a person's actions; cognitive, affective and biological events; and the environment.

There are surplus meanings attached to some alternative terms for "bidirectional causality." Some terms such as *"reciprocal causation," "reciprocal determinism,"* or *"mutual causation"* imply an equality of causal effects between variables that are possible but unnecessary conditions in bidirectional causal relations, as illustrated in Figure 3. The term *"circular causality"* more closely approximates the complex causal chains, also illustrated in Figure 3, that characterize many bidirectional causal relations in psychopathology. Figure 4 illustrates a 3-variable causal chain.

The concept of bidirectional causality has appeared in the writings of philosophers of science for thousands of years. For example, in the third century BC, Aristotle, in his treatise *Physics*, noted: "Some things cause each other reciprocally, e.g. hard work causes fitness and vice versa." (*Physics* by Aristotle, *eBooks@Adelaide 2007)*. Bidirectional causality is also an important element in many domains, such as in the relation between electricity and natural-gas markets (Woo et al., 2006).

RECIPROCAL DETERMINISM IN SOCIAL LEARNING THEORIES

RECIPROCAL DETERMINISM IN PARENT-CHILD INTERACTIONS

Maladaptive Parent: negative, critical, unconstructive; Child Characteristics: negative, critical Behaviors toward parent; Maladaptive child: internalizing and externalizing disorders
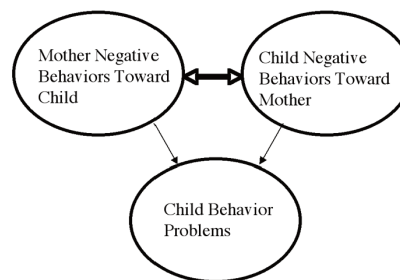
Figure 3: *Triadic reciprocal determinism* in social learning models of behavior. Note that the causal diagrams illustrate *individual differences* in the strength of bidirectional relations (between-person differences in the strength of causal relations between "Actions" and "Internal Processes"). The variable "Internal Processes" refer to cognitive, physiological, and emotional processes.

Figure 4: Bidirectional relations between parent and child negative behaviors and their effect on child behavior problems (Based on Hammen et al.,1990)

A study by Lemay and Clark (2008) illustrates bidirectional causal relations. The authors noted that some persons chronically doubt

their romantic partner's acceptance despite substantial evidence of the partner's love and appreciation. Such doubts can undermine their feelings about, and the quality of, their relationship. Lemay and Clark presented six studies (including a longitudinal and a manipulation study) to test a cognitive *circular causal model*, to account for the perpetuation of insecurity in an intimate relationship. They hypothesized that when person A is insecure about a partner's (person B) acceptance, person A is more likely to express emotional vulnerabilities to person B. Person B will then express positive regard to person A in order to reassure him or her. However, these expressions or reassurance can lead person A to believe that person B expresses positive regard and conceals negative sentiments because he or she views person A as insecure. Further, person A may also develop a belief that person B's expressions of positive regard are inauthentic and that person B actually believes that he or she is emotionally labile and overly dependent on approval. Person A then underestimates person B's true level of positive regard and authentic sentiments which, in turn, increases his or her insecurity.

Several studies have also examined bidirectional relations between marital distress and depression. For example, Choi and Marks (2008) evaluated self-report data from about 1800 Americans, acquired in three waves (1984–1985, 1991–1992, and 2001–2002). The authors evaluated the concurrent and time-lagged correlations (the degree to which measures of one variable were associated with measures of other variables at a later time) among marital disagreements, depressive behaviors, and functional limitations. They found a *positive feedback loop*. In this case, depressive symptoms led to a subsequent increase in marital conflict, which, in turn, led to a subsequent increase in depressive symptoms. Of

relevance to our earlier discussion of causal mechanisms, they also found that marital conflict affected depressive symptoms through two causal paths: (a) directly and (b) indirectly, through its effect on functional impairment, which then affected depressive symptoms.

As noted in Haynes and O'Brien (2000), the concept of bidirectional causality also promotes a positive, constructive focus on the client's goals and behavioral skills during clinical assessment and in the functional analysis. The clinician and client attend to the ways that the client's thoughts or actions may be contributing to his or her behavior problems, and what he or she can do to attain positive intervention goals and maintain positive behavior change over time. Similar to a *task analysis*, and based on the assumptions that behavior problems are a partial function of the client's behavioral repertoire, the clinician can focus on the identification of the patient's skills, skills deficits, and the new skills that are necessary to attain positive treatment outcome.

A Causal Relation Is Still An Inference
Even when the clinician adopts a science-based approach to clinical assessment and uses assessment strategies that are appropriate for the idiographic, multivariate, and dynamic nature of causal relations, an estimate of a causal relation is still a judgment. Consequently, characteristics of the clinician as well as of the variables being measured affect causal judgments. The clinician's judgment of the causal relations associated with a client's behavior problems and goals can be affected by his or her theoretical orientation, selection of assessment instruments that are invalid or insensitive, recent clinical experiences, limitations in knowledge or cognitive abilities, a propensity to use short-cuts in decision-making, and cultural biases. Garb (1998; 2005) reviewed findings on clinicians' cogni-

tive processes in assessment. For example, he noted that making causal judgments can be extremely difficult for clinicians (cf. also Waldmann, 2010) and that clinicians often decide on a diagnosis by comparing clients to a prototype rather than by following a hypothetico-deductive approach to identifying signs and symptoms associated with particular diagnostic categories. As stressed in Haynes et al. (2011a) we presume that judgments about causal relations in clinical assessment are most likely to be valid and beneficial to the client when based on sound clinical assessment evidence and when the clinical assessment and judgment process is guided by scientific principles and findings.

Summary

Inferences about causality are critical components of assessment and case formulation in psychopathology. Causal inferences are important because they allow the clinician to identify variables that influence behavior problems and positive alternatives. Once causal variables and causal relations are identified, a clinician can design treatments aimed at modifying them in order to bring about improved client functioning.

While important in assessment, functional analysis, and other models of clinical case formulation, the identification of causal variables and the measurement of causal relations is a complex task involving many component judgments and decisions. The complexity partially arises from the diverse types of causation and the necessary conditions for inferring causal relations. A necessary and sufficient relation would be considered the strongest evidence for causation between two variables. However, most causal relations encountered in clinical assessment contexts are neither necessary nor sufficient. Instead, complex chains of probabilistic interactions are commonplace.

Further complicating matters, causal effects can differ in their latency of effects. Finally some causal effects occur only when one or more other causal events have occurred.

In order to identify causal relations, a clinician must evaluate four characteristics of a relation between two variables: (a) covariation, (b) temporal precedence, (c) the exclusion of a plausible alternative explanation for the covariation and precedence, and (d) the logical basis for inferring a causal relation. These four criteria for causal inference require complex measurement strategies that target multiple constructs, are able to take into account dynamic variation among variables, and can assess the possibility of bidirectional causality.

A functional analysis requires that we identify the mechanism of the causal relation—the means through which a causal effect operates. Essentially, we are asking the question, "How does X influence Y?" The answer to this question can improve our causal models of psychopathology and lead to more powerful and efficient interventions. However, the identification of the causal mechanism is complicated by several aspects of causal relations: Causal relations are often *multivariate, dynamic, bidirectional,* and they *share variance* amongst causal variables in affecting a behavior problem. Dynamic and bidirectional causal relations are particularly important components of causal models of psychopathology. They inform the clinician about what assessment strategies might be most useful in identifying causal relations; they promote the frequent implementation of sensitive-to-change assessment strategies and measures, and encourage a focus on individual differences in clinical case formulation. Although causal inferences in clinical assessment are based on the clinician's judgments, the validity and utility of those judgments will be strengthened to the degree

that the clinician uses sound clinical assessment strategies.

References

Allen, D. M. (1993). Unified therapy. In G. Stricker & J. Gold (Eds.), *Comprehensive handbook of psychotherapy integration* (pp.125-137). New York: Plenum Publishers.

Asterita, M. F. (l985). *The physiology of stress*. New York: Human Sciences Press, Inc.

Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R. & Fiore, M. C. (2004). Addiction motivation reformulated: An affective processing model of negative reinforcement. *Psychological Review, 111*, 33-51.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Beach, S. R. H., Sandeen, E. E., & O'Leary, K. D. (1990). *Depression in marriage: A model for etiology and treatment*. New York: Guilford Press.

Beck, A., & Brad, A.A. (2009). *Depression: Causes and treatment* (2nd ed.). Philadelphia, PA: University of Pennsylvania Press.

Bryant, R. A. (2006). Post-traumatic stress disorder. In F. Andrasek (Ed.), *Comprehensive handbook of personality and psychopathology* (pp. 189-206). Hoboken, NJ: Wiley.

Bunge, M. (2009). *Causality and modern science*. New York: Dover.

Choi, H. & Marks, N. F. (2008). Marital conflict, depressive symptoms, and functional impairment. *Journal of Marriage and Family, 70,* 377-390.

Danks, D. (2005). The supposed competition between theories of human causal inference. *Philosophical Psychology*, *18*, 259–272.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment.* Washington, DC: American Psychological Association.

Doherty, W. J., Kouneski, E. F., & Erickson, M. F. (1998). Responsible fathering: An overview and conceptual framework. *Journal of Marriage and the Family, 60*, 277–292.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment.* Washington, DC: American Psychological Association.

Garb, H. N. (2005). Clinical judgment and decision making. *Annual Review of Clinical Psychology, 1,* 67-89.

Giorgio, J. M., Sanflippo, J. Kleiman, E., Reilly, D. Bender, R. E., Wagner, C.A., Liu, R. T., & Alloy, L. B. (2010). An experiential avoidance conceptualization of depressive rumination: Three tests of the model. *Behaviour Research and Therapy*, *48*, 1021-1031.

Hammen, C., Burge, D., & Stansbury, K. (1990). Relationship of mother and child variables to child outcomes in a high-risk sample: A causal modeling analysis. *Developmental Psychology, 26*, 24-30.

Hasking, P., Momeni, R., Swannell, S. & Chia, S. (2008). The nature and extent of non-suicidal self-injury in a non-clinical sample of young adults. *Archives of Suicide Research, 12,* 208-218.

Haynes, S. N. (l992*). Models of causality in psychopathology: Toward synthetic, dynamic and nonlinear models of causality in psychopathology*. Des Moines, IA: Allyn & Bacon.

Haynes, S. N., & Yoshioka, D. T. (2007). Clinical assessment applications of ambulatory biosensors. *Psychological Assessment, 19*, 44-57.

Haynes, S. N., Blaine, D., & Meyer, K. (1995). Dynamical models for psychological assessment: Phase-space functions. *Psychological Assessment, 7,* 17-24.

Haynes, S. N. Mumma, G. H., & Pinson, C. (2009). Idiographic Assessment: Conceptual and Psychometric Foundations of Individualized Behavioral Assessment. *Clinical Psychology Review, 29,179-191.*

Haynes, S. N. & O'Brien, W. O. (1990). The functional analysis in behavior therapy. *Clinical Psychology Review, 10,* 649-668.

Haynes, S. N., & O'Brien. W. O. (2000). *Principles and methods of behavioral assessment*: New York: Plenum/Kluwer Press.

Haynes, S. N., O'Brien, W. H., & Kaholokula, J. K (2011a). *Behavioral assessment and case formulation*. Hoboken, N.J.: John Wiley & Sons.

Haynes, S. N., Smith, G., & Hunsley, J. (2011b). *Scientific foundations of clinical assessment*. New York: Routledge.

Hyland, M. E. (1987). Control theory interpretation of psychological mechanisms of depression: Comparison and integration of several theories. *Psychological Bulletin, 102*,109-121.

James, L. R., Mulaik, S. A. & Brett, J. M. (1982). *Causal analysis; Assumptions, models and data*. Beverly Hills: Sage.

Kazdin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston: Allyn and Bacon.

Lemay, E. P., Jr., & Clark, M. S. (2008). "Walking on eggshells": How expressing relationship insecurities perpetuates them. *Journal of Personality and Social Psychology, 95*, 420-441.

Magnavita, J. J. (2006). In search of the unifying principles of psychotherapy: Conceptual, empirical, and clinical convergence. *American Psychologist, 61,* 882-892.

Minke, K., & Haynes, S. N. (2011). Sampling issues. In J. C. Thomas & M. Hersen (Eds.). *Understanding research in clinical and counseling psychology*. Mahwah, NJ: Lawrence Erlbaum Associates.

Newsome, G.L. (2003). The debate between current versions of covariation and mechanism approaches to causal inference. *Philosophical Psychology, 16,* 87-107.

Nezu, A. M., Nezu, C. M., & Lombardo, E. (2004). *Cognitive-behavioral case formulation and treatment design: A problem solving approach*. New York: Springer.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, England: Cambridge University Press.

Reid, J.B., Patterson, G.R., and Snyder, J. (Eds.) (2002). *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention*. Washington, DC: American Psychological Association.

Rehman, U., Gollan, J., Mortimer, A. (2008). The marital context of depression: Research limitations and new directions. *Clinical Psychology Review, 28*, 179-198.

Shadish, W.R., Cook, T.D. & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York: Houghton Mifflin.

Smits, J. A. J., O'Cleirigh, C. M., & Otto, M. W. (2006). Panic and agoraphobia. In M. Hersen & J. C. Thomas (Eds.), *Comprehensive handbook of personality and psychopathology: Adult psychopathology* (Vol. 2, pp. 121-137). Hoboken, NJ: John Wiley & Sons, Inc.

Stulz, N., & Crits-Christoph, P. (2010). Distinguishing anxiety and depression in self-report: Purification of the Beck Anxiety Inventory and Beck Depression Inventory-II. *Journal of Clinical Psychology, 66*, 927-940.

Wagner, E. F., Sobell, M. B. & Sobell, L. C. (2007). Substance-related disorders: Alcohol. In M. Hersen, S. M. Turner, & D. C. Beidel (Eds). *Adult psychopathology and diagnosis* (5th ed., pp.166-200). Hoboken, N.J.: John Wiley & Sons.

Waldmann, M. R. (2010). Causal thinking. In B. M. Glatzeder, V. Goel, and A Muller (Eds.), *Towards a Theory of Thinking, part 2* (pp. 123-134). Heidelberg: Springer.

Wise, D. (2006). Child abuse assessment. In M. Hersen (Ed), *Clinician's handbook of child behavioral assessment* (pp. 549-568). San Diego, CA: Elsevier Academic Press.

Woo, C-K., Olson, A., Horowitz, I., & Luk, S. (2006). Bi-directional causality in California's electricity and natural-gas markets. *Energy Policy*, *34* (15), 2060-2070.

Endnotes

[1] The functional analysis is one model for clinical case formulation and refers to "*The identification of important, controllable, causal and noncausal functional relations applicable to specified behaviors for an individual*" (Haynes & O'Brien, 1990).

[2] In this example, measures of the constructs can show significant covariation but the error would be in inferring a causal relation based on that covariation.

[3] Causal effects could also be additive, in that "depressed" interpersonal actions could affect a spouse only after occurring for an extended period of time.

[4] We illustrate in other sources (Haynes et al., 2011a) that *no harm* would be done to the client by erroneously inferring a causal relation between the behavior problems, as illustrated in Figure 1. Given the use of appropriately validated and focused intervention strategies, a failure to identify the most important and modifiable causal variables in a functional analysis only diminishes the benefits to the client of the intervention.