Causal Paradox: When A Cause Simultaneously Produces and Prevents an Effect

Michael R. Waldmann
Max Planck Institute for Psychological Research
Leopoldstr. 24, 80802 Munich, Germany
waldmann@mpipf-muenchen.mpg.de

York Hagmayer
Department of Psychology, University of Tübingen
Friedrichstr. 21, 72072 Tübingen, Germany
york.hagmayer@student.uni-tuebingen.de

Abstract

Many philosophers and psychologists argue that causal inferences are solely based on the observation of contingencies between potential causes and effects. By contrast, causal-model theory postulates that the interpretation of the learning input is governed by prior causal assumptions. Simpson's paradox is an example of this basic claim of causal-model theory. Identical observations may result in dramatically different causal impressions depending on the partitioning of the event space. Two experiments are presented that show that participants' assessment of a contingency between a potential cause and an effect is moderated by their background assumptions about the causal relevance of additional variables, and the ordering of the learning items. Despite the fact that all participants received identical learning inputs, participants' assumptions about the causal relevance of an additional grouping variable led either to the impression that the cause facilitated the effect or to an impression that it prevented the effect. Thus, the acquisition of new causal knowledge is crucially dependent on causal knowledge that is already available at the outset of the induction process.

Introduction

Causal Induction as Bottom-Up Contingency Learning

The ability to acquire causal knowledge belongs to our most central cognitive competencies. We would hardly be able to survive without knowledge about systematic relationships between events. Causal knowledge allows us to anticipate harmful or gratifying events, or to plan actions to achieve goals. Despite the fact that a large number of philosophers and psychologists have analyzed the concept of causality, no agreement has been accomplished so far. The main question of how we distinguish causal relations from accidental sequences of events remains largely unanswered.

In the past 30 years, philosophers and psychologists have become increasingly interested in probabilistic notions of causality. Our knowledge about causal facts, such as "Aspirin relieves headache" or "Smoking causes lung cancer", is often based on the observation of probabilistic relations. A number of philosophers have proposed a notion of causality that reduces causal relations to observable statistical laws (e.g., Eells, 1991; Salmon, 1971; Suppes, 1970). Roughly, it has been proposed that *causes raise the probabilities of their effects*. This idea has been adopted by psychologists who propose that causal induction involves the acquisition of knowledge about *contingencies* between causes and effects (Jenkins & Ward, 1965; Wasserman, Chatlosh & Neunaber, 1983). Formally, an (unconditional) contingency can be defined as the difference between the conditional probability of a target effect E given the presence of a potential causal factor C and its probability given the absence of the factor (i.e., $p(E|C) - p(E|\sim C)$). If this difference is greater than 0, the contingency suggests that C is a facilitatory cause, if it is smaller than 0 then C may be inhibitory for the occurrence of the effect. Typically, it has been assumed that these probabilities are estimated on the basis of stored frequency information (Cheng & Novick, 1992; Melz, Cheng, Holyoak, & Waldmann, 1993). However, a different approach assumes the acquisition of associative weights (Wasserman, Elek, Chatlosh, & Baker, 1993). Chapman and Robbins (1990) have proven, however, that, at least in situations with one single cause and one effect, the asymptotic associative weight obtained by the Rescorla-Wagner learning rule (Rescorla & Wagner, 1972) corresponds to the results of applying the contingency rule. Thus, associative as well as statistical learning theories model causal induction as bottom-up learning about statistical contingencies.

Unconditional contingencies are only a rough characterization of the theory of probabilistic causality. Unconditional contingencies may even be observed between events that are not directly causally related, or the contingencies may not express the true causal relation between the two events. Problems always occur when there are additional causal factors which are correlated with the two observed events. These additional factors may be the cause for spurious correlations, or they may alter the observed statistical relation and therefore convey a wrong impression of the true causal relation. Philosophers (e.g., Cartwright, 1983; Eells, 1991; Salmon, 1980) and cognitive psychologists (Cheng & Novick, 1992; Cheng, 1993; Melz et al., 1993; Waldmann & Holyoak, 1992) have therefore suggested *conditional contingencies* as indicators of causality. Conditional contingencies assess the contingencies between two events C, E conditional upon alternative causal factors K_i being kept constant, i.e. as $p(E|C.K_1.K_2....K_n) \quad p(E|\sim C.K_1.K_2. ...K_n)$. An isolated period denotes an "and", and each K_i a choice between the presence or the absence of the factor. For example, suppose we want to test the hypothesis that smoking (C) causes heart disease (E). Furthermore, we assume that smoking is correlated with coffee drinking which may also be a cause of heart disease. In order to test our hypothesis, we should assess the conditional contingencies between smoking and heart disease in the subpopulation of coffee drinkers (K_1)

and people who do not drink coffee ($\sim K_l$). If we then discover that smoking equally leads to heart disease in both subpopulations, we may conclude that smoking is an independent cause of this disease.

Causal-Model Theory

No attempt to reduce causal laws to purely statistical relations has so far been successful (see Cartwright, 1983, 1989; Eells, 1991). One of the reasons for this failure is the fact that statistical relations are typically insensitive to one of the most fundamental properties of causal relations, *causal priority*: We know that causes precede their effects, and not the other way around. By contrast, contingencies are typically symmetric. In situations in which a cause raises the probability of its effect, it is *ceteris paribus* true that the effect raises the probability of its cause. Thus, probabilistic relations do not represent the fundamental property of causal priority. In more complex situations with multiple causes and effects, the interpretation of the observed statistical evidence is crucially dependent on the assumed pattern of causal directionality.

Waldmann and Holyoak (1992) have therefore proposed the causal-model theory which postulates an interaction between hypothesized causal models and the learning input. The main idea is that the structure of causal models directs the interpretation of the learning input which in turn may modify the initial causal model. Statistical information about conditional and unconditional contingencies is still important but causal-model theory postulates a "tight coupling" (Wisniewski & Medin, 1994) between bottom-up input and top-down assumptions. The impact of prior assumptions about causal directionality has been demonstrated in a number of experiments. In general, these experiments have shown that identical learning inputs were treated differently depending on the structure of the causal model invoked for the interpretation of the learning experiences (Waldmann & Holyoak, 1990, 1992; Waldmann, Holyoak, & Fratianne, in press).

Previous work on the impact of top-down processes has demonstrated the impact of *domain-specific* knowledge on causal induction and learning (Chapman & Chapman, 1969; Medin, Wattenmaker, & Hampson, 1987; Pazzani, 1991). In general, these studies have shown that people use domain-specific, prior knowledge about relations between specific entities in learning situation. This knowledge may help people see relations in the learning material they otherwise would have missed (Medin et al., 1987; Pazzani, 1991). Furthermore, prior knowledge that is incompatible with the learning input may also lead to distortions and selective processing (Chapman & Chapman, 1969).

Even though causal-model theory is compatible with these types of knowledge influences, its major focus is on more abstract top-down influences. For example, knowledge about the causal direction between two events does not have any implications about the strength of the causal relation between these events. Thus, this type of knowledge does not lead to selective processing and distortions as in the cases in which specific knowledge about the strength of relations is directly transferred. Nevertheless, Waldmann et al. (in press) showed that abstract knowledge about patterns of causal directionality strongly affects the ease of learning of otherwise identical learning material.

Simpson's Paradox

Causal directionality is only one aspect of abstract prior causal knowledge that influences the interpretation of the learning input. A further problem of purely statistical theories of causality is a consequence of the fact that contingencies between two events may be affected by other causal factors. One solution for this problem, the *conditional contingency* approach, has already been mentioned. According to this theory, contingencies should not be computed over the universal set of events but over subsets of events. However, Cartwright (1983) points out that this method only yields correct results when the subsets are properly selected (see also Cheng, 1993). Conditionalizing on the wrong variables may lead to erroneous contingency estimates. An instance of this problem is known in the philosophical and statistical literature as Simpson's paradox (see Cartwright, 1983; Eells, 1991; Simpson, 1951).

Simpson's paradox describes the fact that a given contingency between two events which holds in a given population can disappear or even be reversed in all subpopulations, when the population is partitioned in certain ways. Cartwright (1983) cites a study on the graduate admissions of Berkeley that demonstrates the problem (Bickel, Hammel, & O'Connell, 1977). The graduate school of Berkeley was accused of discriminating against women. And indeed, at first sight the probabilities seemed to support the causal hypothesis that being a woman causes one to be rejected at Berkeley: the probability of admissions were higher for male students than for female students. However, the researchers looked at the data more carefully. When the admissions were analyzed separately for each department, one by one, the correlation between gender and admission completely disappeared. The reason for this was that women tended to apply to departments with higher rejection rates. Department by department women were admitted in the same ratio as male applicants, whereas across all the departments proportionally fewer women were admitted.

Table 1: An example for Simpson's paradox. The Berkeley admission problem (after Eells, 1991, p. 63).

	Dept. 1	Dept. 2	Total
Male	81/90	2/10	83/100
Female	9/10	18/90	27/100

Table 1 from Eells (1991, p.63) gives an example of how this can happen. In this example, Department 1 accepts 90% of the female and of the male applicants. Department 2 only accepts 20% of the female and the male applicants. Thus, within each department male and female applicants are accepted in the same proportions. However, more female applicants apply to Department 2 which is harder to get in. Therefore, overall, across all departments, more than three times as many male applicants (83%) are admitted than female applicants (27%).

This example may lead to the methodological suggestion that it is always a good idea to partition into subsets of events, and compute conditional contingencies. However, this strategy may also lead to false assessments. The reason why in the Berkeley admissions case the analysis should be based on the department level is that the departments are *causally relevant* for the effect under investigation. The departments decide about the admissions, and not the whole university. If, by contrast, it had been shown that the contingencies reverse when the applicants were partitioned on the basis of their roller skating skills, this would not count as an argument against sex discrimination (Cartwright, 1983). Only partitions by causally relevant variables are relevant for evaluating causal laws. If causally irrelevant variables also mattered, almost any contingency can be obtained by choosing the right partition of the event space.

What this example shows is that causal induction is crucially dependent on prior causal knowledge. New causal relations may be induced using contingency estimates. However, the contingencies only reflect causal relations when the observations are partitioned on the basis of causally relevant rather than irrelevant variables. The causal relevance of *these* partitioning variables has to be established prior to the new induction task. Thus, Simpson's paradox exemplifies the basic assumption of causal-model theory that the interpretation of the learning input is based on prior assumptions about general properties of the causal situation.

Simpson's paradox is an interesting example of how specific knowledge interacts with abstract causal strategies. It is true that knowledge about the causal relevance of the partitioning variable is domain-specific (e.g., the fact that departments decide about admissions). However, unlike in the previous research on transfer of specific knowledge, this type of knowledge does not directly bias estimates about the strength of the causal relation between the target cause (e.g., sex of applicants) and the target effect (e.g., admission). In order to get the correct results, abstract knowledge has to be activated that conditional contingencies based on causally relevant subgroups should be computed. Interestingly, the dramatic reversals obtained in situations exemplifying Simpson's paradox are not due to selective processing of individual cases or distortions of the contingency judgments. They rather are a natural consequence of unbiased processing of differentially grouped cases.

The following two experiments are designed to assess whether participants are sensitive to the potential impact of causally relevant as opposed to causally irrelevant grouping variables when assessing contingencies between a putative cause and an effect.

Experiment 1

Method

Participants' task in this experiment was to assess the strength of the causal relation between the irradiation of tropical fruit and the quality of the fruit. Participants received written instructions in which they were told that importers of tropical fruit are trying to improve the quality of the fruit by irradiating them. However, so far it is unknown whether the irradiation has a positive, a negative, or no effect on the quality of the fruit. To assess the efficacy of irradiation participants received information about the quality of samples of fruit that either had or had not been irradiated. The participants were handed a two-page list which contained information about 80 samples of fruit. Each sample was represented on one line, and for each sample participants could see whether or not the sample had been irradiated ("yes" or "no"), and whether the quality of this sample was "good" or "bad". Participants were instructed to study the list carefully in order to be able to assess whether irradiation has an effect or not. They were requested to express their impression on a rating scale that ranged from -10 ("irradiation leads to a strong deterioration of the quality") to +10 ("irradiation leads to a strong improvement of the quality").

Participants were assigned to one of three conditions. Participants in all conditions saw the same list with the 80 cases, and received the same rating instructions. Thus all participants were requested to rate the strength of the causal relation between irradiation and quality of fruit. The only difference was that in two of the conditions an additional grouping variable was mentioned which either was causally relevant or irrelevant. The third condition represents a control condition in which no grouping variable was introduced.

Participants in the condition with the causally relevant variable were told that there are two types of fruit, Taringes and Mamones. Additionally it was pointed out that it was expected that irradiation affects these two types of fruit differently. Furthermore, information was added to the list which indicated that one of the two pages showed Taringes, and the other page Mamones. In the condition with the causally irrelevant grouping variable, participants were told that, due to the large number of tests, the samples of fruit described on the two pages were randomly assigned to two different investigators. The participants in this condition saw the same list as in the condition with the causally relevant variable except that "Mamones" and "Taringes" were respectively replaced by "A" and "B" as a shorthand for the two investigators.

Table 2: Relative frequency of fruit with good quality after irradiation and no irradiation within and across the subcategories A and B of the grouping variables.

	A	B	Total
irradiation	16/36 (.44)	0/4 (.00)	16/40 (.40)
no irradiation	3/4 (.75)	5/36 (.14)	8/40 (.20)

In order to test whether participants' contingency judgments reflect their prior assumptions about the additional grouping variables, the organization of the list corresponded to a variant of Simpson's paradox.

Table 2 displays how the cases were distributed. The table displays the proportion of fruit that were of good quality after they were irradiated, and the proportion of fruit that were good without being irradiated. For example, within subgroup A 36 fruit samples were presented that were irradiated. 44% of these samples (i.e., 16 out of 36) had good quality after irradiation. As can be seen in Table 2, the arrangement of the cases resulted in a reversal of the sign of the contingencies within as opposed to across the grouping variable. Disregarding the grouping variable yields a positive contingency between irradiation and quality of fruit (.40 - .20 = .20). By contrast, the contingency within each of the subgroups is negative (A: .44 - .75 = -.31; B: .00 .14 = -.14). For half of the participants, the mapping between irradiation and quality of fruit was switched so that these participants saw a symmetric situation with a negative overall contingency, and positive contingencies within the subgroups.

It was expected that participants in the control condition who did not receive any grouping information should rely on the total proportions. Thus, participants who received the arrangement displayed in Table 2 should get the impression that irradiation *raises* the quality of fruit. The two other conditions are more interesting. If participants' behavior conformed to the normative suggestions regarding contingency assessments, they should conditionalize the contingency estimates on causally relevant grouping variables. Thus, it may be expected that participants in the condition with the causally relevant grouping variable would assess the causal impact of irradiation separately for each subgroup (Mamones and Taringes). Since the contingencies within each subgroup are negative, participants should get the overall impression that irradiation *lowers* the quality of fruit. Since participants in all conditions were requested to give an overall assessment, it was expected that participants would average the impressions they obtained for each subgroup. Finally, participants in the condition with the irrelevant grouping variable (A and B) should, according to normative standards, ignore this variable. Thus their assessments should be similar to the ones obtained in the control condition.

Results

Table 3 shows the results of this experiment based on 36 students from the University of Tübingen. The signs of the ratings of the group who saw the negative overall contingency were reversed so that the two subgroups were comparable.

Table 3: Mean ratings of the causal relation between irradiation and quality of fruit in the control condition, and the conditions with the causally relevant and irrelevant grouping variable.

relevant	irrelevant	control
-4.33	5.17	4.75

The results show that participants behaved according to normative standards. The ratings in the control condition and in the condition with the irrelevant grouping variable were positive, and statistically indistinguishable from each other. Thus, participants in these two conditions believed that irradiation *raises* the quality of fruit. This finding indicates that participants based their assessments on the total distribution of cases while disregarding subgroups. By contrast, participants in the condition with the causally relevant grouping variable got the impression that the cause prevents the effect. These participants concluded that irradiation *lowers* the quality of fruit. The negative mean rating indicates that many participants computed contingencies for each subgroup separately before these contingency estimates were integrated. The ratings of this group were very different from the ratings of the two other groups, $F(1, 33) = 71.1$, $p < .001$, $MSE = 9.71$.

Experiment 2

Method

In Experiment 1 two different grouping variables were compared, type of fruit (Mamones and Taringes) and investigators (A and B). Experiment 2 attempted to replicate the results of Experiment 1 with a grouping variable that was kept constant across the two conditions. Thus, all participants saw identical cases, received identical rating instructions, and were informed about identical subcategories. Participants were told that their task was to assess the causal efficacy of a new watering technique tried on different plants from the North African desert. Two instruction conditions were compared which only differed in one sentence. In both conditions, it was pointed out that the watering technique is applied to two types of plants, Eleusina and Setaria. But only in one of the two conditions the hint was added that the technique might have different effects on the two plant types. (This manipulation is less strong than in Experiment 1 because participants in the condition without the hint may still believe that the type of plant is causally relevant.)

Similar to Experiment 1, participants received a list of 80 cases which provided participants with information about the type of plant ("Eleusina" or "Setaria"), whether the plant was watered or not ("yes" vs. "no"), and whether the particular plant grew (+) or not (-). Again, participants' task was to rate on a rating scale between +10 and -10 whether watering generally leads to an improved or decreased growth of the plants. The assignment of the cases to the two subcategories corresponded to Table 2. Thus, calculating contingencies for each subgroup separately should cause the impression that watering leads to decreased growth of the plants, whereas disregarding the two plant types should lead to the impression that applying the new watering technique actually helps the plants to grow.

Experiment 2 also introduced a second factor, the ordering of the cases. One condition presented the information about the two plants in a blocked fashion. Thus, on one page all the Eleusinas and on the other page all the Setarias were listed. In the second condition, the presentation of the cases was unordered. On both pages information about the two plant types was given in a random order. It was expected that the hint about the causal relevance of the grouping variable should have particularly strong effects in the condition in which the cases were not grouped. Separating out the two groups is much harder in this condition so that participants should only attempt to calculate separate contingencies when they believed it was absolutely necessary. By contrast, the blocked presentation of the two groups makes the grouping variable more salient as a potential causal factor even when no hint was given.

Again for half of the participants the sign of the observed contingencies was reversed (their ratings were recoded for the analyses).

Results

Table 4 shows the results based on 48 participants from the University of Tübingen. This table displays participants' mean ratings of the causal efficacy of the new watering technique in the conditions with and without the hint about the potential causal relevance of Eleusinas and Setarias, and in the conditions with blocked versus unordered presentation of the two plant types. The negative mean ratings in three of the four conditions indicate that many participants assumed that plant type is causally relevant even when no explicit hint was given. In the two conditions with the blocked presentation of the cases participants tended to give negative ratings which indicate that they assessed the contingencies for each plant type separately before the two contingency estimates were integrated. The ratings of these two groups did not differ significantly (F<1). Grouping the cases apparently led to the impression that the plant type may be causally relevant even when no explicit hint was given. However, in the groups in which the cases were presented in an intermixed fashion, omitting the hint about the causal relevance resulted in a significant shift of the ratings toward the positive side, $F(1, 44) = 11.2$, $p < .01$, $MSE = 11.2$. Thus, participants tended to ignore the subgroups when the groups were not very salient, and when no hint about the causal relevance of the grouping variable was given.

Overall, this pattern of results yielded a reliable interaction between the factors hint and amount of grouping, $F(1, 44) = 6.25$, $p < .025$, $MSE = 11.2$.

Table 4: Mean ratings of the causal relation between watering and plant growth in the groups with or without the hint about the causal relevance of the plant type, and with or without category-based groupings of the cases.

	with hint	without hint
grouped	-2.33	-2.58
ungrouped	-3.75	.83

Discussion

Simpson's paradox is an example of the basic assumption of causal-model theory that the interpretation of the learning input is crucially dependent on prior causal knowledge. Identical cases may result in dramatically different causal impressions depending on the partitioning of the event space. Unlike experiments that demonstrate direct influences of domain specific knowledge about causal relations, Simpson's paradox is an example of the importance of abstract causal knowledge. Even though knowledge about the causal relevance of the grouping variables is domain specific, this knowledge does not directly predetermine the strength of the causal relation that is being assessed. Rather, more abstract kinds of knowledge about interactions between contingency assessments with the causal relevance of partitioning variables has to be invoked.

The two experiments showed that participants' assessment of a contingency between a potential cause and an effect is moderated by their background assumptions about the causal relevance of additional variables, and the mode of presentation of the learning items. Despite the fact that the participants of the experiments received identical learning inputs, participants' assumptions about the causal relevance of an additional grouping variable led either to the impression that the cause facilitated the effect or to an impression that it prevented the effect. These results demonstrate that the acquisition of new causal knowledge is based on old causal knowledge which is already available at the outset of the induction process.

References

Bickel, P. J., Hammel, E. A., & O'Connell, J. W. (1977). Sex bias in graduate admissions: Data from Berkeley. In W. B. Fairley & F. Mosteller (Eds.), *Statistics and public policy.* Reading, MA: Addison-Wesley.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid diagnostic signs. *Journal of Abnormal Psychology, 74*, 271-280.

Cheng, P. W. (1993). Separating causal laws from causal facts: Pressing the limits of statistical relevance. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 215-264). San Diego: Academic Press.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99*, 365-382.

Cartwright, N. (1983). *How the laws of physics lie. Essay 1.* Oxford: Clarendon Press.

Cartwright, N. (1989). *Nature's capacities and their measurement.* Oxford: Clarendon Press.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition, 18*, 537-545.

Eells, E. (1991). *Probabilistic causality.* Cambridge: Cambridge University Press.

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs, 79 (whole volume X).*

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology, 19*, 242-279.

Melz, E. R., Cheng, P. W., Holyoak, K. J., & Waldmann, M. R. (1993). Cue competition in human categorization: Contingency or the Rescorla-Wagner learning rule? Comments on Shanks (1991). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1398-1410.

Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 416-432.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F Prokasy (Eds.), *Classical conditioning II. Current research and theory* (pp. 64-99). New York: Appleton- Century-Crofts.

Salmon, W. (1971). Statistical explanation. In W. Salmon (Ed.), *Statistical explanation and statistical relevance.* Pittsburgh: University of Pittsburgh Press.

Salmon, W. C. (1980). Probabilistic causality. *Pacific Philosophical Quarterly, 61*, 50-74.

Simpson, E. H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society, Ser. B, 13*, 238-241.

Suppes, P. (1970). *A probabilistic theory of causality.* Amsterdam: North Holland.

Waldmann, M. R., & Holyoak, K. J. (1990). Can causal induction be reduced to associative learning? *Proceedings of the Twelfth Annual Conference of the Cognitive Science*

Society (pp. 190-197). Hillsdale, NJ: Erlbaum.
Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*, 222-236.
Waldmann, M. R., Holyoak, K. J., & Fratianne, A. (in press). Causal models and the acquisition of category structure. *Journal of Experimental Psychology: General.*
Wasserman, E. A., Chatlosh, D. L., & Neunaber, D. J. (1983). Perception of causal relations in humans: Factors affecting judgments of response-outcome contingencies under free-operant procedures. *Learning and Motivation, 14*, 406-432.
Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, andCognition, 19*, 174-188.
Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science, 18*, 221-281.