FACULTY OF BU

RESEARCH AND WORKING PAPER SERIES

AN INTEGRATION OF KELLEY'S COVARIATION AND CONFIGURATION THEORIES.

John W. Medcof
McMaster University

Working Paper No. 248
March, 1986

M R UNIVERSITY
1 t West
H o, Canada L8S 4M4
T) 525-9140

AN INTEGRATION OF KELLEY'S
COVARIATION AND CONFIGURATION
THEORIES.

John W. Medcof
McMaster University

Working Paper No. 248
March, 1986

Running Head: Integration of Kelley

Abstract

Kelley's two attribution theories are much more intimately linked than has previously been demonstrated. A core probability model is presented which shows that covariation theory and configuration theory operate upon the same basic principles and that identical terminology can be applied to them.

An Integration of Kelley's Covariation and Configuration Theories

Attribution theory (see Harvey and Weary, 1984; and Kelley and Michela, 1980; for recent reviews) is a body of ideas in Psychology which attempts to describe how observers interpret the events they see in everyday life. In the past twenty or so years attribution theory has enjoyed considerable attention in the literature, has generated a great deal of research and has been shown capable of explaining a wide range of empirical phenomena.

Despite this success (or, perhaps, because of it) attribution theory has remained a rather loose federation of theories and research. Recently the three most influential theorists have been Jones and Davis (1965), Kelley (1973) and Weiner (1972), but there are a great many attribution phenomena which fall outside the bounds of these theories. The grouping of these theories and phenomena under the title, "attribution theory", is widely accepted but is based as much upon a rather vague intuition that they belong together as upon any systematic demonstration that there is a common, elegant core of ideas underlying them all.

Although attribution research has enjoyed considerable success in its vague form, the development of an elegant conceptual core for it is a desirable aim. The primary reason for this is that such a core is one of the signs of a mature science. Physics, for example, is characterized by being able to describe and explain a wide variety of empirical phenomena, and by being able to conceptually relate most of those explanations to a single elegant statement, $E = MC^2$. Furthermore, the relationship between the conceptual core and the other concepts and theories is clearly and explicitly drawn, it is not left to vague intuition. As Psychology matures as a science it must develop itself in both of these ways as well, explaining and predicting a wider and wider range of empirical phenomena and refining its conceptual core. At this point in time Psychology does not have a single conceptual core but it does have theories. It is these which must be refined.

There can be little doubt that the aim of establishing a clear and elegant core for attribution research is believed to be a desirable and attainable end by a number of researchers. It is generally accepted that most of what we call attribution research and theory springs ultimately from the work of Heider (1958). But although Heider is accepted as the intellectual forefather of this field of study, Heider's work, in its richness and suggestiveness, is not necessarily elegantly stated. Neither are the connections between present activities and Heider's work necessarily drawn with much clarity. Consequently, a number of writers have attempted to show more clearly the links between various parts of attribution theory, including Ajzen and Fishbein (1975), Jones and McGillis (1976), Kruglanski (1980) and Read and Stephan (1979).

In keeping with this spirit of conceptual clarification this paper will describe a single elegant core for both parts of Kelley's attribution theory, covariation and configuration. The idea that these two theories are intimately connected is not new (Kelley, 1973). What has not yet been done, though, is a thorough demonstration of this, showing in detail, the interchangeability of concepts and terminology. This analysis will make it clear that the two theories need not be described separately, as Kelley has done in the past (Kelley, 1973; and Kelley and Michela, 1980), but can be treated as a single theoretical entity.

To accomplish this task within the space of a single journal article will require some clear restrictions on what is discussed. Kelley's two theories are primarily concerned with causal attributions, although, as will be seen below, in some places attributions about dispositions come into his analysis. This paper will, therefore, focus primarily upon the causal attribution aspects of Kelley's theories and avoid, wherever possible, the discussion of attributions of dispositions. At two places in this article the attribution of dispositions will be discussed because it is absolutely necessary to do so in order to do justice to Kelley's theory. However, these discussions will be as brief as possible. One thing that will become clear in these brief discussions, though, is that Jones and Davis', as well as Kelley's theories, can be reduced in a clear and systematic way to the core model of attributional processes which is presented here.

The primary conclusion of this paper will be that covariation theory and configuration theory are much more closely connected than has been previously demonstrated. The two theories are more or less interchangeable, although they do tend to emphasize different aspects of the same attribution processes. It will also be shown that the core model presented here points out some things missed in past discussions of both of Kelley's theories. The primary such demonstration will be that there is a free-flowing and intimate interaction between the attribution of dispositions and the attribution of cause.

The Core Model

The first step in the integration of Kelley's two theories will be the presentation of a core model. This core model is a brief set of assumptions about how observers store and use information about their environments which will be shown to underlie both covariation theory and configuration theory.

The core model will borrow some terminology from Kelley (1973). As in his system, causal agents can be actors or entities. Actors are, of course, people; and entities are non-human things. However, an entity can be a specific, non-human thing, such as an automobile or an animal, or it can be a complex configuration of specific entities and/or persons. For example, a cocktail party is a situation which consists of a number of actors (host, hostess, friends, acquaintances, etc.) and entities (drinks, food, too few chairs, etc.). Such complex entities will be called situations. So the term, entity, will refer to both specific, individual entities and to situations. Also, the word agent will be used to refer generically to any possible or actual causal agent, be it actor, specific entity or situation.

The first major assumption of the core model is that people observe the world and store their observations as probability statements. For most events that they are aware of observers will develop some impression of how likely those events are, across a variety of circumstances. For example, people have an impression of how likely snowstorms are, and they will remark if in a particular year snowstorms are more frequent than usual. This will occur without any conscious attempt to systematically record and compare the frequencies of snowstorms. These impressions of likelihood are not always accurate but they are there and people spontaneously make statements about them. The core model assumes that these beliefs about the general probabilities of events are stored as unconditional probabilities, of the form p(event) = X. The core model assumes further that observes note that the probabilities of some events are higher in the presence of some agents than they are in the presence of others. For example, snowstorms are more probable given the presence of clouds than they are in the presence of sunshine. These observations are stored in the form of conditional probabilities, p(event/A) = X. Observers are capable of storing probabilities involving multiple conditions as well. For example, they may have some impression of p(accident/rain $\cap$ night $\cap$ narrow road). These probability statements, which summarize past experience and can be used to predict future events, are also used by observers as a basis for labelling actors and entities and for attributing cause.

Labelling Actors and Entities

One assumption of the core model is that the labelling of actors and entities is intimately related to the attributing of cause. Therefore, it is useful to discuss this labelling process before discussing cause. Labels which are applied to actors will be called dispositions; those applied to entities will be called constraints.

Dispositions. The assignment of dispositions to actors begins with observers perceiving reliable associations between specific individuals and certain events or behaviours. These associations are stored as probabilities. Observers notice that Jack Smith, for instance, often gets into arguments. This can be stored as, p(argument/Jack Smith) = .75. On the other hand, Jack's brother Bill seldom gets into arguments, or, p(argument/Bill Smith) = .05. In addition, observers can compare probabilities and notice whether the probability of an event, given the presence of a particular actor, is greater or less than the unconditional probability of the event. If p(argument) = .45, then Jack Smith is above this norm and Bill Smith is below it.

When observers perceive that the probability of an event, given a particular actor, is greater or less than the unconditional probability, they tend to say that that person has a certain trait or disposition. They label Jack Smith as "argumentative" and Bill Smith as "nice". Such assertions seem to provide some observers with the feeling that they "understand" Jack and Bill Smith. The actions of the Smith brothers are "explained" by the fact that they have certain traits or dispositions. But these are pseudo-explanations. They are really just labels which reflect observers' beliefs about the past and future behaviours of the individuals observed. A general definition of disposition will now be given. So far, the only events discussed here have been human behaviours. However, other kinds of events might also be associated with an individual, for example, a car crash. For this reason, this general definition is stated in terms of events, rather than in terms of the particular kind of event which is of most interest here, human behaviour.

An actor is said to have a DISPOSITION if:

$p(\text{event/actor}) \neq p(\text{event})$

It is further assumed that the greater the deviation of the actor's probability from the unconditional probability, the stronger the actor's disposition is perceived to be. For example, if the p(argument/Jack Smith) = .75, the p(argument/Bill Bailey) = .95, and the unconditional probability of argument is .45, then observers will say that Bill Bailey is more argumentative than Jack Smith. This gives a second definition.

The perceived STRENGTH OF AN ACTOR'S DISPOSITION is directly proportional to:

$|p(\text{event/actor}) - p(\text{event})|$

This model of the nature of dispositions allows observers to perceive individual actors as influencing the probabilities of events. A particular actor can be perceived to have a disposition which makes a particular kind of event very probable, or to have a disposition which makes a particular kind of event very unlikely. In the former case the actor may be perceived to be the cause of the event, in the latter the actor may be perceived to be preventing the occurrence of the event. These processes will be discussed more fully below when attributions are described.

Different observers may assign different disposition labels to the same actor and these differences may come from a number of sources. Most obviously, if different observers have different beliefs about p(event/actor) they are likely to assign different dispositions. But differences in perceived p(event) could lead to different disposition labels, even when there is agreement upon p(event/actor). This is because perceived dispositions depend upon variations around p(event). Differences in the nature of disposition labels themselves could also influence the labelling process. Sometimes disposition labels designate quite obviously the particular behaviours being associated with an actor, eg. talkative. Very often though, the disposition term refers to a class of behaviours which are not clearly designated, eg. careless. When such general terms are used, observers will have to do more interpreting to determine if a particular behaviour fits into a class of behaviours. This may lead to some variation between observers in what is stored and therefore in statements about what is stored.

Presumably, observers' beliefs about the probabilities of certain behaviours, given a particular actor, develop over time. When observers get their first piece of information about an actor they establish probability statements about that individual. When subsequent information is received, by direct experience with the person or through indirect sources, the probability statements are adjusted appropriately. Sometimes the information about an individual may be rather fragmentary. For example, as one hears about Joe Smith from first one person and then another, and another, the probability statements are developed by the successive pieces of information. In some cases, though, comprehensive probability statements about an individual may be acquired all at once. Stereotyping is just such a case. Stereotypes usually include statements about the physical characteristics of the individual involved (eg. female, black) and statements about their typical or most probable behaviours (eg. non-dominant, musical). On this model a stereotype is a collection of probability statements about the kinds of behaviours likely to be emitted by certain classes of people. Once a person is identified as belonging to a certain stereotyped group, a whole host of assumptions about behaviour probabilities is made. These behaviours are sometimes "explained" by stating that the person is a member of the stereotyped group, for example, "She acts that way because she is a women". This kind of explanation operates on the same basis as explaining behaviour in terms of dispositions. To the user it seems to explain, but it is really just a statement about the user's belief about the probability that some behaviour will occur.

Constraints. The assignment of constraints to entities follows the same general principles as the assignment of dispositions to actors. Observers notice that actors tend to act in predictable ways when in the presence of certain entities. For example, there may be a great deal of audience laughter at a particular movie. The movie is therefore likely to be labelled a comedy. This constraint label indicates a certain probability of laughter associated with this particular movie. Such constraint labels are used when the probability of the event, given the presence of the entity, is different from the unconditional probability. The definitions for constraint labelling are therefore analogous to those for disposition labelling.

An entity is said to have a CONSTRAINT if:

$p(\text{event}/\text{entity}) \neq p(\text{event})$

The perceived STRENGTH OF AN ENTITY'S CONSTRAINT is directly proportional to:

$|p(\text{event}/\text{entity}) - p(\text{event})|$

Observers use constraint labels in the same way that they use disposition labels. They are based upon the probability statements which are used to summarize the past and predict the future. Constraint labels seem to provide explanations of events when they are really just summaries of observations. Observers get their beliefs about constraints from both direct experience and by being informed by other people. Those beliefs can come from a cumulation of experiences or from a single overall statement analogous to the stereotype. For example the "stereotypical" haunted mansion has a whole host of human behaviours and other events associated with it. Observers assume that the constraint of an entity makes particular events very likely or very unlikely. In the former case the entity may be seen as the cause of the event, in the latter case the entity may be seen as preventing or inhibiting the occurrence of the event.

The process of labelling agents with dispositions or constraints is analogous to attributing dispositions as that process is described by Jones and Davis (1965). It has just been shown that in the core model this is based upon stored probability statements. Shortly it will be shown that causal attributions can also be dealt with in terms of probabilities. One important advantage of the core model is that both kinds of attributions are reduced to a common basis, probabilities. With this common basis it can be seen that the two kinds of attributions operate in close cooperation with one another.

This discussion of dispositions and constraints has given an important place to the idea of the unconditional probability of an event. This idea is vague and useful theoretically but may be difficult to define precisely in an operational way. Presumably there are population norms for smiling, walking, talking, weeping, etc., but measuring these norms across all actors and all entities is practically impossible. In addition, any group of real observers will probably not be aware of the true probabilities of the occurrence of events, but will have only beliefs, which they have acquired through their own experiences. Researchers interested in exploring the specifics of perceived norms and their relationship to dispositions, constraints, and their relative strengths, would have to empirically establish the local norms of the populations they are using. Researchers not interested in that particular question but who want to manipulate strength of disposition or constraint, or who want to establish perceived constraints or dispositions, will have to work on a trial and error basis. For example, to establish in observers' minds the perception that an actor is dour, the actor, when presented to observers, would have to exhibit extreme behavioural signs of dourness, to ensure that the population norm (whatever it might be) is avoided. Trial and error and manipulation checks could refine this process.

A summry of this discussion of labelling with dispositions and constraints goes as follows. Observers store information about their observations in the form of probability statements. These statements are used to label actors with dispositions and entities with constraints. These disposition and constraint labels are treated as explanations by most observers, but are really just verbal labels attached to probability statements which summarize beliefs. The cognitive mechanisms for establishing and using probability statements are the same for both dispositions and constraints, the only difference is in the kinds of agents referred to.

<u>Attributing Causes of Events</u>

A major assumption of the core model is that when observers make causal attributions they base the attributions upon the same kind of stored probability statements that they use when labelling dispositions and constraints. This assumption provides a basis for integrating the two parts of Kelley's theory, as will be shown below, and provides a point of contact between attribution theory and theories dealing with labelling processes such as stereotypes (eg., Buss and Clark, 1983).

Observers base their causal attributions upon these stored probabilities because their concept of cause is based upon predictability. If the presence of an agent is a reliable predictor that a certain kind of event is likely to occur, observers will be prepared to state that that agent caused the event. For example, if observers notice repeatedly that when the temperature (an agent) is below 0^{o} C, water turns to ice, they are prepared to say that the cold caused the ice to form. This basis for assigning cause is far from perfect. For example, the ancients discovered that the act of throwing a virgin into an overly active volcano was usually followed, sooner or later, by a return to the normal level of volcanic activity. They therefore assumed that it was the sacrifice which caused the return to normal. They did not adopt the interpretation that the volcano would have returned to normal, sacrifice or no sacrifice. The repeated co-occurrence of the two events was sufficient to convince them of cause. Extended observation and experimentation my lead to an improved understanding of what the "real" causes of an event are, and may even lead to complicated scientific theories involving chains of causes, but the intuitive practice of the naive observer is to base statements of cause upon predictability.

Given that observers base their causal attributions upon predictability, and that predictability is based upon stored probability statements, it follows that the information in the probability statements is used to make causal attributions. The basic mechanics of that process are as follows. It has already been established that observers store information about the unconditional probabilities of events, p(event); and conditional probabilities of events, p(event/agent). If observers believe that p(event/agent) is greater than p(event), they have a basis for predicting that the event will occur. For example, if I store the information that people seldom stare at me on the street, p(stare) = .001; but that people stare at me a great deal when I wear my clown suit on the street, p(stare/clown suit) = .999; then I can predict that if I wear my clown suit people will stare at me. Further, I will conclude that it is the clown suit which causes the staring. In essence, if the stored probabilities suggest that the presence of an agent is an indication that there is an above normal level of probability that a particular kind of event will occur, and on a particular occasion that event occurs in the presence of that agent, observers are prepared to state that in that particular case the agent caused the event.

Once such a causal agent is identified observers usually feel that they have an adequate explanation for the occurrence of the event. That agent is a sufficient cause of the event. Given this, the following definition holds.

An agent will be perceived to be a SUFFICIENT CAUSE of an event when:

p(event/agent) > p(event)

If there is only one agent present, and that agent is a sufficient cause of the event, observers will accept it as the cause. But in most cases there are multiple agents available and the decision about cause is not that simple. When several agents are available, observers will scan the probability statements for each of those agents and name as the cause of the event the agent whose presence is the best predictor of the event. For example if a teacher is writing on the blackboard with his back to the class and is hit on the back of the head with a piece of chalk, he must decide which of his several students is the cause of the thrown chalk. He has several possibilities, looking over his beliefs about the past behaviours associated with each of his pupils. Little Willie Frolic has never done anything bad in his life (at least in the mind of the teacher) so p(chalk missle/Willy) = .001. Mary Smith is a more likely suspect, p(chalk missle/Mary Smith) = .40. But she is not nearly as bad as her brother, Jack, p(chalk missle/Jack Smith) = .99. Jack Smith has the highest probability and is therefore the most likely cause. The teacher will probably blame his for the chalk missle.

In some case, though, an agent may not be a predictor that the event is more likely than usual, the agent may be a predictor that the event is less likely than usual. Observers will be prepared to believe that such agents tend to cause the event not to occur, that such agents tend to prevent the occurrence of the event. For example, low temperatures generally predict that ice cream sales will be less than average, so observers will be prepared to attribute lower sales on a particular occasion to the cold weather (if the cold weather is present on that occasion). This leads to the following definition.

> An agent will be perceived to be a SUFFICIENT PREVENTOR of an event when
>
> p(event/agent) < p(event)

In situations in which several preventors are available, observers will tend to assign blame for the below normal occurrence of the event to the preventor which is the best predictor of non-occurrence.

Observers will attempt to make causal attributions which are consistent with all of the stored relevant probability statements, but sometimes an event occurs which is not consistent with the stored information. If an event occurs (or does not occur) which is contrary to the expectations provided by the stored probability statements, observers will tend to deal with this unexpected occurrence in one of three ways. The first is to accept the event as unexpected and to attribute it to luck, chance, circumstances, or some other ill-defined or hypothetical agent. The second course is to revise the probability statements for one or more of the agents present. The revised probability statement(s) can then be used to explain the event. This revision of probability statements involves a change in a perceived disposition or constraint of an agent. The third alternative for observers faced with events which are contrary to expectations is to make use of agents which are present but about which little is known. For example, suppose Kent is known to be a very poor tennis player and engages in a match with a new club member about whom nothing is known. If Kent wins, people will immediately attribute the disposition, "poor tennis player", to the new club member, even if they have never seen him play. In the core model this means they will establish a probability statement for the new member, perhaps p(win/new member) = .05. Observers prefer to apply this attribution to this hitherto unknown person than to revise their beliefs (probabilities) about Kent, in order to explain the win. These latter two methods of dealing with unexpected events involve the formation and/or change of attributions of dispositions (or constraints), a topic treated in Jones and Davis' (1965) theory of attributions.

Now that the basic processes for attributing cause have been described, the implications for more complex situations can be considered. To do this, situations involving two agents will be analysed, for illustrative purposes, because they are relatively easy to understand. The same basic principles can be applied to three or more agents. A further advantage of illustrating using two agents is that Kelley's theories are, for the most part, also stated in terms of two agents, so the correspondences between the core model and Kelley's will be clearer.

The following analysis will show that attributions can be classified into four basic types, and that each type is characterized by a particular core pattern of probability statements. When two agents are involved (represented by A and B), and they are known to the observers (i.e. observers have stored probability statements about them), the patterns are made up of four probability statements, p(event), p(event/A), p(event/B) and $p(\text{event}/A \cap B)$. It is the relative magnitudes of these probabilities which are crucial in determining which attributions are made.

The general rules that observers follow when determining attributions will now be described. These general rules will be described for situations in which observers have observed an event and now wish to ~~exploit~~ explain it. Below, it will be shown that the same general logic is used when the non-occurrence of an event is to be explained. When asked to explain an event observers will scan their stored probabilities to determine if the occurrence of the event is predicted given the presence of both of the known agents. In other words they will check to see if p(event/A $\cap$ B) > p(event). If this is so, they then scan their probabilities to see which of the known agents, alone or in combination, best predicts the occurrence or of the event. The agent or agents which is (are) the best predictor(s) will be named as the cause(s). If the initial scan shows that this event is not predicted given the presence of the known agents, i.e. p(event) > p(event/A $\cap$ B), observers will use one of the three explanations for unexpected events which were described above, attribute to chance, change perceived dispositions and/or constraints, or attribute to an unknown agent. Of the four types of attribution to be described here, the first three involve expected events and yield attributions in terms of the known agents. The fourth pattern involves unexpected events.

<u>Type I, Single Cause</u>. The pattern of probabilities for this type of attribution is:

1. p(event/A) $\geq$ p(event/A $\cap$ B) > p(event)
2. p(event/A) > p(event/B)

Given this pattern of stored information, and the knowledge that the event occurred while A and B were present, observers will make their causal attributions as follows. A scan of the probabilities in Statement 1 shows that p(event/A $\cap$ B) > p(event). This shows that the event is predictable in these circumstances and a causal explanation in terms of the known agents present is possible. Now observers must decide which specific agent or combination of agents is the actual cause of the event. Another scan of Statement 1 shows that p(event/A) > p(event) and p(event/A $\cap$ B) > p(event). Thus both A and A $\cap$ B are sufficient causes of the event. However, statement 1 shows that in some cases p(event/A) > p(event/A $\cap$ B). When this happens A is preferred over A $\cap$ B as the cause because it is the better predictor of the event. Statement 1 also shows that in some cases p(event/A) = p(event/A $\cap$ B). When this occurs A is still preferred over A $\cap$ B as the cause. This is because the core model assumes that when two probability statements provide equal predictability, observers will choose the one with fewer conditionals as the cause of the event. This is not only parsimonious, but also, the one with fewer conditionals is the better general predictor of the event across a variety of circumstances. The two conditionals mean that this prediction is restricted to cases in which both agents are present and this set of circumstances will usually be smaller than the set in which only one agent is present. So the scan of statement 1 shows that A is preferred as the cause over A $\cap$ B. A scan of statement 2 shows that p(event/A) > p(event/B). Again A will be preferred as the cause because it is the better predictor. So, given the Type I pattern of probabilities, observers will name agent A as the cause of the event.

An example of observations fitting Type I is as follows. Suppose the event to be explained is the fact that Jack got a good catch of fish in a particular creek this morning. Jack is one agent and the creek is the other. Suppose that the observers believe that if one goes fishing the probability of getting a good catch is about 50-50, or, p(event) = .50. Suppose also that the observers believe, on the basis of their own experiences and from talking to others, that when people fish in this particular creek the probability of a good catch is relatively low, p(good catch/creek) = .20. In other words this creek is a poor fishing hole. Suppose that the observers believe that Jack is a good fisherman, from their own observations and from talking to Jack and others. This might be represented as, p(good catch/Jack) = .95. Finally, suppose the observers believe that when Jack has fished in this particular creek in the past he has done very well, p(good catch/creek $\cap$ Jack) = .80. When asked to name the cause of Jack's good catch this morning the observers will scan their probabilities and find that they fit the essentials of Type I.

1. p(event/Jack) = .95 > p(event/Jack $\cap$ creek) = .80 > p(event) = .50
2. p(event/Jack) = .95 > p(event/creek) = .20

Jack will therefore be named as the cause because the presence of Jack is the best predictor of the event. The choice of Jack can be determined strictly by looking at the probabilities, but if observers were asked to explain their attributions they would probably reply somewhat as follows. This is a really bad fishing creek but Jack got a good catch in it today for the same reason he usually does, he is a really good fisherman.

Although this particular example has concerned the predictable occurrence of an event, p(event/Jack $\cap$ creek) > p(event), the same logic can be found in cases involving a predictable non-occurrence of an event. In these latter cases, p(event) > p(event/A $\cap$ B) and, as predicted by this probability statement, the event did not occur on this occasion in these circumstances. The logic of this kind of interpretation can be demonstrated by modifying the example of Jack and the fishing creek. In the previous example p(event/Jack $\cap$ creek) = .80 > p(event) = .50. Suppose that instead it happens that p(event) = .50 > p(event/Jack $\cap$ creek) = .30. The other elements of the example can be left unaltered, p(event/Jack) = .95, p(event/creek) = .20. Suppose that this morning Jack did not get a good catch. That these circumstances represent a Type I pattern is best seen if the event is redefined to be the lack of a good catch. Given this, p(event/Jack) = .05, p(event/creek) = .80, p(event/Jack $\cap$ creek) = .70 and p(event) = .50. This information is easily rearranged to show that it is congruent with Type I.

1. p(event/creek) = .80 > p(event/creek Jack) = .70 > p(event) = .50
2. p(event/creek) = .80 > p(event/Jack) = .50

This pattern of probabilities suggests that Jack's poor catch today will be attributed to the fact that the creek is a poor fishing place. If asked to explain this attribution, observers would probably reply that although Jack is generally a good fisherman he, like everyone else, does poorly in this creek because it is a poor fishing place. Today's poor catch is due to the poor nature of the creek.

This example shows that the Types being described in the core model apply to both the occurrence and non-occurrence of events. In the first fishing creek example Jack was a sufficient cause of the event, p(event/Jack) = .95 > p(event) = .50, and the creek was a sufficient preventor of the event, p(event) = .50 > p(event/creek) = .20. However, in that first example, when Jack and the creek had come together in the past, it was Jack who had had the more potent influence upon the outcome, as is shown by, p(event/Jack ∩ creek) = .80 > p(event) = .50. Thus when the event occurred this time it is attributed to Jack. In the second fishing creek example Jack is still a sufficient cause, p(event/Jack) = .95 > p(event) = .50, and the creek is still a sufficient preventor, p(event) = .50 > p(event/creek) = .20, but now it is the creek which has been the most potent influencer of events in the past, as is shown by, p(event) = .50 > p(event/Jack ∩ creek) = .30. Jack's lack of a good catch this time is consistent with these past events and so the poor catch is attributed to the creek.

Although the two core statements for Type I specify the essentials for that type, there are several variations of it. The core statements do not state explicitly the relationship between p(event/B) and any of the other probabilities except that p(event/A) > p(event/B). Therefore, p(event/B) could assume any of several relationships with p(event) and p(event/A $\cap$ B). Three variants upon the core statements, beginning with that used in the fishing creek example are:

p(event/A) > p(event/A $\cap$ B) > p(event) > p(event/B)

p(event/A) > p(event/A $\cap$ B) > p(event/B) > p(event)

p(event/A) > p(event/B) > p(event/A $\cap$ B) > p(event)

In all of these, and their non-occurrence equivalents, agent A alone will be named as the cause of the event.

Type II, Synergy. The core pattern of probabilities for this type is:

p(event/A $\cap$ B) > p(event/A), p(event/B), p(event)

In this pattern the only requirement is that p(event/A $\cap$ B) be greater than all of the other probabilities. No particular relationship is required between any of the others. Since p(event/A $\cap$ B) > p(event) this event is predictable in these circumstances and observers will look for a causal explanation in terms of the known agents present. Scanning the probabilities shows that p(event/A $\cap$ B) is greater than all of the others, and therefore the best predictor of the event. The event is therefore attributed to the joint presence of the two agents.

An example of Type II goes as follows. Suppose Melda and Melvin are two children who got into trouble this morning by painting a cat with blue paint. The event to be explained is this mischief. Suppose the observers have the following stored probabilities. Generally speaking, Melda is a good little girl, p(mischief/Melda) = .20, and Melvin is a good little boy, p(mischief/Melvin) = .25. Suppose that in general children get into mischief more often than these two angels, p(mischief) = .35. However, in the past, when Melda and Melvin have played together, they have got into quite a bit of mischief, p(mischief/Melda $\cap$ Melvin) = .65. These probabilities can be arranged to show that they fit the pattern for Type II: p(event/Melda $\cap$ Melvin) = .65 > p(event/Melda) = .20, p(event/Melvin) = .25, p(event) = .35. The event is therefore best predicted by the joint presence of the two agents and they will be named as the cause. If the observers were the distraught parents they would probably mumble something about the two being bad influences upon each other and that if you can only keep them apart they are no trouble. It is the confluence of the two agents which is responsible for the mischief.

As long as p(event/A $\cap$ B) is the highest probability present the conditions for Type II are met. Some variations on it, beginning with that for the Melvin and Melda example, are as follows:

p(event/A $\cap$ B) > p(event) > p(event/A) > p(event/B)

p(event/A $\cap$ B) > p(event/A) > p(event) > p(event/B)

p(event/A $\cap$ B) > p(event/A) > p(event/B) > p(event)

In all of these variations the event will be attributed to the joint presence of the two agents.

An interesting characteristic of Type II patterns is that they suggest some sort of synergy between the agents. In all of these patterns the two agents together give outcomes which are quite different from (or at least more extreme than) what is normally associated with the agents taken individually. Examples from every day life include two mediocre hockey players who, when put on the same line, "click" to produce above average play; a married couple who get no enjoyment from anyone but each other, and a graduate student and a professor who argue with one another incessantly but are on very good terms with everyone else. In all of these examples there is synergy because the whole is different from the sum of the parts.

When preventors are involved in a Type II pattern the basic logic is the same but "anti-synergy" is involved. Events which are expected to occur when the two agents are individually present are less likely to occur when both are present. For example, all-star hockey teams are noted for consistently doing more poorly than is expected based upon the individual performances of the players involved. The unexpectedly poor performance of a star player on the team cannot be explained by referring to incompetence on the part of team-mates, for they are all stars as well. Most explanations of such things involve the idea that these individuals do not play well together.

Type III, Redundancy. The core pattern for this type of attribution is:

$$p(\text{event}/A) = p(\text{event}/B) \geq p(\text{event}/A \cap B) > p(\text{event})$$

This third type is characterized by a single probability pattern for which there are no variations. The probabilities for the two single agents are equal and they are the best predictors of the event. They will be named as the causes of the event by observers because they are the best predictors and because this event is predictable in these circumstances, $p(\text{event}/A \cap B) > p(\text{event})$.

An example of Type III is as follows. The event to be explained is an argument between two people. It is known that these two people, when they meet, are much more likely to argue than are other people, $p(\text{argument}/\text{Mary} \cap \text{Mike}) = .50 > p(\text{argument}) = .05$. It is also known that Mary and Mike are, individually, very argumentative people, $p(\text{argument}/\text{Mary}) = .50 =$

p(argument/Mike). When asked why Mary and Mike argued on this particular day observers may say, "They're argumentative", or, "They always do, they just have to argue no matter who they meet." Alternately observers may focus on one or the other of the agents as the cause and name that one. Relatively transitory factors such as salience and attention can affect which agent is focused upon and therefore how the causal attributions are made. Type III is called redundancy because two equally acceptable causal explanations are available and are redundant with each other. The logic for preventors in this Type is essentially the same as it is for causes, as it was in the first two types.

Type IV, Surprise. The pattern of probabilities for this type of attribution is:

$p(\text{event}) > p(\text{event}/A \cap B)$

The core statement defining Type IV attributions shows that the presence of agents A and B predicts that the event has a less than usual probability of occurrence. In other words the presence of the two agents is a sufficient preventor of the event because in the past there has been a predictable non-occurrence of the event in the presence of the two agents. But on this occasion the event has occurred, despite their presence. This surprise event is therefore not explainable in terms of the known characteristics of these agents. Given that the event is unpredicted and therefore cause cannot be attributed to the agents as they are perceived at present, observers will turn to one of the three strategies for unexpected events which were described earlier; to attribute to chance or some other ill-defined agent, to attribute to an unknown agent, or to change beliefs about the dispositions and/or constraints of known agents.

An example of a Type IV attribution can be seen by slightly modifying the second fishing creek example used above. In that example, p(good catch/Jack) = .95, p(good catch/creek) = .20, p(good catch/Jack ∩ creek) = .30, and p(good catch) = .50. These probabilities will form a Type IV pattern if it is assumed that on this occasion Jack gets a good catch (when this example was used above as a variant of Type I, it was assumed that Jack, as expected, did not get a good catch). In the Type IV version Jack, against expectations, gets a good catch. In other words, p(event) > p(event/A ∩ B). This unexpected event cannot be explained in terms of the known agents, so observers will turn to one of the three courses of action outlined above for unexpected events. They might say that Jack got lucky today, or that an unknown agent (perhaps a new lure that Jack bought yesterday) caused the good catch, or that somehow Jack or the creek changed. It might be decided that, after all, Jack is a good enough fisherman to get a good catch in this creek.

It is important to notice the difference between Type IV attributions and the preventor versions of Types I to III. In Type IV patterns an event is predicted not to occur but it does anyway. Observers therefore cannot explain the event on the basis of the information which led them to believe in the first place that the event would not occur. In the preventor versions of Types I to III an event is predicted not to occur and it does not occur. This non-occurrence can be explained in terms of the information which was used to predict it.

There are many variants upon the Type IV pattern because the core pattern is so simple. Here are some of them.

$$p(\text{event}) > p(\text{event}/A \cap B) > p(\text{event}/A) > p(\text{event}/B)$$

$$p(\text{event}) > p(\text{event}/A \cap B) > p(\text{event}/A) = p(\text{event}/B)$$

$$p(\text{event}) > p(\text{event}/A) > p(\text{event}/B) > p(\text{event}/A \cap B)$$

<u>Summary of the Core Model</u>

The essence of the core model is as follows:

1. Observers store information about events and agents as probability statements.

2. These probability statements serve as summaries of past events, predictors of future events and are used by observers to "explain" events.

3. If the presence of an agent is a predictor that a particular event is more or less likely than usual to occur, observers label the agent with a disposition if it is a human and a constraint if it is an entity or situation.

4. The better the agent is as a predictor, the stronger will be the disposition or constraint assigned to that agent.

5. Observers base their attributions of cause on the ability of their stored probability statements to predict events.

6. If an agent is a predictor of an event it is a sufficient cause of the event and if an agent is a predictor that the event is unlikely to occur it is a sufficient preventor of the event.

7. When asked about the causes of events observers will scan their stored probabilities concerning the agents present when the events occurred.

8. If an event is predictable given the presence of the agents, observers will name the agent or combination of agents with the highest probability of association with the event, as the cause of the event.

9. If an event is not predictable given the presence of the known agents, observers will:
    a. attribute to chance or other vague agents
    b. attribute to agents about which little is known
    c. revise their probability statements about the constraints and/or dispositions of agents.

10. The many possible patterns of stored probability statements can be classified into four Types. Each Type leads to a particular kind of attribution when observers use the rules just stated in 7, 8 and 9. The Types and their associated attributions are summarized in Figure 1.

Insert Figure 1 about here

Configuration Theory

The core model will now be shown to be a basis for Kelley's configuration theory. After this is accomplished the same will be done for Kelley's covariation theory so that Kelley's two theories will be integrated.

According to Kelley (1973), configuration theory applies when observers must make attributions after only a single observation of the event. But Kelley points out that in such circumstances the observers are seldom completely ignorant, they have observed similar events before and have some notions of the causal processes involved. It is by using this stored knowledge that observers are able to make attributions.

The core model assumes that this knowledge which is relevant to the interpretation of the event is stored by observers as probabilities. These probabilities are used by observers to derive attributions in the ways described above. The various concepts of configuration theory will now be restated in terms of the core model.

Plausible Cause, Facilitative Cause, Sufficient Cause

These three terms are synonymous in the core model. The way in which Kelley (1973) uses the term plausible cause suggests that a plausible cause is an agent which is perceived by observers to be capable of causing the event. Therefore, when observers are looking for a causal explanation for the event, they will find this agent a likely candidate. In the core model an agent will be considered as a cause for an event when $p(\text{event/agent}) > p(\text{event})$. This is also the definition of a sufficient cause. It essentially states the assumption of the core model that perceptions of cause are based upon predictability and an agent which is a predictor of an event can be seen as the cause of the event. Thus, to say that an agent is a plausible cause of an event is the same as saying that it is a sufficient cause. When Kelley (1973) uses the term facilitative cause it refers to an agent which acts as a force to cause the event to occur. In the core model, this is also the same as saying that the agent is a sufficient cause. Given the synonymy of these three terms, the following definition holds.

> An agent will be labelled a PLAUSIBLE CAUSE, FACILITATIVE CAUSE or a SUFFICIENT CAUSE when:
>
> p(event/agent) > p(event)

<u>Inhibitory Cause, Sufficient Preventor</u>

According to Kelley (1973) an inhibitory cause is one which acts to suppress the occurrence of the event. In the core model this means that observers believe that this particular agent tends to be associated with the less than usual occurrence of the event. This is essentially the same as the core model's definition of a sufficient preventor, so the following definitions holds.

> An agent will be labelled an INHIBITORY CAUSE or a SUFFICIENT PREVENTOR when:
>
> p(event) > p(event/agent).

Kelley (1973) uses the idea of the causal schema to represent the cognitive structures of observers which are used to organize information while making attributions. The frameworks of these schemata provide structures into which observers put isolated bits of information in order to organize them into meaningful patterns from which attributions can be made. These schemata will now be analyzed in terms of the core model.

Multiple Necessary Causes

In this schema, "Both causes must be present or favourable if an effect is to occur," (Kelley, 1973). In the core model this is an extreme case of Type II, Synergy, as can be seen when the probabilities implied by Kelley's statement of the schema are examined. First, the event will not occur if only one of the agents is present, or, $p(\text{event}/A) = 0$, $p(\text{event}/B) = 0$. Second, when both agents are present the event always occurs, $p(\text{event}/A \cap B) = 1.0$. The only thing missing is a value for p(event). It is safe to assume that the events which most observers will encounter will have a probability of occurrence less than 1.0 and more than 0, so $p(\text{event}/A \cap B) = 1.0 > p(\text{event})$. The four probability statements just listed can be easily arranged to show that they fit the Type II pattern. This yields the following definition.

Agents A and B are MULTIPLE NECESSARY CAUSES of an event when:

$$p(\text{event}/A \cap B) = 1.0 > p(\text{event}) > p(\text{event}/A) = p(\text{event}/B) = 0.0.$$

Thus, in the core model Kelley's multiple Necessary Cause schema is a particular example of the Type II pattern of probabilities. The advantage of the core model is that it not only allows a statement of multiple necessary causes which is as clear and elegant as Kelley's, it also shows that there is a family of related schemata or probability patterns, some of which may deserve investigation. Not all of the related patterns will be as easily understood as the one identified by Kelley.

Multiple Sufficient Cause

According to Kelley (1973) this schema occurs when there is more than one cause present and each is alone capable of causing the event. Translated into the core model, this yields the following definition.

Agents A and B are MULTIPLE SUFFICIENT CAUSES of an event when:

p(event/A), p(event/B) > p(event)

The multiple sufficient causes schema cannot be tied as uniquely to one of the four basic probability patterns as the multiple necessary causes schema could be. Some variants of the Type I pattern will fulfill the defining condition of the multiple sufficient causes schema which is shown above. To be a Type I pattern p(event/A) must be greater than all of the other probabilities and p(event/A $\cap$ B) > p(event). Within these parameters it is quite possible for p(event/A), p(event/B) > p(event), which is the requirement for multiple sufficient causes. But the presence of the multiple sufficient causes schema is not essential for the Type I pattern. In the Type II pattern, in which the only requirement is that p(event/A $\cap$ B) is greater than all the other probabilities, it may also happen that p(event/A), p(event/B) > p(event). But again, this is not an essential aspect of the type. The Type III pattern has only one version and an essential part of that pattern is that p(event/A) = p(event/B) and that these two probabilities be greater than all of the other probabilities, including p(event). Thus, multiple sufficient causes is an essential aspect of the Type III pattern of probabilities. The definition of Type IV patterns is that p(event) > p(event/A $\cap$ B). The required conditions for multiple sufficient causes can occur in such patterns but the schema is not an essential aspect of the pattern.

Given all of this, in the core model the multiple sufficient causes schema is most closely associated with the Type III pattern because it is the only one in which the schema plays an essential role. Although the core model can define the multiple sufficient causes schema in a very precise way it does not give it as unique a position as it does the multiple necessary causes schema.

<u>Compensatory Causes</u>

Kelley (1973) uses an example of task success to demonstrate the idea of compensatory causes. Using an actor and a task as the agents, and success on the task as the event, his model shows that characteristics of the actor and task can trade off with each other to produce the event. Figure 2 shows Kelley's (1973, p. 114) diagram representing this schema.

Insert Figure 2 about here

The event (represented by E in the Figure) is the occurence of a successful outcome when a person of either high, medium or low ability works on a task that is difficult, moderate or easy to do. When a person of medium ability tackles a moderate task, E occurs. The person is successful. When a person of medium ability takes on difficult task, the event does not occur, the actor is unsuccessful.

In the core model Kelley's diagram serves the purpose of showing the probabilities of events associated with various disposition and constraint labels. These probabilities can be used to attribute cause, dispositions and/or constraints. Bearing in mind that Kelley presented this oversimplified diagram for illustrative purposes, the core model will now be applied to it. This does not assume that the actual probabilities associated with the labels by real observers are as represented in this diagram. The actual probabilities would have to be determined empirically.

In Figure 2 the following probabilities are represented. In the column for the difficult task, only one of the three cells has an E in it. In other words, p(E/difficult task) = .33. Analogous logic applies to the other two columns, so that if abbreviations are used to represent difficult moderate and easy tasks,

$p(E/DT) = .33$ (DT = difficult task)

$p(E/MT) = .66$ (MT = moderate task)

$p(E/ET) = 1.0$ (ET = easy task)

This kind of representation allows observers to do at least two things. First, if they notice that a particular task has a certain probability of success associated with it they can label it as difficult, moderate or easy, thus applying a constraint label. Second, if someone else informs observers that a task is moderate, for example, they can easily attach a probability of success statement and use it later to make attributions.

The labels and probability statements for the dispositions of actors, as represented in the rows of Figure 2, are as follows.

$p(E/LA) = .33$ (LA = low ability)

$p(E/MA) = .66$ (MA = medium ability)

$p(E/HA) = 1.0$ (HA = high ability)

The individual cells in the Figure can also be represented as probabilities.

$p(E/DT \cap HA) = 1.0$
$p(E/DT \cap MA) = 0$
$p(E/DT \cap LA) = 0$
$p(E/MT \cap HA) = 1.0$
$p(E/MT \cap MA) = 1.0$
$p(E/MT \cap LA) = 0$
$p(E/ET \cap HA) = 1.0$
$p(E/ET \cap MA) = 1.0$
$p(E/ET \cap LA) = 1.0$

Finally, looking at all nine cells and their E's,

$p(E) = .66$

The core model assumes that all of these probabilities are stored by observers and can be used to make causal attributions. For example, if observers see a person who is labelled as having high ability succeed on a task which is labelled high difficulty, the following probabilities will be consulted.

$$p(E/HA) = 1.0 = p(E/HA \cap DT) > p(E) = .66 > p(E/DT) = .33$$

This pattern is a variant of Type I in the core model and observers will name as the causal agent the one with the highest probability, namely the actor of high ability. If a person of low ability succeeded on a task that was difficult, the appropriate probabilities would be:

$$p(E) = .66 > p(E/LA) = .33 = p(E/DT) > p(E/LA \cap DT) = 0$$

This is a variant of Type IV, surprise. Observers would say the person got lucky or perhaps decide that he or she has ability after all. Alternately some unknown agent might be invoked.

Kelley (1973) also shows that the compensatory causes schema can be used by observers to infer the characteristics of an unknown agent when the nature of the other agent and of the event are known. This process of inferring dispositions or constraints goes beyond the purely causal attributions aspects of Kelley's theories and cannot be fully discussed in terms of the core model until the link between the core model and Jones and Davis' (1965) theory is demonstrated. As stated earlier this goes beyond the confines of the present paper. However, since this inference process is an important use of this schema, the core model's explanation of it will be sketched briefly. The core model assumes that observers can compare probabilities and can use symmetry around the value of p(E) to infer the characteristics of unknown agents. Suppose an agent of unknown ability succeeds at a difficult task. Observers will first assume that this event is indicative of the enduring characteristics of both the person and the task. They will assume that in past and future encounters, if any, the outcome would be the same. In other words they assume, $p(E/Actor \cap DT) =$

1.0. Taking this, and the probabilities available from Figure 2, the observers now have the following probabilities to scan,

$$p(E/\text{actor} \cap DT) = 1.0 > p(E) = .66 > p(E/DT) = .33.$$

There is no probability here for the unknown actor alone. The process of inferring a disposition for the actor involves the process of deciding what probability of success should be assigned to that person. The observers believe that the difficult task exerts an inhibitory or preventive influence upon success and for them the magnitude of that inhibition is represented by the difference between the overall probability of success, $p(E) = .66$, and the probability of success given the difficult task, $p(E/DT) = .33$. The inhibiting effect is thus represented by the value -.33. In order for the event to have occurred, the inhibitive effect of the task must have been overcome by the actor. The facilitative effect of the actor must, therefore, be equal to or greater than the inhibition and must act in the opposite direction. It must be a facilitative effect of +.33. Given this, and the knowledge that $p(E) = .66$, the probability associated with the actor must be $.66 + .33 = 1.0$. The actor's ability disposition is therefore represented by the probability statement, $p(E/\text{Actor}) = 1.0$. Only one class of actors has this probability statement attached, actors labelled high ability. Observers will therefore infer that the unknown actor has high ability and was thus able to succeed at the difficult task. Notice that this inference process involves two attributions. The first is that the actor caused the success, the second is that the actor has the disposition, high ability. These two attributions are an integral part of the observers' explanatory processes.

It is important to notice how the core model permits this facile operation of both dispositional and causal of attributions, simultaneously.

By representing the concepts and processes in terms of probabilities, the whole model is seen to operate in an elegant and fluid manner.

The core model can also be used to show that the compensatory causes schema is not really an entirely new schema. It is really a representation of data which contains the simpler schemata of configuration theory. Depending on the circumstances one or the other of these simpler schemata will be appropriate. For instance, observers who see a successful event when a person of medium ability encounters a task of moderate difficulty, can draw from Figure 2 the following probabilities, $p(E/MT \cap MA) = 1.0 > p(\text{event}) = p(E/MT) = p(E/MA) = .66$. This is a variant of a Type II pattern which is associated with the multiple necessary causes schema. Likewise, observers who see an actor of high ability having success on an easy task can draw upon the following probabilities, $p(E/ET) = p(E/HA) = p(E/ET \cap HA) = 1.0 > p(\text{event}) = .66$. This is a variant of Type III, which is closely associated with the multiple sufficient causes schema.

This discussion of the compensatory causes schema has shown that its workings can be entirely described in terms of the core model. The making of causal and dispositional attributions were described and the intimate cooperation of these processes, when viewed in terms of the core model, pointed out. The core model also was used to show that the compensatory causes schema is a way of representing, simultaneously, other simpler schemata.

There are two other important concepts in Kelley's configuration model, discounting and augmentation. These are not schemata, but rather phenomena, which occur because of the way in which configuration works.

<u>Discounting</u>

Kelley (1973) uses the term discounting to refer to the fact that the more plausible causes there are available for an event, the weaker will be the attribution to any one of those causes. In Kelley's example (Kelley, 1973, p. 113), which is based upon a study by Thibaut and Riecken (1955), the observer has a choice of attributing to a particular actor or to a particular entity. If the actor is the only plausible cause, the attributions to the actor are higher than if both the actor and the entity are plausible causes. In Kelley's terms, the presence of the plausible entity causes a discounting of the attribution to the actor. Although in Kelley's example there is only a single alternate plausible cause, discounting need not be restricted to such cases. There may be several alternate plausible causes. For example, if an airplane which is manned by a pilot who is known to be incompetent crashes despite the fact that the plane was in good repair, the weather was very good and the control tower was operating well that day, most observers would attribute the crash to the only plausible cause available, the incompetent pilot. However, if the pilot was incompetent, the airplane was in poor repair, the weather was stormy and the control tower was manned by inexperienced people, there are several plausible causes available. The discounting principle predicts that attribution to the pilot will be less under the multiple plausible causes than under the single plausible cause circumstances. A question remains as to whether there will be more discounting the more alternate plausible causes there are. This could be checked empirically. Kelley's example of discounting involves an actor as a focal agent and an entity as the alternate plausible agent. But the focal agent could be an entity and the alternate agent a person. For example, when a faulty airplane crashes the crash will be more strongly attributed to the airplane when the pilot is competent than when the pilot is incompetent.

In the core model, discounting involves a comparison of attribution Types I and III. This will be demonstrated by contrasting the two examples of the fishing creek which demonstrate Types I and III respectively. Suppose the event to be explained is Jack's good catch in the creek this morning and the following probabilities had been previously stored by observers.

p(good catch) = .50

p(good catch/creek) = .20

p(good catch/Jack) = .95

p(good catch/creek $\cap$ Jack) = .80

These are the same probabilities as were used above to demonstrate the Type I pattern and they therefore fulfill the basic conditions of Type I.

1. p(event/Jack) = .95 > p(event/Jack $\cap$ creek) = .80 > p(event) = .50

2. p(event/Jack) = .95 > p(event/creek) = .20

Observers will therefore unequivocally attribute the good catch to Jack. Discounting is demonstrated when this situation is compared to a similar one in which there are two equally plausible causes present, rather than just one. In the fishing example this comes about when the creek, rather than being an inhibitor of good catches, is a facilitative cause equal to Jack. This is represented in probabilistic terms by assuming that p(good catch/creek) = .95, rather than .20, as was assumed for the Type I example. This change gives a set of probabilities which fit Type III.

p(event/Jack) = p(event/creek) = .95 > p(event/Jack $\cap$ creek) = .80 > p(event) = .50

These probabilities show two equally plausible causes and observers must choose one, the other, or both, when stating their causal attributions. Some observers will say it is the creek, others that it is Jack, and others that it is both the creek and Jack. The average attribution to Jack in these circumstances will be less than it is with the Type I pattern because in Type I all of the observers will be attributing only to Jack. Stated in another way, discounting occurs when a Type I situation is converted into a Type III situation by adding plausible causes.

> DISCOUNTING refers to the fact that when an agent is the single most plausible cause of an event, attributions to that agent will be greater than when that agent is only one of two or more equally most plausible causes.

Augmentation

Kelley (1973) presents the augmentation principle as the opposite of the discounting principle. In discounting, an alternate agent which facilitates the occurrence of the event has the effect of reducing attributions to the focal agent. In augmentation, an alternate agent which inhibits the occurrence of the event has the effect of increasing attributions to the focal agent. In Kelley's example, an actor who succeeds at a very difficult task (the difficult task is a strong inhibitor of the occurrence of the event, success) receives stronger ability attributions than an actor who succeeds at a very easy task (the easy task is not an inhibitor of the event, success).

A careful analysis of augmentation using the core model shows that it can involve a change in the locus of a causal attribution and/or a change in the nature of a disposition or constraint attributed to an agent. In order to deal with augmentation fully, then, another trespass into the realm of disposition and constraint attribution is necessary. In order to keep to the prime theme of this paper, causal attributions, this trespass will be made as brief as possible.

Kelley quite rightly pointed out that augmentation is based upon the compensatory causes schema, so the basic nature of augmentation can be demonstrated using Kelley's example of augmentation and the earlier discussion of compensatory causes. Kelley's example states that an actor who succeeds at a difficult task receives stronger ability attributions than an actor who succeeds at an easy task. The case of the actor of unknown ability succeeding at a difficult task was analyzed in the discussion of compensatory causes above. In it, by assuming symmetry of potency of agents around p(event), it was shown that in order to balance the -.33 associated with the difficult task, the observers had to attribute at least a +.33 potency to the actor. Given this, the p(event/actor) was at least 1.0. Therefore the label, high ability, was applied to the actor. The second part of Kelley's example can also be analyzed in this way. In it an unknown actor has success at an easy task. Using the probabilities from Figure 1 gives

$$p(\text{event/easy task}) = 1.0 > p(\text{event}) = .66$$

Here, easy task is a sufficient cause of the event, by the definition of sufficient cause given earlier. Observers therefore have an event, an agent which is unknown (the actor) and an agent which is a sufficient cause of the event (the easy task). Therefore they will choose the easy task as the cause of the success. This shows the first aspect of augmentation. In the case in which the task was difficult, the task was a sufficient preventor (inhibitory cause) of the event so observers had to turn to the unknown actor to explain the event. In order to say that the actor was the cause of the event they had to assume that the actor had high ability. In ~~their~~ this case the available data demanded an interpretation which named the actor as the cause and the attribution of high ability to the actor. In the case of the easy task, however, the causal locus was most readily perceived to be the easy task, and no particular assumptions about dispositions of the actor were necessary. Observers could assume the actor to have low, medium or high ability, and it would not influence their causal attributions. Given this pressure to name the entity as the cause, and the freedom to choose one of three dispositions for the actor, it follows that the case of easy task will have weaker attributions to the actor, both causal and dispositional. Another way of stating it is to say that when an event occurs in the presence of a known inhibitory agent and an unknown agent, observers must attribute the cause of the event to be the unknown agent if they are to avoid contradicting the knowledge they have about the inhibitory agent. Further, they must attribute a disposition or a constraint to the unknown agent which is of sufficient potency to overcome the inhibitory agent. Therefore the more inhibitory the one agent the more facilitory the other agent must be assumed to be. This is the crux of augmentation.

> AUGMENTATION refers to the fact that when an inhibitory agent and an unknown agent are present when an event occurs, the unknown agent will be named as the cause of the event and a facilitory disposition or constraint will be attributed to the unknown agent. The stronger the inhibitory influence of the inhibitory agent the stronger will be the facilitory disposition or constraint attributed to the unknown agent.

The core model has the potential for allowing a more rigorous analysis of the relationship between augmentation and discounting than has so far been presented by anyone. The discussion here has shown that these two concepts, as described by Kelley, can be analyzed in terms of the core model. But several things coming out of that analysis suggest that the relationship between the two may not be as simple as is suggested by the statement that they are more or less opposites. For example, in discounting all of the agents involved are known, they have probability statements attached to them by observers and therefore have established dispositions and/or constraints. In augmentation one of the agents is unknown and is therefore likely to have disposition or constraint labels (and therefore probabilities) attached during the process of attribution. Also, in discounting only the assignment of a causal attribution is involved, while in augmentation both causal and disposition/constraint attributions are assigned. These differences suggest that more theoretical analysis here will bear fruit. However, such an analysis is beyond the scope of this paper.

This completes the discussion of Kelley's configuration theory. It has shown that the concepts of that theory can be reduced to the probability statements of the core model. In addition, the processes involved in attribution making can also be described by the core model. The multiple necessary causes scheme was shown to be an example of the Type II pattern of probabilities and the multiple sufficient causes schema was shown to be most closely related to Type III. The compensatory causes schema was shown to be made up of simpler schemata and its operation for attributing causes and dispositions/constraints described. Discounting was shown to result from a comparison of Types I and III. Augmentation was shown to involve both causal and dispositional/constraint attributions. Some differences between discounting and augmentation were made apparent by the use of the core model. This statement of Kelley's configuration theory using the core model showed a clear and simple basis for it. In addition, several insights into the operation of configuration were provided. The most important of these is that in the core model, causal and disposition/constraint attribution processes operate in close cooperation with one another in a fluid and elegant way.

Covariation Theory

Now Kelley's covariation theory will be analyzed in terms of the core model and this analysis will show that covariation theory and configuration theory are very closely related. Although it is well known that they are related and several writers (eg. Kelley, 1973) have mentioned close links, there has so far been no systematic demonstration of this. It will be shown here that the attributions described in the two models are based upon information that is stored by observers in the same form (probabilities), that the process by which attributions are formed from this information are the same (comparison of probabilities), that identical terminology can be applied to the two models and that about the only difference is in the source of the information used to form attributions. This difference was originally pointed out by Kelley. In covariation theory this information comes from the observation of the event repeatedly, while in configuration theory it comes from virtually any source, including past experience with similar agents and events and being told by other people. In the core model this difference is a relatively trivial one.

Covariation theory states that people base their attributions upon three different kinds of information, which Kelley (1973) calls consensus, distinctiveness and consistency. Kelley suggests that each of these kinds of information can occur in one of two states, high or low. With three kinds of information, each of which can take on two different values, there are eight possible combinations of information. McArthur (1972) has derived predictions about the attributions that should occur in each of these eight information configurations and has demonstrated quite good empirical support for them.

The correspondence between the core model and each of the three kinds of covariation information will now be shown. After that predictions about the attributions which should occur with each of the eight combinations of the information will be made.

Covariation theory reduces the attribution situation to one in which a single actor and a single entity interact and an event occurs. The question the observer must answer is whether the event was caused by the actor, the entity or by the circumstances of the interaction. An attribution to circumstances occurs when neither the actor nor the entity provides a plausible explanation for the event.

Consensus

Consensus information states whether the event in question also occurs when other actors interact with the focal entity. If other actors get the same event, consensus is high; if other actors do not get the same event, consensus is low. According to McArthur (1972) high consensus tends to produce attribution to the entity while low consensus tends to produce attribution to the person.

The core model assumes that in the case of high consensus because the event is associated with this entity across a variety of actors, then this entity is a good predictor of the event. No matter what actor the entity is paired with, it can be predicted that the event will occur. In other words, the presence of the entity is an indication that the event has an above usual probability of occurring. This leads to the following definition.

The term HIGH CONSENSUS applies to situations in which:

p(event/entity) > p(event)

Notice that, by this definition, if high consensus is present the entity is a sufficient cause of the event, given the probabilistic definition of sufficient cause as stated in the core model. In configuration theory high consensus therefore means the entity is a facilitative cause or a plausible cause.

According to Kelley low consensus is present when the event tends not to occur when this entity is paired with other actors. In the core model this would mean that the entity is a good predictor that the event will not occur. This event is unusual when this entity is present. This gives the following definition.

The term LOW CONSENSUS applies to situations in which:

p(event) > p(event/entity)

By this definition, when low consensus is present the entity is a sufficient preventor of the event. In configuration terms it is an inhibitory cause. Low consensus therefore means that there was an inhibitory cause present when the event occurred. The presence of an inhibitory cause will cause augmentation of attributions to any other agents which are present and plausible causes. In some cases that plausible agent is the actor. In those cases low consensus will therefore support attribution to the actor, as McArthur (1972) suggested.

In summary, with high consensus the entity is highly associated with the event, is a sufficient and plausible cause of the event, and so attributions to the entity tend to be high when consensus is high. With low consensus the entity has a low association with the event, is neither a plausible nor sufficient cause of the event, is a constraint to prevent the event occurring (is an inhibitory cause) and so attributions to the entity tend to be low when consensus is low. Low consensus tends to augment attributions to the actor.

Distinctiveness

Distinctiveness works on the same principles as consensus but for distinctiveness the focal agent is the actor rather than the entity. According to Kelley (1973) distinctiveness refers to whether or not the event occurs when the same actor interacts with other entities. If the same event tends to occur when the actor is with other entities, distinctiveness is low. If the event does not occur with other entities, distinctiveness is high.

According to the core model distinctiveness has to do with the probability that the event will occur in the presence of the actor. High distinctiveness means that there is a low association between the actor and the event.

The term HIGH DISTINCTIVENESS applies to situations in which:

p(event) > p(event/actor)

Low distinctiveness occurs when the event tends to be associated with the focal actor, regardless of what entities are present.

The term LOW DISTINCTIVENESS applies to situations in which:

p(event/actor) > p(event)

These probability statements can be related to the earlier discussion of dispositions. High distinctiveness means there is little association between the actor and the event. In this situation the actor is not a plausible cause of the event. If the probability of the behaviour is very low it suggests that the individual has a disposition to NOT emit the focal behaviour. For example, an individual might be known for never smiling. With such a strong disposition the augmentation principle can play a role, just as it did in the case of low consensus. It operates upon the same principle as Kelley originally stated for the augmentation principle but here the inhibitory cause is internal (disposition of the person) and the enhanced attributions are to the entity. For example, take a person who is known to rarely laugh, even when others do, $p(laugh/actor) < p(laugh)$. If that person goes to a movie and is seen to laugh, observers are likely to say the movie is very funny. They are likely to say it is funnier than they would if they observed a person laughing at the movie who has a "normal" history of laughter. As McArthur (1972) suggested, high distinctiveness therefore enhances attribution to the entity. In contrast, low distinctiveness means that the actor is likely to emit the behaviour in a variety of circumstances. The actor is therefore a plausible and sufficient cause of the event. The actor will be perceived to have a disposition to act in the way demonstrated by the focal event and will be perceived to be the cause of the event.

In summary, with high distinctiveness the actor has a low association with the event, is neither a plausible nor a sufficient cause of the event, has a high disposition to hinder the occurrence of the event (provides an inhibitory cause) and so attributions to the actor tend to be low with high distinctiveness. Further, high distinctiveness may augment attribution to the entity. With low distinctiveness the actor is highly associated with the event, is a sufficient and plausible cause of the event, has a strong disposition to produce the behaviour which is the event and so attributions to the actor are high with low distinctiveness.

Consistency

Consistency information states whether the event has occurred when this particular actor and this particular entity have interacted in the past

(Kelley, 1973). High consistency means they usually get this event, low consistency means that they have not got this event in the past. According to McArthur (1972), high consistency enhances attributions to the actor and/or entity.

According to the core model, consistency states the probability of the event given BOTH the entity AND the actor. High consistency means there is a high association between the event and the confluence of actor and entity.

The term HIGH CONSISTENCY applies to situations in which:

$p(\text{event}/\text{actor} \cap \text{entity}) > p(\text{event})$

Low consistency occurs when the event has low association with the confluence of the actor and entity.

The term LOW CONSISTENCY applies to situations in which:

$p(\text{event}) > p(\text{event}/\text{actor} \cap \text{entity})$

The probability statements for consistency cannot be as closely tied to the concepts of constraint and disposition as were the statements about consensus and distinctiveness. If consistency is high, it means that this event usually occurs when these two agents interact; the event might occur because either the actor or the entity alone is a sufficient cause of the event, or because the actor and the entity are multiple necessary causes of the event. The consistency information alone does not tell observers which is the case, so they must consult consensus and distinctiveness information to try to resolve the question. If consistency is low it means that this event does not usually occur when these agents interact. In this case observers will turn to one of the three courses of action available for such circumstances.

In configuration theory there is no concept analogous to consistency. It has just been shown that configuration theory does have concepts which explicitly correspond to the various states of consensus and distinctiveness information. But with consistency, although assumptions about it are made in some parts of configuration theory, and the core model can show where those parts are and what the assumptions are, configuration theory does not explicitly give it a place. In the discussion that follows this will be seen to be a weakness of configuration theory which makes it unable to predict some things which covariation theory and the core model can predict.

Covariation Cases

Now that the meaning of each of the three kinds of covariation information has been stated in terms of the core model, the eight combinations of information suggested by Kelley (1973) can be analyzed. The eight combinations are listed in Figure 3. Each combination is called a

----
Insert Figure 3 about here
----

case and will be treated in turn.

Case One. The states of the three kinds of covariation information and their respective probability statements, for Case One, are as follows.

low consensus	p(event) > p(event/entity)
low distinctiveness	p(event/actor) > p(event)
high consistency	p(event/actor $\cap$ entity) > p(event)

These three statements capture the nature of Case One as Kelley defined it. If one more assumption is made it can be shown that the case fits the Type I pattern of the core model. That assumption is that p(event/actor) $\geq$ p(event/actor $\cap$ entity). This assumption will be discussed more in a moment, but for now we will just proceed with it. Given this additional assumption, the probability statements for Case One can be arranged to show the Type I pattern.

1. $p(\text{event/actor}) \geq p(\text{event/actor} \cap \text{entity}) > p(\text{event})$

2. $p(\text{event/actor}) > p(\text{event/entity})$

The core model predicts that this pattern will yield attribution to the actor as the cause of the event. McArthur (1972) and Kelley (1973) both predicted this also, and McArthur's empirical research confirmed it. Thus the core model agrees with Kelley's statement of covariation theory.

Configuration theory can also be used to predict in this case. As was shown earlier, low consensus means that the entity is an inhibitory cause. It is therefore not a plausible cause of the event, is unlikely to be named as the cause, and may even cause augmentation of causal attributions to another agent. Low distinctiveness means that the actor is a plausible, sufficient and facilitory cause. Since it is the only plausible cause available, configuration theory, like the core model and covariation theory, predicts that attribution of cause will be to the actor in Case One. Notice that in this case configuration theory makes this correct prediction without reference to the concept of consistency, although both the core model and covariation theory use the concept to make the same prediction. At this point the lack of consideration of consistency does not harm configuration theory. Below it will be seen that it does not always fare so well.

It was stated above that the core model, in order to deal with Case One, required an assumption that Kelley and McArthur did not mention. The core model said that in order to make the same prediction as Kelley and McArthur, a further assumption, not mentioned in Kelley's basic statement of Case One in terms of consensus, distinctiveness and consistency, is necessary. That assumption was that $p(\text{event/actor}) \geq p(\text{event/actor} \cap \text{entity})$. The reason for this assumption can be seen in Table 1. If this assumption is violated it means that $p(\text{event/actor} \cap \text{entity}) >$

p(event/actor). If this is so then in the set of probabilities for Case One the largest probability is p(event/actor $\cap$ entity). If this happens this pattern fits Type II and not Type I. Given this the core model would predict joint attribution to both the actor and the entity, rather than to the actor alone, as is predicted for Type I.

The discussion of this extra assumption touches upon one advantage of the core model over Kelley's statement of theory. By reducing all the concepts to probability statements the core model makes it possible to meaningfully compare consensus, distinctiveness and consistency information with each other. Using the verbal form articulated by Kelley makes such comparisons virtually meaningless. For example, it would be meaningless to ask how low distinctiveness has to be before it exceeds high consistency. But in the probabilistic terms of the core model this is a perfectly legitimate and clear question.

This discussion of Case One has shown how the core model can be used to disentangle some of the fine points in Kelley's theory and show its basic, elegant core logic. It was found that the configuration concepts inhibitory cause and facilitory cause could be applied to the consensus and distinctiveness concepts of covariation theory. This application showed that configuration theory and covariation theory made the same predictions for Case One. This discussion, though, showed that configuration theory does not really provide a place for consistency information, which is a part of covariation theory. It was also shown that Kelley's covariation theory does not make explicit a particular assumption which the core model shows is necessary if the covariation theory prediction for Case One is to be valid. Thus, the reduction of both covariation theory and configuration theory to the core model has shown how intimately connected they are and what some of their strengths and weaknesses are.

Case Two. The states of the three kinds of covariation information and their respective probability statements are as follows.

high consensus	p(event/entity) > p(event)
high distinctiveness	p(event) > p(event/actor)
high consistency	p(event/actor $\cap$ entity) > p(event)

If it is further assumed that p(event/entity) $\geq$ p(event/actor $\cap$ entity) analogous to the extra assumption for Case One the pattern of probabilities can be seen to fit Type I.

1. p(event/entity) > p(event/actor $\cap$ entity) > p(event)
2. p(event/entity) > p(event/actor)

The core model therefore predicts that Case Two will give attribution to the entity as did Kelley and McArthur, and as confirmed empirically by McArthur (1972).

In Case Two the high consensus means that the entity is a plausible cause. The high distinctiveness means that the actor is not a plausible cause but an inhibitory cause. Configuration theory therefore also predicts attribution to the entity in Case Two.

The core model has shown that Cases One and Two are both examples of Type I. The only difference is that in Case One the actor is the most plausible cause while in Case Two it is the entity which is most plausible.

Case Three. The states of the three kinds of covariation information and their respective probability statements are as follows.

low consensus	p(event) > p(event/entity)
high distinctiveness	p(event) > p(event/actor)
high consistency	p(event/actor $\cap$ entity) > p(event)

These probability statements can be rearranged to show that they fit the basic Type II pattern.

$$p(\text{event}/\text{actor} \wedge \text{entity}) > p(\text{event}/\text{actor}),\ p(\text{event}/\text{entity}),\ p(\text{event})$$

The core model therefore predicts that causal attributions in Case Three will be to both the actor and the entity. McArthur (1972) made the same prediction and found empirical support for it.

For this case configuration theory does not work nearly as well as do covariation theory and the core model. This becomes obvious if the terminology of plausibility of causes is used. Low consensus means that the entity is not a plausible cause of the event and high distinctiveness means that the actor is also not a plausible cause. Given this information there is no causal explanation for the event. This is as far as one can go using the terminology of plausible causes from configuration theory. This dead end shows the weakness of configuration theory which arises from its lack of explicit consideration of consistency information. It is by considering consistency information that the core model and covariation theory are able to make their correct predictions. The core model does show that buried in configuration theory there is a possibility of dealing with Case Three, but the method is indirect and not all of the necessary assumptions are made explicit. It was earlier shown that the multiple necessary causes schema is a variant of the Type II pattern. Case Three is also a variant of Type II. From this basis it can be shown that Case Three can fit into a multiple necessary causes schema. When discussing this schema Kelley (1973) implies, but does not explicitly demonstrate, that with this schema attributions are to both agents. Thus, by a roundabout route, configuration theory can deal with Case Three, but the route is far from elegant and has more than one important unstated assumption. This analysis has shown a connection between a schema of configuration theory and a case from covariation theory, as well as an important advantage of the core model and covariation theory over configuration theory.

Case Four. The states of the three kinds of covariation information and their respective probability statements are as follows.

high consensus	$p(event/entity) > p(event)$
low distinctiveness	$p(event/actor) > p(event)$
high consistency	$p(event/actor \cap entity) > p(event)$

McArthur (1972) predicted that attributions in this case could be to either agent alone or to both agents. Empirical data supported this prediction for the most part.

However, as with Case One, the core model shows that McArthur's predictions are not valid unless certain additional assumptions are made. If Case Four can be shown to be on example of Type III, the core model will agree with McArthur's prediction. But two additional assumptions are necessary if that is to happen. Covariation theory says that high consensus, low distinctiveness and high consistency are present. If it is assumed on the basis of this that $p(event/entity), p(event/actor) \geq p(event/actor \cap entity)$, then McArthur's prediction is valid. But the state of the covariation information as given in the basic statement of this case is also consistent with;

$$p(event/entity), p(event/actor) < p(event/actor \cap entity)$$

If this is what occurs, then this is a Type II pattern and only a prediction of joint causality is possible in the core model. This prediction of only joint causality is not the same as McArthur's prediction for Case Four. So, for McArthur's prediction to be valid, not only must the covariation information be as stated, it must also be assumed that, p(event/entity),

$p(\text{event/actor}) \geq p(\text{event/actor} \cap \text{entity})$. Further, Case Four will be an example of Type III only if, $p(\text{event/actor}) = p(\text{event/entity})$. If the two agents do not have equal probabilities, then one agent is a more plausible cause than the other and this will be a Type I pattern. As a result attribution will be to only the agent which is the more plausible cause. In order to be consistent with McArthur's predictions it must be assumed that high consensus and low distinctiveness mean that the actor and entity have equal probabilities associated with them. If these two additional assumptions are allowed; that $p(\text{event/entity})$, $p(\text{event/actor}) \geq p(\text{event/actor} \cap \text{entity})$ and $p(\text{event/entity}) = p(\text{event/actor})$; then these, combined with the probabilities from the basic statement of Case Four, can be arranged to yield a Type III pattern; $p(\text{event/actor}) = p(\text{event/entity}) \geq p(\text{event/actor} \cap \text{entity}) > p(\text{event})$. From this the core model predicts attribution to both or either of the agents, to agree with McArthur's prediction.

McArthur's (1972) data do not fully support her predictions for Case Four. This might be explained if it were shown that in designing her experimental materials for Case Four she did not take into account the two additional assumptions which the core model says are needed.

For Case Four configuration theory provides quite good predictions. The high consensus and low distinctiveness mean that both the actor and the entity are plausible causes of the event. Configuration theory would predict that whichever of the two is more plausible will be chosen as the cause. If they are equally plausible both could presumably be named. For this case configuration theory predicts quite well, despite its lack of consideration of consistency information.

Now that Case Four has been described the configuration theory concept of discounting can be analyzed in terms of covariation theory. This correspondence has not been described before and arises from the earlier discussion of discounting in terms of the core model. Earlier it was shown that discounting, in the core model, is demonstrated by comparing causal attributions to an agent in Types I and III. Now it has been shown that Cases One and Two are examples of Type I and that Case Four is an example of Type III. It follows that in covariation theory discounting can be demonstrated by comparing attributions in Case Four to those in Cases One and Two. In Case One (Type I) only the actor is a plausible cause of the event while in Case Four (Type III) both the actor and the entity are plausible causes. It is therefore expected that in Case One attributions to the actor will be greater than they are in Case Four. Attributions to the actor are thus discounted in Case Four because of the presence of the additional facilitative cause, the entity. In Case Two only the entity is a plausible cause so attributions to the entity will be higher there than in Case Four in which attributions to the entity are discounted due to the presence of the alternate plausible cause, the actor.

<u>Case Five</u>. The states of three kinds of covariation information and their respective probability statements are as follows.

low consensus	$p(\text{event}) > p(\text{event/entity})$
low distinctiveness	$p(\text{event/actor}) > p(\text{event})$
low consistency	$p(\text{event}) > p(\text{event/actor} \cap \text{entity})$

These probabilities fit the core pattern required for Type IV, surprise.

$$p(\text{event}) > p(\text{event/actor} \cap \text{entity})$$

The core model therefore predicts that this will lead to attributions of luck, attribution to unknown agents or a change in beliefs about the dispositions and/or constraints of the entities. McArthur (1972) found that her subjects attributed to "circumstances" in this Case. The response format she used allowed observers only certain attribution possibilities. They could attribute to the actor or entity alone, to some combination of the two, or to "circumstances". Thus circumstances was the alternative to use when the known agents did not provide a reasonable explanation. Thus circumstances probably corresponds to the core model's attribution to luck or vague agents. McArthur's response format did not allow observers to atttribute to unknown agents or to change the dispositions and or constraints of actors or entities. Given these restrictions, McArthur's data support the predictions of the core model.

The discussion of this case in terms of configuration theory will show again the limitation imposed upon configuration theory by its lack of consideration of consistency. Low consensus means that the entity is not a plausible cause of the event. Low distinctiveness means that the actor is a plausible cause. Given this, configuration theory would predict attribution to the actor in this Case. But McArthur's data do not support this. The reason why becomes clear when consistency is considered. Although the actor is generally a good predictor of the event; as shown by low, distinctiveness, $p(\text{event}/\text{actor}) > p(\text{event})$; when this entity is with the actor this usual state of affairs does not hold. When this entity and actor are together the event is unlikely to occur; as shown by low consistency, $p(\text{event}) > p(\text{event}/\text{actor} \wedge \text{entity})$. So the presence of the actor is not a good predictor of what will happen when this actor and entity are together. However, the presence of the entity is a good predictor under such circumstances because the presence of the entity is a predictor that the event will not occur; as is shown by low consensus, $p(\text{event}) >$

p(event/entity). This suggests that the constraint of the entity to prevent the event is normally more potent than the disposition of the actor to cause the event. This is the enduring, predictable nature of these two agents and their interaction. Now, on this occasion, this predictable pattern is violated, the event occurs. This cannot be explained on the basis of the known characteristics of the agents, so observers will turn to one of the three strategies outlined above for Type IV patterns, according to the core model. Because configurations theory takes account of only the usual characteristics of the actor and entity, and not what has happened when these two have come together in the past, it is incapable of providing a correct prediction for this case. The usual state of affairs when these two agents come together is described in consistency information.

Cases six to eight will now be discussed as a group. They all are Type IV patterns with considerable complications if discussed in detail using the core model. However, these detailed discussions will not be pursued since they all lead to essentially the same conclusions as were arrived at for Case Five. The event is unpredicted and so the causal attributions will not be to enduring characteristics of the agents involved.

Case Six. The states of the three kinds of covariation information and their respective probability statements are as follows.

high consensus	p(event/entity) > p(event)
high distinctiveness	p(event) > p(event/actor)
low consistency	p(event) > p(event/actor ∩ entity)

Because of low consistency this Case fulfills the requirements of Type IV, surprise.

p(event) > p(event/actor ∩ entity)

Case Seven. The states of the three kinds of covariation information and their respective probability statements are as follows.

low consensus p(event) > p(event/entity)

high distinctiveness p(event) > p(event/actor)

low consistency p(event) > p(event/actor $\cap$ entity)

Because of low consistency this case also fulfills the requirements of Type IV.

p(event) > p(event/actor $\cap$ entity)

Case Eight. The states of the three kinds of covariation information and their respective probability statements are as follows.

high consensus p(event/entity) > p(event)

low distinctiveness p(event/actor) > p(event)

low consistency p(event) > p(event/actor $\cap$ entity)

Because of low consistency this case also fulfills the requirements of Type IV.

p(event) > p(event/actor $\cap$ entity)

All three of these cases, six, seven and eight, because they are Type IV, are predicted by the core model to lead to attributions of luck/circumstances (if observers are not allowed to attribute beyond the named agents and cannot change the nature of the agents). McArthur (1972) agreed with this prediction and confirmed it empirically.

Configuration theory is inadequate to deal with Cases Six to Eight for the same reasons as it was unable to deal with Case Five, configuration does not take into account consistency information. The detailed working through of this will not be attempted here.

This completes the discussion of covariation theory and the core model. It has shown that the core model easily covers covariation theory, that many of the concepts of covariation theory and configuration theory are interchangeable, that the core model clarifies some parts of covariation theory and that configuration theory's lack of inclusion of consistency information is a serious weakness.

Conclusions

This paper has described a core model of attribution processes, stated in terms of probabilities, which provides a single basis for causal and dispositional/constraint attributions. This model explains how observers label actors and entities and specifies four basic types of causal attribution, based upon patterns of probabilities. It was further shown that the vocabulary of configuration theory is easily applied to the probabilities of the core model so that this core model can be seen to underlie Kelley's configuration theory. The analysis of configuration theory in terms of the core model showed some of its weaknesses but also showed an elegant intimacy of operation of causal attribution and dispositional/constraint attribution. The core model was also shown to be compatible with the logic of covariation theory. This analysis showed that in many ways the core model, configuration theory and covariation theory are identical in their logic. Although the core model underlies them all, configuration and covariation theory represent important elaborations upon the core. Certain strengths, weaknesses and assumptions of the theories were thereby demonstrated. It should now be clear that these theories need no longer to be treated separately.

This paper represents one small but important step in the clarification of an elegant core theory for attribution research. Now that Kelley's theory has been integrated it is possible to integrate Kelley's theory with that of Jones and Davis. That this is so was shown by the discussion of dispositions in covariation theory. The integration with Jones and Davis is beyond the scope of this paper but an important topic for a subsequent one.

References

Ajzen, I., and Fishbein, M. (1975) A Bayesian analysis of attribution processes. Psychological Bulletin, 82 (2), 261-277.

Buss, D.M., and Craik, K.H. (1983) The act frequency approach to personality. Psychological Review, 90 (2), 105-126.

Harvey, J.H., and Weary, G. (1984) Current issues in attribution theory and research. Annual Review of Psychology, 35, 427-459.

Heider, F. (1958) The Psychology of Interpersonal Relations. New York: Wiley.

Jones, E.E., and Davis, K.E. (1965) From acts to dispositions. The attribution process in person perception. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology, (Vol. 2). New York: Academic Press, 219-266.

Jones, E.E., and McGillis, D. (1976) Correspondent inferences and the attribution cube: a comparative reappraisal. In J.H. Harvey, W. Ickes and R.F. Kidd, (Eds.), New Directions in Attribution Research, Vol. 1. Hillsdale, N.J.: Erlbaum, 389-420.

Kelley, H.H. (1973) The process of causal attribution. American Psychologist, 28, 107-128.

Kelley, H.H. and Michela, J.L. (1980) Attribution theory and research. Annual Review of Psychology, 31, 457-501.

Kruglanski, A.W. (1980) Lay epistemo-logic process and contents. Psychological Review, 87, 70-87.

McArthur, L.A. (1972) The how and what of why: Some determinants and consequences of causal attributions. Journal of Personality and Social Psychology, 22 (2), 171-193.

Read, S.J., and Stephan, W.G. (1979) An integration of Kelley's attribution cube and Weiner's achievement attribution model. Personality and Social Psychology Bulletin, 5, (2), 196-200.

Thibaut, J.W. and Riecken, H.W. (1955) Some determinants and consequences of the perception of social causality. Journal of Personality, 24, 113-133.

Weiner, B. (1972) Theories of Motivation. Chicago: Rand-McNally.

Figure Captions

Figure 1. The four types of probability patterns and their associated attributions

Figure 2. The compensatory causes schema (based upon Kelley, 1973, p. 114).

Figure 3. Eight combinations of covariation information (Consen. = Consensus, Distinct. = Distinctiveness, Consis. = Consistency).

Name	Core Probability Pattern	Attribution
Type I Single Cause	1. $p(\text{event}/A) \geq p(\text{event}/A \cap B) > p(\text{event}/B)$ 2. $p(\text{event}/A) > p(\text{event}/B)$	Agent A
Type II Synergy	$p(\text{event}/A \cap B) > p(\text{event}/A)$, $p(\text{event}/B)$, $p(\text{event})$	Both Agent A and B
Type III Redundancy	$p(\text{event}/A) = p(\text{event}/B) \geq p(\text{event}/A \cap B) > p(\text{event})$	Agent A and/or B
Type IV Surprise	$p(\text{event}) > p(\text{event}/A \cap B)$	1. vague agent such as luck 2. unknown agent 3. revise probabilities

ABILITY	DIFFICULT	MODERATE	EASY
HIGH	E	E	E
MEDIUM		E	E
LOW			E

TASK

Case	Covariation Information			Core Model's Predicted Attribution	McArthur's Findings
	Consen	Distinct.	Consis.		
1	low	low	high	actor	actor
2	high	high	high	entity	entity
3	low	high	high	actor & entity	actor & entity
4	high	low	high	actor, entity, actor & entity	actor, actor & entity
5	low	low	low	–	
6	high	high	low	1. vague agent 2. unknown agent 3. revise probabilities	circumstances
7	low	high	low		
8	high	low	low		

Faculty of Business

McMaster University

WORKING PAPERS - RECENT RELEASES

221. Joseph B. Rose and Gary N. Chaison, "The State of the Unions: United States and Canada", April 1984.

222. Naresh C. Agarwal, "Economic Costs of Employment Discrimination", May 1984.

223. R.J. Adams, "Estimating the Extent of Collective Bargaining in Canada", May 1984.

224. Robert G. Cooper, "New Product Strategies: What Distinguishes The Top Performers", May 1984.

225. Norman P. Archer, "Selection of a Simulation Language for an MBA Modeling and Simulation Course", June 1984.

226. Robert F. Love, William G. Truscott, John H. Walker, "A Facility Location Decision: Thibodeau-Finch Transport Ltd.", June 1984.

227. R.J. Adams, "Two Policy Approaches to Labour-Management Decision Making at the Level of the Enterprise: A Comparison of the Wagner Model and Statutory Works Councils", July 1984.

228. George Saunders, "Employment and the Productivity Slowdown 1958 to 1980", September 1984.

229. George Saunders, "Industrial Relations in Canada", September 1984.

230. George Saunders, "Wages and Interest Arbitration: A Case Study of the Canadian Federal Public Service", September 1984.

231. Norman P. Archer, "Optimal Control of an Ornstein-Uhlenbeck Diffusion Process, with Applications", November 1984.

232. Robert G. Cooper, "Selecting Winning New Product Projects: The Newprod System", January 1985.

233. R.J. Adams, "Industrial Relations and the Economic Crisis Canada Moves Towards Europe", January 1985.

234. G. John Miltenburg and Itzhak Krinsky, "Evaluating Flexible Manufacturing Systems", January 1985.

235. G.J. Miltenburg, "The Transient Behaviour of Transfer Lines With Buffer Inventories", February 1985.

236. Paul D. Dowling and Robert F. Love, "An Evaluation of the Dual as a Lower Bound in Facilities Location Problems", March 1985.

237. Min Basadur, George B. Graen and Terri A. Scandura, "Improving Attitudes Toward Creative Problem Solving Among Manufacturing Engineers", March 1985.

238. Christopher K. Bart and Roderick E. White, "Managing To Make Business Unit Strategies Work", May 1985.

239. Harish C. Jain, "Sexual Harassment: Issues and Policies", June 1985.

240. George Steiner, "Searching In 2-Dimensional Partial Orders", July 1985.

241. Norman P. Archer, "Measuring the productivity of office organizations", August 1985.

242. P.L. Abad, "Coordination of Pricing, Advertising, Production and Work-Force Level Decisions in a Functionally Decentralized Firm", September 1985.

243. Robert G. Cooper, "Defining the New Product Strategy", January 1986.

244. R.G. Cooper and E.J. Kleinschmidt, "An Investigation into the New Product Process: Steps, Deficiencies, and Impact", January 1986.

245. Elko J. Kleinschmidt, "Does Foreign Ownership Matter in the Export Behaviour of Exporting Industrial Firms in Canada?", January 1986.

246. John C.C. Macintosh, "Reporting to Employees: Establishing the Demand by Employees for Information", February 1986.

247. Itzhak Krinsky and A. Leslie Robb, "Approximating the Statistical Properties of Elasticities Derived from Translog and Generalized Leontief Cost Functions", February 1986.