Journal of Experimental Psychology: General
2000, Vol. 129, No. 4, 530–545

0096-3445/00/$5.00 DOI: 10.1037//0096-3445.129.4.530

Dynamic and Consequential Consistency of Choices Between Paths of Decision Trees

Jerome R. Busemeyer, Ethan Weg, Rachel Barkan, and Xuyang Li
Indiana University

Zhengping Ma
Purdue University

The generally prescribed procedure for choosing a decision strategy from a decision tree employs a backward induction analysis that entails 3 fundamental consistency principles: dynamic, consequential, and strategic. The first requires the decision maker to follow through on plans to the end, the second requires the decision maker to focus solely on future events and final consequences given the current state of events, and the third is the conjunction of the first 2. Five experiments were reported to test these principles using different subject populations, different procedures for estimating consistency, and different factors for manipulating the attractiveness of the gamble at the final stage of the tree. The main findings were that strategic and dynamic consistency principles were violated at rates that exceeded choice inconsistency.

Most real-life decisions involve multiple stages, that is, a sequence of actions and events resulting in a series of consequences over time. Strategic decisions in business, medical, military, and foreign policy all require planning across multiple-stage future scenarios. During the past 30 years, decision researchers have learned a great deal about the basic principles of single-stage decisions. We are beginning to understand the general conditions under which rational principles such as transitivity or independence break down using single-stage gambles (see Luce, 2000, for a recent comprehensive review). However, very little is known about the principles of multistage decision making, and the purpose of this article is to test some basic assumptions about the way human decision makers plan strategies across multiple stages.

Decision Trees

Multistage decisions are usually described in terms of a graphical representation of the problem called a *decision tree* (cf. Keeney & Raiffa, 1976; Raiffa, 1968; Von Winterfeldt & Edwards, 1986). Figure 1 is a textbook example adapted from a popular book on decision analysis (Behn & Vaupel, 1982, pp. 281–293).

In this example, a 40-year-old woman is considering the possibility of bearing a child. Rectangular boxes, [ ], represent *decision nodes* (the decision maker picks the next move); ellipses, ( ), represent *event nodes* (nature picks the next move), and the solid dots, •, represent *terminal nodes* (final consequences). Branches, ----, connecting decision and event nodes form a *path* for a possible future scenario. Each decision node defines a new *stage* in the decision tree, and the maximum number of stages along a path defines the decision or planning horizon.

This particular decision tree involves three interdependent decision stages: Stage 1, whether or not to get pregnant; Stage 2, whether or not to take an amniocentesis test if pregnant; and Stage 3, whether or not to have an abortion if the test is positive (indicating a birth defect). The nodes at the very end of the tree are the final consequences. For example, the final consequence labeled C_6 represents a consequence such as "not aborting a birth defected baby." A decision *strategy* is a plan that specifies each action that will be taken at each stage as the woman travels down a path of the tree from the root to a terminal node. For example, considering Figure 1, one strategy would be to plan to get pregnant, take the test, and abort if it's positive; another strategy would be to plan to get pregnant, but not take the test.

Backward Induction

The generally prescribed procedure for choosing a decision strategy from a decision tree is called *backward induction analysis* (also known as dynamic programming or the "averaging out and folding back" method; see Bertsekas, 1976; DeGroot, 1970; Keeney & Raiffa, 1976; Raiffa, 1968; Von Winterfeldt & Edwards, 1986). For over 30 years, medical, business, and management schools have taught backward induction analysis to their students as an unquestionable method for improving planning and decision making (see, e.g., Clemen, 1996; Weinstein & Fineberg, 1980).

The analysis can be summarized briefly as follows. First, each terminal node, •, is assigned a real number representing the utility or worth of the final consequence to the decision maker. For example, in Figure 2, consequence C_6 is assigned a number symbolized as u_6 representing the utility of "not aborting a birth

Jerome R. Busemeyer, Ethan Weg, Rachel Barkan, and Xuyang Li, Department of Psychology, Indiana University; Zhengping Ma, Department of Psychology, Purdue University.

This research was supported by National Institute of Mental Health Perception and Cognition Grant R01 MH55680 and National Science Foundation Decision Risk Management Science Grant SBR-9602102.

We thank Barbara Mellers and Peter Wakker for their comments on earlier versions of this article.

Correspondence concerning this article should be addressed to Jerome R. Busemeyer, Psychology Building, Indiana University, Bloomington, Indiana 47405. Electronic mail may be sent to jbusemey@indiana.edu.

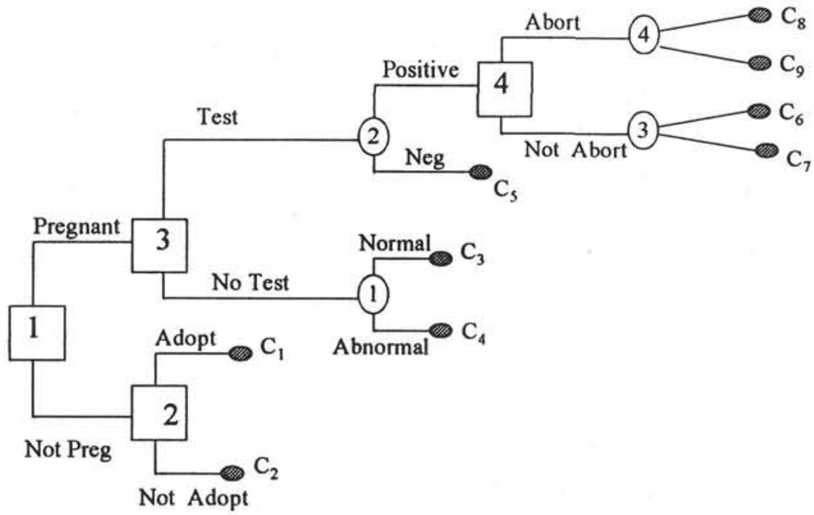

Figure 1. A decision tree involving three interdependent decision stages: Stage 1, whether or not to get pregnant (preg); Stage 2, whether or not to take an amniocentesis test if pregnant; and Stage 3, whether or not to have an abortion if the test is positive (indicating a birth defect). Neg = negative.

defected baby." Second, each event node is assigned a real number called the *expected utility* of the node, which represents the weighted average utility of the event node. For example, the event node (4) is assigned a utility denoted $u(4)$ computed by multiplying the utility u_8 by its probability p_8, multiplying the utility u_9 by its probability p_9, and summing the two products. Third, each decision node, [ ], is assigned a real number computed by taking the maximum value of the nodes that branch out of the decision node. For example, the decision node [4] is assigned a utility $u[4]$ equal to the maximum of the utilities $u(3)$ and $u(4)$ corresponding to event nodes (3) and (4).

Finally, the decision maker selects a strategy by working backward. First plan the action at decision node [4]: If abortion is planned, that is, $u(4) > u(3)$, then cut off the bottom branch after node [4] and plan to move from decision node [4] to event node (4). Second, plan the action at decision node [3] using the tree trimmed by the first plan: If the test is taken, that is, $u(2) > u(1)$, then cut off the bottom branch after node [3] and plan to move from [3] to (2). Finally, choose the action at node [1] using the tree trimmed by the second plan: If the decision to get pregnant has higher utility, then move from the root node [1] to node [3]. Thus, the key idea is to trim off unwanted branches starting at the end of the tree and working backward to the beginning.

1. Assign utilities to end nodes •,

$$C_j \rightarrow u_j$$

2. Apply MIX operator at chance nodes ( )

$$u(4) = p_8 u_8 + p_9 u_9$$
$$u(2) = p_+ u[4] + p_- u_5$$

3. Apply MAX operator at decision nodes [ ]

if $u(3) > u(4)$ then
$u[4] = \text{MAX}[u(3), u(4)] = u(3)$
if $u_1 > u_2$ then
$u[2] = \text{MAX}[u_1, u_2] = u_1$

4. Follow Maximum Path Working Backwards

If $u(3) > u(4)$ then move [4] to (3)
If $u(2) > u(1)$ then move [3] to (2)
If $u[3] > u[2]$ then move [1] to [3]

Figure 2. Outline of backward induction analysis.

The above backward induction analysis entails three fundamental consistency principles: *dynamic, consequential,* and *strategic* consistency (Hammond, 1988; Machina, 1989; Sarin & Wakker, 1998). Intuitively, dynamic consistency requires the decision maker to follow through on plans to the end. This is required for the working-backward planning strategy. Consequential consistency requires the decision maker to focus solely on the future events and final consequences given the current state. This is required for the estimation of utilities at each node. Strategic consistency results when both dynamic and consequential consistencies are satisfied. More rigorous operational definitions used to test these principles are provided below.

Empirical Tests of Consistency Principles

The principle of dynamic consistency requires that two different types of choices produce consistent preferences. For the first choice, suppose the woman is currently at the root, node [1], of the decision tree, but she is planning ahead about what to do if she reaches the decision at node [4], and suppose that she plans to abort. For the second choice, suppose the same woman does get pregnant, takes the amniocentesis test, and the test is positive, so that she has traveled down the tree and finally reaches the final decision at node [4]. To be dynamically consistent, she must follow through on her plan and carry out the abortion. If the woman violates dynamic consistency by changing her mind after taking the test and now decides not to abort regardless of the test outcome, then the cost and information of the test is wasted, and the initial decision is no longer justified.

Although the decision tree in Figure 1 is a socially relevant example, it is less than ideal for providing pristine and uncontroversial tests of dynamic consistency. For this purpose, we have devised a more highly controlled decision task, illustrated in Figure 3.

This example is a four-stage decision tree, but the number of stages can be easily manipulated. At the root of the tree, located at the lower left node [1], the decision maker must choose to either stop the trial early and take a monetary payment, denoted t, or pay a trivial price to work up the tree toward an attractive final stage. If the up branch is chosen at node [1], then the decision maker faces an event node (1), which may end the trial with no gain or loss (with some known probability $1 - p$) or continue the trial and allow the decision maker to keep working up the tree (with known probability p). If the up branch occurs at this event node, then the decision maker once again has to decide whether to pay a trivial price to continue working up the tree or stop the trial and receive the payment t. This process continues until the decision maker reaches the final stage [D], at which point he or she is faced with a choice between a sure thing, s, or a gamble (G). If the gamble is chosen, then there is a .50 probability of winning a monetary reward, R, or suffering a punishment, P (perform a tedious and boring arithmetic task that was designed to produce negative emotional affect).

The decision illustrated in Figure 3 is an abstraction of real-life decisions that entail working toward a higher goal under risks of failure or temptations of quitting early. For example, this can be

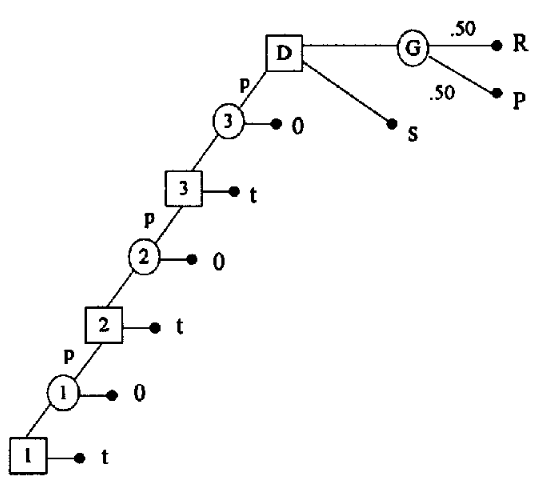

Figure 3. Decision tree used in Experiments. The first decision at node [1] is a choice between stopping the trial and taking a sure payment, denoted t, versus paying a small amount to try to work up the tree toward a final gamble. If the decision maker chooses to pay to go up on the first decision, then he or she faces an uncertain event. With some known probability, $1 - p$, the event may go down, in which case the trial ends and the decision maker receives no payment; or with probability p the event may go up, in which case the decision maker faces another decision. If the decision maker finally achieves the last stage [D], then he or she faces a choice between a sure thing, denoted s, or a gamble. If the gamble is chosen, then the decision maker can win a reward of money, denoted R, or suffer a punishment, denoted P.

considered an abstract representation of the plight of a graduate student working toward a PhD degree or an athlete working toward an Olympic medal.

Dynamic consistency is tested in this task by asking the decision maker to make two choices: A *planned* choice is made while the decision maker is still waiting at node [1] by asking him or her to make a *planned* choice for the final node [D]. This planned choice is automatically carried out later by the computer if and when the decision maker reaches the final stage. A *final* choice is made on another trial that the decision maker actually travels up the tree to node [D] and makes the decision directly facing node [D]. Dynamic consistency requires the same action to be taken on both the planned and final choices. Note that the only difference between the planned and final decisions is that the former is based on planned series of events and the latter is based on the actual realization of these same events.

Consequential consistency is a logically independent property, which entails a different pair of choices: The first choice is an *isolated* choice, in which node [D] is clipped off (by the computer) and presented by itself as a single-stage gamble and the decision maker is asked to make a choice between the sure thing versus the gamble in isolation. The second choice is identical to the *final* choice described above. Consequential consistency requires the same action to be taken on both the isolated and final choices. Note that the only difference between the final and isolated decisions is that the former occurs after experiencing a series of events and the latter omits this experience.

Strategic consistency is defined in terms of another pair of choices: The first choice is identical to the planned choice described above, whereas the second choice is identical to the isolated choice described above. Strategic consistency requires the same action to be taken on both the planned and isolated choices. Note that dynamic consistency implies that the planned choice equals the final choice and consequential consistency implies that the final choice equals the isolated choice. Simultaneously, they imply that the planned choice equals the isolated choice.

In sum, the decision-tree paradigm outlined in Figure 3 provides tests of three different types of consistencies: dynamic, consequential, and strategic. A fourth type of consistency also needs to be considered. Decision makers frequently change their preferences even when they are presented with exactly the same choice on two different occasions. For example, exactly the same planned choice may be presented on two different trials. Choice consistency requires the same action to be taken on both replications. Choice consistency provides a baseline rate for comparisons with the other three types of consistency rates.

The empirical status of the dynamic consistency principle is a critical issue for decision theorists. Expected utility theory requires dynamic, consequential, and strategic principles to be satisfied (Hammond, 1988; Machina, 1989; Sarin & Wakker, 1998). Backward induction may be used without consequential consistency, but dynamic consistency is still required (Machina, 1989; Sarin & Wakker, 1998). Dynamic consistency also provides the foundation for working-backward search procedures that underlay planning heuristics used in problem solving (see Nilsson, 1980). Backward induction analysis is difficult to justify for individuals that intentionally reject the dynamic consistency principle. In particular, backward induction may be unsuitable for emotionally laden decisions such as the one illustrated in Figure 1.

Despite the importance of these consistency principles for planning and decision-making theory, only a single experiment has ever examined them empirically. A seminal study by Cubitt, Starmer, and Sugden (1998) tested these principles using a between-subjects design, in which each participant made a single choice between a pair of gambles displayed in one of the three forms—planned, final, or isolated (similar to that described above). Although this design did not permit direct estimates of consistency rates, it did permit indirect tests by comparing choice probabilities produced by exactly the same pair of gambles across the different choice displays. Cubitt et al. (1998) found a significant difference between the choice probabilities produced by planned versus final decisions and a marginally significant difference between choice probabilities produced by planned versus isolated decisions, but they found no significant difference between the choice probabilities produced by the final versus isolated decisions.

One purpose of this article is to generalize and extend the landmark findings of Cubitt et al. (1998) by using a within-subject design. The advantages of using a within-subject design are that it provides both *indirect* tests based on choice probabilities, as well as *direct* tests based on inconsistency rates (i.e., the probability that participants change their preference across a test pair). A direct test is performed by comparing the inconsistency rate of each principle (dynamic, consequential, strategic) to the base rate of choice inconsistency.

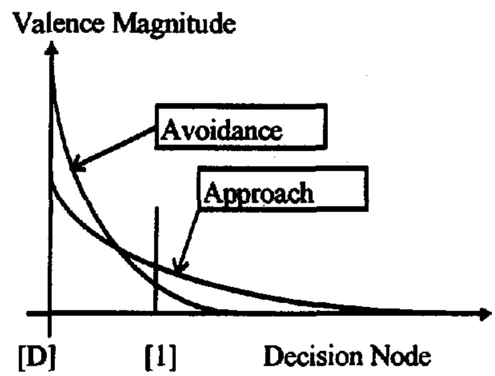

Figure 4. Application of the goal-gradient hypothesis to the decision tree illustrated in Figure 3. The horizontal axis indicates the decision nodes in the tree, where node [1] is at the beginning of the tree and node [*D*] is the final decision leading to potential gains or losses. The vertical axis represents the valence magnitude produced by the anticipated gains and losses. The steeper curve represents the avoidance gradient produced by the potential losses, and the flatter curve represents the approach gradient produced by the potential gains. When the decision maker is making a plan for node [*D*] at node [1], he or she is many stages away from the future consequences, and the approach gradient exceeds the avoidance gradient. Later when the decision maker finally arrives at node [*D*] and is faced with immediate consequences, the avoidance gradient exceeds the approach gradient.

Empirical Tests of Decision Field Theory

A second purpose of this article is to test a set of theoretical predictions derived from a psychological theory of decision making called decision field theory (DFT; Busemeyer & Townsend, 1993; Townsend & Busemeyer, 1989). DFT is one of the few psychologically based theories that make formal a priori predictions concerning violations of the previously discussed consistency principles. It is shown below that DFT predicts systematic violations of dynamic and strategic consistencies, but DFT predicts that consequential consistency will be satisfied for this experimental paradigm.

DFT is a formal application of earlier approach–avoidance theories of conflict (Miller, 1944; Lewin, 1935) to decision-making research. One of the main assumptions incorporated into DFT from earlier approach–avoidance theories is the concept of the *goal gradient*. Following Lewin's (1935) and Miller's (1944) original conflict theories, attention to the final consequences of a decision are assumed to increase as one gets closer to becoming committed to these final consequences. Empirical evidence supporting the goal-gradient hypothesis has been found with human decision makers using laboratory tasks with money (Losco & Epstein, 1977) as well as field studies of sport parachuting (Fenz & Epstein, 1967).

Figure 4 illustrates two goal gradients, and in this case, the avoidance gradient decreases with distance more rapidly than the approach gradient. Applying this idea to decision trees, we assume that the planned decision (made at node [1]) is psychologically further away from commitment to the final consequences than the final decision (made at node [D]). At the root of the tree, node [1], when the decision maker is far removed from the final consequences, the approach gradient is stronger than the avoidance, and the attractive aspects of the gamble are preferable to the sure thing. But at the final stage, when the decision maker is at node [*D*] and directly facing the consequences, the avoidance gradient now exceeds the approach, and the aversive aspects of the punishment become more salient. This change in gradients causes the decision maker to reverse his or her preference: initially planning to take the gamble but later rejecting the gamble at the final stage, thus violating the dynamic consistency principle.

Figure 4 illustrates the most common case where the slope of the avoidance gradient is steeper than the slope of the approach gradient (see Fenz & Epstein, 1967; Losco & Epstein, 1977). However, for some individuals, the opposite may be true, and the slope of the approach gradient may be steeper than the slope of the avoidance gradient. In this alternative case, the change in gradients causes the decision maker to reverse his or her preference in the other direction: initially planning to reject the gamble but later taking the gamble at the final stage, once again violating dynamic consistency. The critical point is that in either case, the goal-gradient hypothesis predicts violations of dynamic consistency.

Furthermore, the goal-gradient hypothesis predicts that isolated choices and final choices should be the same. This prediction follows from the fact that the decision maker directly faces a commitment to identical consequences under final and isolated choice conditions. Thus, the goal-gradient hypothesis predicts no violations of consequential consistency in this decision-tree paradigm.

More formally, if we apply DFT to the simple-choice problem shown in Figure 3, then probability of choosing the gamble over the sure thing will be an increasing function of the mean valence difference between the gamble versus the sure thing. For the planned-choice condition, the mean valence difference can be written as follows (cf. Equations 6a and 6b from Busemeyer & Townsend, 1993):

$$\delta_{\text{Plan}} = [(.5)g_R(3)u(R) - (.5)g_P(3)u(P)] - g_R(3)u(S), \quad \text{(Plan)}$$

where $u(R)$, $u(P)$, and $u(S)$ represent the subjective evaluations for the reward, punishment, and certain consequence; and $g_R(3)$ and $g_P(3)$ represent the weights for the approach and avoidance gradients at a distance of three stages away from the final consequences. For the final choice condition, the difference in mean valence is written similarly:

$$\delta_{\text{Final}} = [(.5)g_R(0)u(R) - (.5)g_P(0)u(P)] - g_R(0)u(S), \quad \text{(Final)}$$

where $g_R(0)$ and $g_P(0)$ represent the weights for the approach and avoidance gradients at zero distance from the final consequences. For the isolated choice condition, the mean valence difference is exactly the same as that for the final choice,

$$\delta_{\text{Isolated}} = [(.5)g_R(0)u(R) - (.5)g_P(0)u(P)] - g_R(0)u(S) = \delta_{\text{Final}}. \quad \text{(Isolated)}$$

This is because the distance is exactly the same (zero) for both the isolated and the final choices. Therefore, DFT predicts systematic violations in dynamic consistency ($\delta_{\text{Plan}} \neq \delta_{\text{Final}}$) and strategic consistency ($\delta_{\text{Plan}} \neq \delta_{\text{Isolated}}$) but no systematic violations of consequential consistency ($\delta_{\text{Final}} = \delta_{\text{Isolated}}$). Note that these predictions hold regardless of specific parameters used for the goal gradients, $g_R(D)$ and $g_P(D)$. In other words, these predictions follow for both the cases—when the avoidance slope exceeds the approach slope and when the approach slope exceeds the avoidance slope.

More specific predictions are possible, contingent on individual differences in the goal gradients. If the avoidance gradient exceeds the approach gradient, then payoff values can be selected that satisfy the inequality

$$g_R(3)/g_P(3) > [u(R) - 2u(S)]/u(P) > g_R(0)/g_P(0). \quad (1)$$

In this case, the probability of planning to choose the gamble will exceed .50, but the probability of finally choosing the gamble will fall below .50. If the approach gradient exceeds the avoidance gradient, then payoff values can be selected that satisfy the reverse inequality

$$g_R(0)/g_P(0) > [u(R) - 2u(S)]/u(P) > g_R(3)/g_P(3). \quad (2)$$

In this case, the probability of planning to choose the gamble will fall below .50, but the probability of finally choosing the gamble will exceed .50. Finally, if the approach and avoidance gradients are equal, so that $g_R(D) = g_P(D) = g(D)$, then choice probabilities for planned and final choices will still differ systematically. This follows from the assumption that $g(3) < g(0)$ so that

$$|\delta_{\text{Plan}}| = [g(3)/g(0)] \times |\delta_{\text{Final}}| < |\delta_{\text{Final}}|. \quad (3)$$

In other words, the mean valence difference must be smaller in magnitude for planned compared with final choices. In this case, it is difficult to discriminate between the gamble and sure thing during planning, but the strength of preference changes at the final choice, becoming clearer and stronger. Consequently, the probabilities must be less extreme for planned compared with final choices.

In summary, the experiments provide a priori and parameter free tests of predictions derived from DFT. Choice probabilities obtained from planned choices are predicted to differ systematically from those obtained from final and isolated choices, and the latter two are predicted to be equal. Furthermore, dynamic and strategic inconsistency rates are predicted to exceed choice inconsistency rates because of changes in goal gradients with distance between planned and final or planned and isolated choices, respectively. But consequential inconsistency rates are not expected to exceed choice inconsistency rates because the isolated and final choices used in the test of consequential consistency are both at the same (zero) distance. These predictions do not depend on any specific assumptions about the direction of differences in approach and avoidance gradients.

More specific predictions can also be made contingent on the direction of differences between approach and avoidance gradients. For the most common case where the avoidance gradient exceeds the approach, choice probabilities will change from planning to take the gamble toward finally rejecting the gamble. For the less common case where the approach gradient exceeds the avoidance, choice probabilities will change from planning to reject the gamble toward finally taking the gamble. If there are no differences between approach and avoidance gradients, then the choice probabilities are predicted to be less extreme for planned compared with final choices.

Experiments

Overview

Five experiments are reported using a within-subject design and the decision-tree task illustrated in Figure 3 to estimate dynamic, consequential, strategic, and choice consistency measures. The attractiveness of the gamble was parametrically manipulated to test for reversals of preference between planned and final choices as predicted by decision field theory. Across the five experiments, we varied participant populations, procedures for testing consistency, and factors for manipulating preferences.

First, the experiments used two different participant populations. Experiments 1a and 1b used participants from Purdue University composed of students that were 24 years old on average, and the modal student was a male (65%), Asian (56%), graduate (56%) student from engineering, a hard science, or management (77%). Experiments 2a, 2b, and 3 used participants from Indiana University composed of students that were 20 years old on average, and the modal student was a female (65%), Caucasian (86%), undergraduate (100%) student from humanities and social sciences (62%).

Second, the experiments used two different procedures for testing consistency. Experiments 1a, 2a, and 3 used what we call a *between-trial* test of consistency. For example, we tested dynamic consistency by comparing a planned choice taken from an early trial with a final choice taken from a later trial but using the same decision problem for both trials. This test places a large number of distracting problems in between the two test pairs, making it difficult to recall previous choices for the same decision problem. Experiments 1b and 2b used what we call a *within-trial* test of consistency. This procedure allowed a more stringent test of dynamic consistency by asking participants to make a plan at the beginning of a tree and then make a final choice at the end of the same tree within a single trial. Although this test has the drawback of making it easier to recall a previous choice, it permits an examination of changes in preference without interference from other extraneous decision problems.

Third, the experiments manipulated the attractiveness of the final gamble relative to the sure thing at node [*D*] in two different ways. This manipulation was designed to test the predictions from DFT that preferences for planned and final choices should reverse at some intermediate level of gamble's attractiveness relative to the sure thing. For Experiments 1 and 2, we manipulated the punishment level and provided feedback after each trial. For Experiment 3, we manipulated the sure-thing value rather than the punishment level, used larger stakes for the final gamble, and withheld all payoffs until the end of the experiment.

The results of each experiment are organized into two parts. First, we report the proportion of trials that the gamble was chosen at node [*D*] as a function of the experimental factor (punishment or sure-thing value) and display type (planned, final, isolated). These choice proportions provide an opportunity to replicate Cubitt et al. (1998), as well as test predictions derived from decision field theory.

Second, we report the comparisons between each type of consistency rate (dynamic, consequential, strategic) with the baseline choice consistency. Statistical tests of these comparisons are summarized at the end of all three experiments to maximize statistical power and minimize Type I error rates. These tests provide the first attempt ever to directly test whether or not violations of each consistency principle significantly exceed the rate expected by choice consistency alone. The theoretical implications of the results for decision field theory and other alternative explanations are examined in the General Discussion section.

Methods for Experiments 1a and 1b

Participants. There were 40 participants in Experiment 1a and 40 participants in Experiment 1b from Purdue University. All participants were students who volunteered for payment contingent on their performance. Each student participated in one session that lasted about 1.5 hr and earned about $8.00 on average depending on their performance (as described below).

Between-trial procedure. Each session began with detailed general instructions presented by the computer and extensive practice with a computer simulated spinner that was later used to represent the event nodes in Figure 3. Although the probabilities associated with event nodes were always displayed directly on each decision tree, we wanted to make sure that participants understood that the events were randomly sampled. So the computer gave them 30 practice trials with the spinner, with the spinner probabilities fixed at .25, .50, and .75 for 10 trials each.

Following the general instructions and spinner practice, participants received 11 practice problems with decision trees displayed in a form very similar to that shown in Figure 3. This practice gave participants experience with various numbers of stages, spinner probabilities, rewards, and punishments. During the practice problems, participants did not win any money. However, it was important to give participants experience with the punishments, and so they experienced punishments ranging from 0 to 60 arithmetic problems on any given practice trial, where each problem required correctly adding a pair of two digit numbers.

For the between-trial test trials, participants then received 30 decision trees, once again displayed on the computer in a format similar to that shown in Figure 3. On these 30 trials, they actually won money or received a punishment. We used 10 trials to present a five-stage tree with a planned choice at node [1] (and the computer executed the plan if the participant happened to reach node [*D*]). We used 10 trials to present a five-stage tree that permitted the participant to make the final choice (if the participant happened to reach node [*D*]). We used the remaining 10 trials to present a single-stage tree with node [*D*] presented in isolation.

The following specific parameters were used on each five-stage tree: At all decision nodes except the final stage, participants chose between paying $.01 to move up and receiving a payment t = $.04 to stop the trial; the probability that the event node would move up was displayed on the top branch from each event node as $p = .84$ (the probability of four "up" events was therefore $.84^4 = .50$). For all trees, the reward for the final gamble was fixed at R = $1.20; the sure thing was fixed at s = $.50. Participants made choices by clicking a branch on the displayed decision tree using a mouse, and we used a spinner to display the outcomes of the event nodes on the same screen as the decision tree. Thus, participants observed their choice, the random events, and the final payments as they progressed up the decision tree on each trial. If the trial ended with a punishment, then the computer would present arithmetic problems and record the participant's answer, and the next trial would not begin until the required number of arithmetic problems was correctly completed.

The punishment was varied across trials, and on each trial one of the five levels (P = 10, 20, 30, 40, 50 arithmetic problems) was presented. The punishment level was crossed with the type of choice trial (planned, final, isolated) to produce a 5 (punishment) × 3 (choice type) factorial design. Each participant experienced two replications of each cell of this design to produce the 30 experimental trials. The orders used to present these 30 trials are shown in the Appendix.

This procedure provided a between-trial consistency test using a pair of separated trials for each punishment level, where each member of the pair was separated by at least 15 trials (or about 30 min). For example, a planned choice for a five-stage tree with a punishment level of 30 problems was presented on trial 7; 19 trials later, the final choice for exactly the same tree and punishment level was presented on trial 26, and this pair of trials was used to form a test of dynamic consistency.

Within-trial procedure. The following changes were made for the within-trial test of consistency. The first 11 trials were practice trials as before, and 20 experimental trials were used. Ten of these experimental trials were single-stage decisions where node [*D*] was presented in isolation; five trials were plan–final choices concerning node [*D*], and the remaining five trials were plan–plan choices concerning node [*D*]. The orders used to present these 20 experimental trials are also shown in the Appendix.

The plan–final choice trials used a five-stage tree with a planned choice at node [1], but the participant was allowed to change his or her mind if the participant later happened to reach node [*D*]. (Contrast this with the between-trial procedure, in which the computer automatically carried out the plan on the trials that the computer requested a planned choice.) With this procedure, participants made an initial plan, and after working up through four stages, participants were asked to make another choice at the final stage within the same trial. In this case, participants were informed that the computer would randomly select either the planned choice or the final choice with equal probability to determine the payoff at node [*D*]. We used the pair of plan and final choices obtained within the same trial to perform a within-trial test of dynamic consistency.

The plan–plan trials also used a five-stage tree but with a planned choice requested twice at node [1] within the same trial. In between these two planned choices, the participant waited and observed the random outcomes of four spins of the spinner. After the second planned choice, the participant would move up the tree, and if the participant happened to reach the final node, the computer would automatically carry out one of the planned

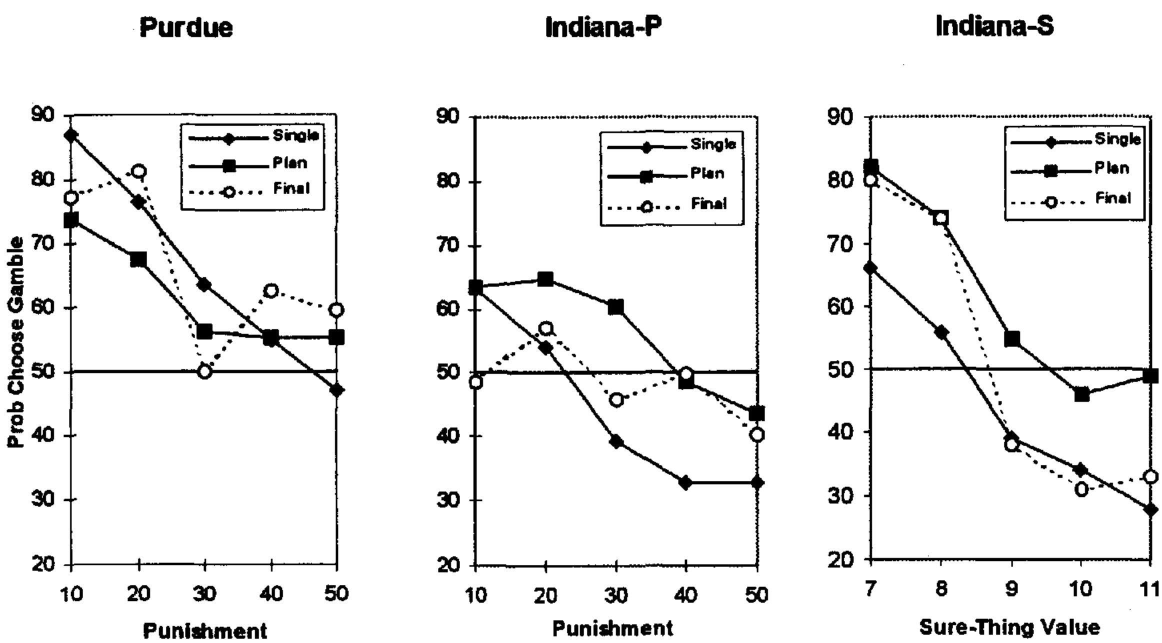

Figure 5. Probability of choosing the gamble plotted as a function of the punishment level (Purdue and Indiana-P experiments) or sure-thing value (Indiana-S experiment) separately for each of the type of choice display. Indiana-P = Experiment 2 participants; Indiana-S = Experiment 3 participants.

choices. In this case, the participants were told that the computer would randomly select either the first or the second planned choice with equal probability to determine the payoff at node [*D*]. With this procedure, a within-trial measure of choice consistency was obtained for planned choices, with approximately the same time and events between the two choices as was required to make the plan–final choice.

The within-trial procedure also provided between-trial tests of consistency (although with fewer intervening trials between pairs than was possible with the between-trials procedure). For example, an isolated choice with a punishment level of 30 problems was presented on Trial 3, then a plan–final choice with the same punishment level was presented on Trial 8, and the isolated and final choice from this pair of trials formed a between-trial test of consequential consistency.

Results for Experiments 1a and 1b

Proportion of gamble choices. For the Purdue population, the gamble was preferred over the sure thing at node [*D*], but this preference was less extreme for the planned choices compared with isolated or final choices (latter two were almost identical). The overall proportions of gamble choices were .62 (N = 1,000) for planned, .66 (N = 800) for isolated, and .67 (N = 255) for final choices. The difference between the isolated- and final-choice proportions was not statistically significant ($Z = -0.29, p > .10$). The difference between the planned-choice proportion versus the pooled average of the isolated- and final-choice proportions is statistically significant ($Z = -2.00, p < .05$).

The left panel in Figure 5 shows the proportion of trials that the gamble at node [*D*] was chosen as a function of the punishment level, broken down by decision type for the Purdue population. The results were collapsed across participants and collapsed across the between- and within-trial experiments. (No significant differences resulted from this procedural change.)

As expected, this preference decreased as the punishment level increased. More interesting, the rate of decrease was greater for single as opposed to planned choices, with the final choices zigzagging in between.[1] We performed statistical tests of the main and interaction effects by using categorical data analysis models (the standard option in the Statistical Analysis System [SAS] procedure CATMOD was used). For the Purdue population, the main effects of decision type, $\chi^2(2, N = 2{,}055) = 7.03, p < .05$, and punishment level, $\chi^2(4, N = 2{,}055) = 71.09, p < .01$, and their interaction, $\chi^2(8, N = 2{,}055) = 19.31, p < .05$, were significant.

Inconsistency rates. An *inconsistency* occurs whenever a participant changes his or her preference across a pair of test trials. Table 1 shows the inconsistency rates for Experiment 1a under the Purdue column, collapsed across punishment levels and participants. All of the inconsistencies in Table 1 were based on the between-trial procedure. The first column indicates the type of comparison: P_1 stands for the first encounter of a particular planned-choice tree, P_2 stands for the second encounter of the same planned-choice tree, F_1 and F_2 stand for the first and second encounters of a particular final-choice tree, and I_1 and I_2 stand for the pair of encounters of a particular isolated-choice tree. Using

[1] Each single-choice proportion in Figure 1 was based on 160 trials, each planned choice was based on 200 trials, and each final choice was based on approximately 50 trials. The reason for the reduced number of trials for the final choices is that participants either chose to drop out or the computer dropped them out when the downward branch was selected. The smaller sample size for the final choices caused these results to be less reliable, which explains why the curve fluctuates more for the final compared with the single or planned choices in Figure 1.

Table 1
Between-Trial Inconsistencies for Experiments 1a, 2a, and 3

Pair	Type	Purdue *M*	Purdue *SD*	Indiana-P *M*	Indiana-P *SD*	Indiana-S *M*	Indiana-S *SD*
$P_1 - P_2$	Choice	.23	200	.31	210	.29	245
$F_1 - F_2$	Choice	.18	39	.28	25	.20	80
$I_1 - I_2$	Choice	.26	200	.32	210	.29	245
$I_1 - F_2$	Consequential	.23	84	.32	71	.31	273
$P_1 - I_2$	Strategic	.33	200	.44	210	.40	490
$P_1 - F_2$	Dynamic	.31	84	.45	71	.37	273

Note. Final choice frequencies are reduced by failures to reach node [*D*]. I = isolated choice, P = planned choice, F = final choice. Indiana-P = Experiment 2 participants; Indiana-S = Experiment 3 participants.

this notation, $P_1 - P_2$, $F_1 - F_2$, and $I_1 - I_2$ symbolize choice inconsistency pairs for planned, final, and isolated choices; $I_1 - F_2$, $P_1 - I_2$, and $P_1 - F_2$ symbolize consequential, strategic, and dynamic inconsistency test pairs.

The dynamic inconsistency rate (.31) and the strategic inconsistency rate (.33) were similar, and both were higher than the choice inconsistency rates for either final (.18) or planned (.23) choices. However, the consequential inconsistency rate (.23) fell below the choice inconsistency rate for isolated choices (.26). (Note: Statistical tests of inconsistency rates are presented in the summary section after all three experiments.)

Table 2 shows the inconsistency rates for the within-trial experiment displayed under the Purdue column in a manner similar to Table 1. The two rows at the top of Table 2 show the results for the within-trial tests of dynamic inconsistency. The within-trial dynamic inconsistency rate for Purdue students only slightly exceeded the corresponding choice inconsistency rate.

The five rows of the bottom of Table 2 show the results for the between-trial tests from this same experiment. Both dynamic and strategic inconsistency rates exceeded the choice inconsistency rates for the Purdue populations. The consequential inconsistency rate was only slightly higher than the choice inconsistency rates. In general, the between-trial test results in Table 2 for the Purdue students replicate the pattern of results from the between-trial test shown in Table 1 for the Purdue students.

Discussion of Experiments 1a and 1b

The Purdue students generally favored taking the gamble across all of the conditions. However, the probabilities based on planned choices were less extreme compared with the probabilities based on either isolated or final choices, and the latter two differed by very little. This pattern of results provides moderate indirect evidence for systematic violations of dynamic and strategic consistency principles, and no evidence for violations of the consequential consistency principle, in accord with Cubitt et al.'s (1998) previous findings.

The present experiment also provides new direct evidence for these two conclusions on the basis of the between-trial estimates of consistency. These results indicated that the dynamic and strategic inconsistency rates exceeded the choice inconsistency rate, but the consequential inconsistency rate did not. However, the results obtained from the within-trial test failed to show much evidence for dynamic inconsistency over and above that expected by choice inconsistency. The latter result is not too surprising given that it is more difficult to observe any inconsistencies at all using the within-trial test procedure. This is because it is easier for participants to recall their previous choice when plan and final choices are made within the same trial. The between-trial test places a large number of distracting problems in between the two test pairs, making it more difficult to recall previous choices for the same decision problem. Given the theoretical significance of the consistency principles being tested, and the paucity of research on this issue, and the unusual characteristics of the Purdue population, it was imperative to replicate these findings with a new participant population, experimenter, and laboratory setting.

Methods for Experiments 2a and 2b

There were 42 participants in Experiment 2a and 37 participants in Experiment 2b from Indiana University. All participants were students who volunteered for payment contingent on their performance. Each student participated in one session that lasted about 1.5 hr and earned about $8.00 on average depending on their performance (as described earlier). Experiments 1b used the between-trial test procedure, and Experiment 2b used the within-trial test procedure for testing consistency. The remaining details are the same as for Experiments 1a and 1b.

Results for Experiments 2a and 2b

Proportion of gamble choices. For the Indiana population, participants preferred to Gamble at node [*D*] for planned decisions, but they changed preferences and chose not to gamble on isolated or final decisions (the latter two were almost identical). The overall proportions of gamble choices were .56 (N = 975) for planned, .43 (N = 790) for isolated, and .48 (N = 193) for final choices. The difference between the isolated- and final-choice proportions is not statistically significant ($Z = -1.25$, $p > .05$). The difference between the planned-choice proportion versus the pooled average of the isolated- and final-choice proportions is statistically significant ($Z = 5.32$, $p < .01$).

The middle panel in Figure 5 shows the proportion of trials that the gamble at node [*D*] was chosen as a function of the punishment

Table 2
Within- and Between-Trial Inconsistencies for Experiments 1b and 2b

Pair	Type	Purdue *M*	Purdue *SD*	Indiana *M*	Indiana *SD*
Within					
$P_1 - P_2$	Choice	.22	200	.19	185
$P_1 - F_2$	Dynamic	.23	96	.27	56
Between					
$I_1 - I_2$	Choice	.28	200	.34	185
$P_1 - P_2$	Choice	.29	200	.29	185
$I_1 - F_2$	Consequential	.32	190	.36	112
$P_1 - I_2$	Strategic	.38	800	.43	740
$P_1 - F_2$	Dynamic	.35	192	.39	112

Note. Final choice frequencies are reduced by failures to reach node [*D*]. For the between-trial comparisons, only the first of the two planned choices from the plan–plan trials was included in the analyses. I = isolated choice, P = planned choice, F = final choice.

level, broken down by decision type for the Indiana students in Experiments 2a and 2b. The results were collapsed across participants, and collapsed across the between- and within-trial experiments. (Once again, no significant differences resulted from this procedural change.)

As expected, preference for the gamble decreased as the punishment level increased. Note, however, that at the intermediate level (30) of punishment, planned choices still favored the gamble (60%), but isolated and final choices both switched to favor the sure thing (39% and 45%, respectively). We performed statistical tests of the main and interaction effects using categorical data analysis models (the standard option in the SAS procedure CATMOD was used). For the Indiana population, the main effects of decision type, $\chi^2(2, N = 1{,}958) = 25.03, p < .01$, and punishment level, $\chi^2(4, N = 1{,}958) = 29.71, p < .01$, were significant, but the interaction failed to reach significance, $\chi^2(8, N = 1{,}958) = 12.31, p > .05$.

Inconsistency rates. Table 1 also shows the inconsistency rates for Experiment 2a under the Indiana-P column, collapsed across punishment levels and participants. All of the inconsistencies in Table 1 were based on the between-trial procedure, and the rows are interpreted in the same way as they were defined earlier during the description of the results for Experiment 1a.

Once again, the dynamic inconsistency rate (.45) and the strategic inconsistency rate (.44) were similar, and both were much higher than the choice inconsistency rates for either final (.28) or planned (.31) choices. Also, the consequential inconsistency rate (.32) did not differ from the choice inconsistency rate (.32) for isolated choices. (Note: Statistical tests are presented in the summary.)

Table 2 shows the inconsistency rates for the within-trial experiment under the Indiana-P column displayed in a manner similar to Table 1. The two rows at the top of Table 2 show the results for the within-trial tests of dynamic inconsistency. The within-trial dynamic inconsistency rate for Indiana exceeded the corresponding choice inconsistency rate.

The last five rows at the bottom of Table 2 show the results for the between-trial tests. Both the dynamic and the strategic inconsistency rates exceeded the choice inconsistency rates for the Indiana population. The consequential inconsistency rate was only slightly higher than the choice inconsistency rates. In general, the between-trial test results in Table 2 for the Indiana participants replicate the pattern of results from the between-trial test shown in Table 1 for the Indiana participants.

Discussion of Experiments 2a and 2b

The results for the Indiana population indicate that probabilities based on planned choices generally favored taking the gamble, but probabilities based on either final or isolated choices generally favored taking the sure thing. This reversal of preference provides strong indirect evidence for systematic violations of dynamic and strategic consistency principles and no evidence for violations of the consequential consistency principle, in agreement with Cubitt et al. (1998).

The present experiment provides even more direct evidence for these two conclusions on the basis of the consistency estimates. The between-trial estimates indicated that the dynamic and strategic inconsistency rates exceeded the choice inconsistency rate, but the consequential inconsistency rate did not. Unlike the first experiment, the within-trial estimates obtained from the Indiana population also produced greater dynamic inconsistency compared with choice inconsistency. This result was not due to a higher within-trial choice inconsistency rate for the Indiana students—in fact the opposite was true. Instead, the higher rate of dynamic inconsistency for Indiana compared with Purdue students may reflect group differences in the attractiveness of the gamble.

Thus far it has been assumed that the observed dynamic inconsistencies were caused by the multiple stages of actions and events that intercede between the plan and final decisions. Alternatively, one might argue that the results are caused by the time required for these events to take place, that is, temporal inconsistencies in preferences.[2] When making a planned decision, participants chose between consequences that were delayed from reception for about 1 min (waiting to get past five stages). However, when making a final decision, participants chose between consequences that were delayed only a few seconds (waiting to get past one stage). Previous studies of intertemporal choice have shown that preferences can change simply by delaying the delivery of consequences (Hoch & Loewenstein, 1991; Thaler, 1981). However, these intertemporal choice studies involved weeks of delay, whereas our participants waited less than 1 min to experience the consequence of the planned decision. Nevertheless, it is possible to eliminate this explanation by delivering the consequences for planned and final decisions after the same amount of time. This was one purpose of the third experiment. Another purpose was to replicate the tests of the three consistency principles using larger stakes (ranging from $7 to $11) for the final gamble.

Method for Experiment 3

There were 49 participants that participated in Experiment 3 from Indiana University. All participants were students who volunteered for payment contingent on their performance. Each student participated in one session that lasted about 1.5 hr and earned about $8.00 on average depending on their performance (as described earlier). Experiment 3 used the between-trial test procedure, and all the remaining details are the same as for Experiments 1a and 2a, except where noted below.

[2] The dynamic-consistency principle needs to be distinguished from the temporal-consistency principle, where the latter is defined in terms of the following pair of choices: For the first choice, an individual is asked at time t_0 to choose between consequence A delivered at a future time $t_A > t_0$ versus consequence B delivered at a future time $t_B > t_0$; for the second choice, the individual is asked again at time t_0 to choose between consequence A delivered at a later time $(t_A + d)$ versus consequence B at a later time $(t_B + d)$, with $d > 0$. To be temporally consistent, the individual should make the same choice for each pair. In other words, adding a constant delay to the reception of both consequences should not change the preference (cf. Hoch & Loewenstein, 1991). Temporal consistency requires consistency across the passage of time intervals, but it does not require consistency across actions and events, whereas dynamic consistency does not require consistency across time intervals but does require consistent plans across actions and events (compare Strotz, 1956 vs. Machina, 1989, for alternative views). The relation of the present work to temporal consistency was brought to our attention by discussions with George Loewenstein and Chris Hsee at the *New Directions in Decision Making* conference sponsored by Northwestern University in 1997, organized by Doug Medin and Max Bazerman.

There were two primary changes in design and procedure for Experiment 3. First, the payoffs for choices at node [D] of the decision tree shown in Figure 3 were changed as follows. Rather than manipulating the punishment as in the previous experiments, this time we manipulated the sure-thing value across five levels ($7, $8, $9, $10, $11). Furthermore, we no longer used arithmetic problems for punishment. Instead, the gamble was presented as a choice between two equally likely monetary outcomes that depended on the value of the sure thing. Consider, for example, a trial on which the sure thing was set equal to $9. In this case, the top branch of the gamble stated, "win more than $9" and the bottom branch stated, "win less than $9." They were told that if the spinner happened to land on the up branch of the gamble, then the computer would "randomly select a dollar value larger than that shown for the sure thing," and if the spinner happened to choose the down branch of the gamble, then the computer would "randomly select a dollar value smaller than that shown for the sure thing." Here the monetary outcomes were uncertain—for example, the participant didn't know how much more or less than $9 they could earn. The purpose of using uncertain outcomes was to make it difficult for decisions to be based on simple numerical calculations like expected value.

Second, participants were not given any feedback at the end of each trial. They were not informed about whether they won or lost, and they were not shown any amounts earned. The final pay, they were told, would be determined at the end of the experiment by averaging the payoffs from six trials randomly sampled out of all of the regular trials of the experiment (excluding practice trials). Unknown to the participants, the computer randomly selected payoffs from a uniform distribution. On the trials that the spinner landed on the up branch, the distribution was uniform over (S, S + $5), and on the trials that the spinner landed on the down branch, the distribution was uniform over (S − $5, S). Using this procedure, the consequences for all three types of decisions (planned, isolated, final) were delayed until the very end of the experiment with the same average time interval for each type of decision.

Results for Experiment 3

Proportion of gamble choices. Similar to Experiment 2, participants preferred to gamble at node [D] for planned decisions, but they changed preferences and chose not to gamble on isolated choices, or they were indifferent on final decisions. The overall proportions of gamble choices were .61 ($N = 490$) for planned, .44 ($N = 490$) for isolated, and .50 ($N = 273$) for final choices. The difference between the isolated and the final choice proportions is not statistically significant ($Z = -1.59$, $p > .05$). The difference between the planned choice proportion and the pooled average of the isolated and final choice proportions is statistically significant ($Z = 5.13$, $p < .01$).

The right panel in Figure 5 shows the proportion of trials that the gamble at node [D] was chosen as a function of the sure-thing value, broken down by decision type for the Indiana students in Experiment 3. As expected, this preference decreased as the sure-thing value increased. Note, however, that at the intermediate value ($9) of the sure thing, planned choices still favored the gamble (55%), but isolated and final choices both switched to favor the sure thing (39% and 38%, respectively). Statistical tests of the main and interaction effects were performed using categorical data analysis models (the standard option in the SAS procedure CATMOD was used). For Experiment 3, the main effects of decision type, $\chi^2(2, N = 1{,}253) = 29.75$, $p < .01$, and sure-thing value, $\chi^2(4, N = 1{,}253) = 117.87$, $p < .01$, were significant, but the interaction failed to reach significance, $\chi^2(8, N = 1{,}253) = 5.62$, $p > .05$.

Inconsistency rates. Table 1 also shows the inconsistency rates for Experiment 3 under the Indiana-S column, collapsed across sure-thing values levels and participants. All of the inconsistencies in Table 1 were based on the between-trial procedure, and the rows are interpreted in the same way as they were defined in the earlier description of the results for Experiments 1a and 2a.

Once again, the dynamic inconsistency rate (.37) and the strategic inconsistency rate (.40) were similar, and both were much higher than the choice inconsistency rates for either final (.20) or planned (.29) choices. Also like the previous results, the consequential inconsistency rate (.31) differed very little from the choice inconsistency rate for isolated choices (.29). (Note: Statistical tests are presented in the summary.)

Discussion of Experiment 3

Despite major changes in design and procedure, the results for Experiment 3 closely matched those found with Experiment 2. First of all, the probabilities based on planned choices generally favored taking the gamble, but probabilities based on either final or isolated choices generally favored taking the sure thing. Again, this reversal of preference provides strong indirect evidence for systematic violations of dynamic and strategic consistency principles and no evidence for violations of the consequential consistency principle, confirming the results of Cubitt et al. (1998).

The present experiment provides even more direct evidence for these two conclusions on the basis of the between-trial consistency estimates. The results clearly indicated that the dynamic and strategic inconsistency rates exceeded the choice inconsistency rate, but the consequential inconsistency rate did not. Note that in the present study, the consequences for planned and final decisions were delivered after the same amount of time. Thus, the dynamic inconsistencies found in this experiment cannot be explained by temporal inconsistencies. Instead, plans based on imaged occurrence of events are sometimes difficult to follow after one actually realizes these same events.

Summary Tests of Consistency Principles

A violation of dynamic consistency occurs if the dynamic inconsistency rate is significantly greater than the choice inconsistency rate, and a violation of consequential consistency occurs if the consequential inconsistency rate is significantly greater than the choice inconsistency rate. A similar definition holds for a violation of strategic consistency.

We performed statistical tests for violations of each principle using the between-trial consistency estimates. The latter were used because more data is available from this method. Also, because of the similarity of the pattern of results found between Table 1 and Table 2 for the first two experiments, these proportions were pooled. For example, the 84 observations from the last row for the Purdue students in Table 1 were pooled with the 192 observations from the last row of the Purdue students in Table 2. Likewise, the 71 observations from the last row of the Indiana-P students in Table 1 were pooled with the 112 observations from the last row of the Indiana-P students in Table 2. This produced a single set of between-trial consistency estimates for each of the three experiments.

Table 3
Z Statistics Used to Statistically Test for Violations of Each Type of Consistency Principle

Pair	Type	Purdue	Indiana-P	Indiana-S
$I_1 - F_2$	Consequential	.88	.44	1.13
$P_1 - I_2$	Strategic	4.73	5.01	3.62
$P_1 - F_2$	Dynamic	2.45	2.73	2.68

Note. Critical $z = 1.65$ to reject H_0 at $\alpha = .05$, one tail. The Z statistic tests the difference between two proportions: One is an estimate of the type of inconsistency indicated by each row of the first column, and the second is the choice inconsistency rate (see footnote 3). I = isolated choice, P = planned choice, F = final choice. Indiana-P = Experiment 2 participants; Indiana-S = Experiment 3 participants.

Table 3 provides statistical tests of differences between proportions for each type of inconsistency (consequential, strategic, and dynamic) versus choice inconsistency.[3] The first two columns indicate the type of inconsistency being tested (consequential, strategic, dynamic) and the last three columns indicate the corresponding Z statistic, separately for Experiments 1, 2, and 3. (Note that the critical $z = 1.64$ for a one-tail test at $\alpha = .05$). In sum, the dynamic and strategic inconsistency rates are both significantly greater than the choice inconsistency rates, but the consequential inconsistency rate is not significantly different from the choice inconsistency rate. This basic pattern held for all three experiments when using the between-trial estimates of consistency.

We performed additional statistical tests to directly compare the rates produced by consequential, strategic, and dynamic inconsistency (pooled across all three experiments to maximize power and minimize Type 1 errors). The difference between the strategic and consequential inconsistency rates is statistically significant ($Z = 4.3$, $p < .01$). The difference between the dynamic and consequential inconsistency rates is also statistically significant ($Z = 2.29$, $p < .05$). However, the difference between the strategic and dynamic inconsistency rates is not statistically significant ($Z = 1.53$, $p > .05$).

One last result concerns the direction of change that occurred when participants were inconsistent across planned and final choices. For this analysis, we estimated the probability of choosing the gamble on the planned choice and then choosing the sure thing on the final choice, given that an inconsistent choice was made.[4] These analyses can be summarized as follows across all the tests from each of the two populations. For the Purdue students, 48% of the 48 inconsistent choices were based on switching from planning to take the gamble to finally taking the sure thing. For the Indiana students, 65% of the 147 inconsistent choices were based on switching from planning to take the gamble to finally taking the sure thing.

General Discussion

Complex strategic decisions involving multiple stages of actions and events can be represented graphically as decision trees. The generally prescribed procedure for choosing a strategy from a decision tree is a backward induction analysis that entails three fundamental consistency principles: dynamic, consequential, and strategic. These principles require the decision maker to have consistent preferences across different types of choices. One type is a planned choice, where the decision maker makes a commitment at the beginning of a tree about a decision occurring at the final stage in the tree. Another type is a final choice, where the decision maker actually travels down the tree and makes the decision at the final stage. A third type is an isolated choice, where the final stage is clipped off and the decision maker makes a choice after omitting the earlier stages of the tree. Dynamic consistency requires the same action to be taken on both the planned and final choices, consequential consistency requires the same action to be taken on both the isolated and final choices, and strategic consistency requires the same action to be taken on both the planned and isolated choices. We reported five experiments to test these principles using different participant populations, different procedures for estimating consistency, and different factors for manipulating the attractiveness of the gamble relative to the sure thing at the final stage of the tree.

Tests of Consistency Principles

The main conclusions from this work are that both dynamic and strategic consistency principles are systematically violated, but there is no evidence for systematic violations of consequential consistency in this paradigm. These conclusions are based on two lines of evidence.

The first line of evidence is indirect, being based on the probability of taking a gamble obtained under different choice displays for the same gamble. In general, the probabilities produced by the planned decision differed significantly from the probabilities produced by the final or isolated decisions, and the latter two were not different. The Purdue participants generally preferred to take the gamble, but the probability was less extreme for planned choices compared with final or isolated choices. The Indiana participants reversed their preferences—they tended to prefer to take the gamble during planning, but they tended to prefer not to take the gamble during the final or isolated decisions. The latter result was

[3] Consequential inconsistency rates are estimated from isolated–final pairs, and they were compared with a choice inconsistency estimate obtained by pooling isolated–isolated pairs and final–final pairs. Strategic inconsistency rates are estimated from plan–isolated pairs, and they were compared with a choice inconsistency estimate obtained by pooling plan–plan pairs and isolated–isolated pairs. Dynamic inconsistency rates are estimated from plan–final pairs, and they were compared with choice inconsistency estimates obtained by pooling plan–plan pairs and final–final pairs.

[4] This statistic was based on the following theoretical rationale. The joint probability of changing from a planned choice for the gamble to a final choice for the sure thing is postulated to be a product of three theoretical probabilities: p_p = the probability of choosing the gamble on the planned decision, m = the probability of not recalling the preference made for the planned decision, and q_f = the probability of choosing the sure thing on the final decision. The joint probability of changing from a planned choice for the sure thing to a final choice for the gamble is also postulated to be a product of three theoretical probabilities: q_p = the probability of choosing the sure thing on the planned decision, m = the probability of not recalling the preference made for the planned decision, and p_f = the probability of choosing the gamble on the final decision. According to these assumptions, we compute $p_p m q_f / (p_p m q_f + q_p m p_f)$. If there are no changes in the basic choice probabilities from planned to final decisions, then $p_p = p_f = p$ and $q_p = q_f = q$. In this case, this ratio should equal .50.

particularly strong for intermediate levels of punishment in Experiment 2 or intermediate levels of sure-thing value in Experiment 3 (see Figure 5). This indirect evidence based on choice probabilities is consistent with previous results reported by Cubitt et al. (1998).

The second line of evidence is more direct, being based on comparisons of dynamic, consequential, strategic, and choice inconsistency rates. We used two different procedures to estimate these consistency rates. The between-trial test compared planned and final choices that were separated by a large number of trials, making it difficult to base the final choice on memory recall of the planned choice. For all five experiments, this procedure produced 8% and 10% higher rates of dynamic and strategic inconsistencies, respectively, compared with choice inconsistency, and only a 1% higher rate of consequential inconsistency compared with choice inconsistency. This is the first study to demonstrate that dynamic and strategic inconsistencies occur at significantly higher rates than what one could expect from choice inconsistency alone (see Table 3).

We also used a within-trial test procedure to estimate and compare dynamic and choice inconsistency rates. This procedure compared planned and final choices made within the same choice trial, making it easy to base the final choice on memory recall of the planned choice. In fact, the choice inconsistency rate dropped from about 30% to 20% from the between- to the within-test procedures. As expected, this procedure also weakened the treatment effect—the dynamic inconsistency rate was only 5% higher than the choice inconsistency rate for the within-trial test.

Limitations and Extensions

The concepts of dynamic and consequential consistency examined in this research were defined in a highly specific manner. This definition was necessary to provide rigorous tests of the axiomatic foundations for backward induction analysis (cf. Machina, 1989; Sarin & Wakker, 1998). It is important, however, to clearly distinguish these highly constrained definitions from the more general and popular use of these terms (cf. Shafir & Tversky, 1992).

Intuitively, dynamic consistency requires decision makers to follow through on their plans to the end. However, this requirement does not imply that decision makers should ignore newly acquired information. For example, suppose a participant initially makes a plan based on instructions that the probability of winning the final gamble is .50. However, later, when she actually reaches the final stage, she learns from another reliable source that the probability of winning the final gamble is actually below .25. In this case, the participant may change her plans, but this reflects learning and not dynamic inconsistency. An antecedent condition for a test of dynamic consistency requires the conditional probabilities, given the entailed events, to remain the same for planned and final decisions.

Intuitively, consequential consistency requires decision makers to focus solely on final consequences and future events. However, this requirement does not imply that decision makers should ignore all past consequences. For example, suppose an individual would accept an offer to buy a lottery ticket for a trip to Rome when it is presented in isolation. However, now suppose that this individual initially purchased a lottery ticket for Paris and won, and afterward he was offered the lottery ticket for Rome. In this case, the individual may change his mind, but this change reflects consumption and not consequential inconsistency. An antecedent condition for a test of consequential consistency requires the final consequences to remain the same for the isolated and final decisions.

The conclusions concerning the consistency principles are also initially limited to the paradigm illustrated in Figure 3, which is an abstraction of real-life decisions that entail working toward a higher goal under risks of failure or temptations of quitting early. These principles also could be tested with more complex trees such as that shown in Figure 1, and such tests are an important direction for future research.

Nevertheless, the decision problem in Figure 3 is a generalization of the standard single-stage gambling decision that is so popular among decision researchers (see Goldstein & Weber, 1995). Furthermore, there is evidence that the present results hold more broadly, and so these conclusions can be extended to a wider range of situations.

First, the present experiments displayed the choice problems graphically as decision trees, and one might question whether or not these results generalize to more common presentations of decision problems. To answer this question, we have recently succeeded in extending our dynamic inconsistency findings to text-based presentations of multistage decision problems (Barkan & Busemeyer, 1999). Furthermore, the violations of dynamic consistency reported by Cubitt et al. (1998) were based on purely text-based presentations.

Second, the present experiments used payoffs ranging from $1 to $9, and one might wonder whether or not the same results occur with larger sums. Cubitt et al. (1998) used larger amounts (up to $40) and obtained the same pattern of results. Thus, the findings generalize across the range of payoffs that are normally used in laboratory experiments.

Third, violations of dynamic consistency have now been obtained with three different participant populations. The Purdue students were primarily older male Asian engineering and management students; the Indiana students were primarily younger female liberal arts students. Furthermore, Cubitt et al.'s (1998) participants were British from a variety of academic backgrounds.

Finally, the previous experiments used highly artificial gambles to test dynamic consistency, and it is reasonable to ask if these results are applicable to real-life situations as exemplified by Figure 1. Some preliminary evidence from a medical decision-making study by Christensen-Szalanski (1984) indicates that dynamic inconsistency occurs in the field as well as the laboratory.

Relation to Other Findings

A phenomenon called the *disjunction effect* (Tversky & Shafir, 1992) has a close relation to dynamic consistency. Tversky and Shafir asked three groups of participants to imagine that they just finished playing a gamble, and they had to decide whether or not to play the same gamble again. One group was asked to imagine they lost the first gamble; a second group was asked to imagine they won the first gamble, and a third group was asked to imagine that they did not know the outcome of the first gamble. The results indicated that the first two groups both preferred playing the second gamble, but the third group preferred not to play the second gamble.

The purpose of Tversky and Shafir's (1992) study was to test a key axiom of subjective expected utility theory (Savage, 1954). In our terminology, the first group made a planned decision conditioned on winning, the second made a plan conditioned on losing, and the third made an unconditional plan. According to the axiom, if the decision maker plans to play the second gamble independent of the first play's assumed outcome, then the decision maker should plan to play the second gamble even when the first outcome is unknown. Tversky and Shafir's (1992) results violated this axiom.

If the above results could be reproduced using outcomes that were actually experienced rather than imagined during the first play, then the results of this modified experiment could be reinterpreted as evidence for violations of dynamic consistency. Recently, Barkan and Busemeyer (1999) carried out this experiment and, as expected, found systematic violations of dynamic consistency.

Another phenomenon called the pseudo-certainty effect (Tversky & Kahneman, 1981) has a direct bearing on the issue of strategic consistency. Tversky and Kahneman (1981) gave one group of participants a simple choice between a gamble and a sure thing, corresponding to an isolated type of decision. A second group was faced with a two-stage decision (a two-stage version of Figure 1) and was asked to make a planned choice about the second stage before knowing the outcome of the first stage. A third group was presented a single-stage decision, but with payoff probabilities equated to those obtained from the two-stage display. Tversky and Kahneman found no difference between the first and second groups (both preferred the sure thing) but a large difference between the second and third groups (the latter preferred the gamble).

Note that the comparison between the first (isolated decision) and second (planned decision) groups provides an indirect test of strategic consistency. The lack of difference between these two groups indicates that the strategic consistency principle was obeyed in the Tversky and Kahneman (1981) study. The comparison of the second group with the last group constitutes a test of the reduction principle, which is a key axiom of expected utility theory (cf. Machina, 1989).

The findings regarding strategic consistency reported by Tversky and Kahneman (1981) seem to conflict with the violations of strategic consistency found in the present research. One explanation may be that the present study used a larger number of stages (five) than used by Tversky and Kahneman (1981). However, Cubitt et al. (1998) also used two-stage problems and methods very similar to Tversky and Kahneman (1981). However, they found that 38% ($N = 50$) of the participants preferred the gamble in the isolated choice, and 57% ($N = 51$) preferred the gamble in the planned decision, that is, violations of strategic inconsistency. Further research is needed to clarify this apparent discrepancy.

One last phenomenon, called the *sunk cost effect* (Thaler, 1980), seems to have implications for the consequential consistency principle. As an example, Arkes and Blumer (1985) gave one group of participants a hypothetical choice between investing an additional million dollars to complete a project that already cost them 9 million or to discontinue investing in the project. A second group was asked whether or not they would be willing to step into the middle and invest 1 million to complete the same project (but without any prior investment in the project). Arkes and Blumer (1985) found that the first group preferred to invest the extra million but the second did not. If we treat the first group as providing the final choice in a two-stage decision and the second group as providing the isolated choice, then the final and isolated choices are quite different.

A closer inspection of the sunk cost effect reveals that it does not necessarily imply a violation of the consequential consistency principle because the consequential consistency test requires identical payoffs for the final and isolated choices. In the Arkes and Blumer (1985) study, the final choice faced by the first group is to continue the project (receiving an uncertain return from the project minus the 10 million cost of the investment) versus discontinue the project (losing the initial 9 million invested). The isolated choice faced by the second group is to step into the middle of the project (receiving an uncertain return from the project minus the 1 million cost of the investment) versus not step into it (lose nothing). Note that the consequences of the final choice for the first group (e.g., uncertain return and a cost of 10 million) are not the same as the consequences faced by the participants in the second group (uncertain return and a cost of 1 million).[5] Therefore, the two groups used in the sunk cost paradigm do not satisfy the requirements for a test of consequential consistency. In sum, the sunk cost effect does not constitute a violation of the consequential consistency principle required for backward induction analysis. More generally, consequential consistency does not imply that the decision maker has to ignore past consequences because the final outcomes must include all of the consequences that occur along a path of a decision tree.

Tests of DFT

The pattern of inconsistency rates provides a strong a priori test of DFT. Recall that planned decisions are made several stages away from the final consequences, but final and isolated decisions are both made immediately facing the final consequences. Accordingly, DFT predicts that dynamic and strategic inconsistency rates should exceed consequential and choice inconsistency rates, and the latter two should not differ.[6] Furthermore, this prediction holds regardless of whether or not the approach and avoidance gradients are assumed to differ. The results of all three experiments, summarized in Table 3, clearly support this general prediction of the theory.

[5] The difference between the utility of uncertain return and a cost of 10 million and the utility of a cost of 9 million is not necessarily ordered the same as the difference between the utility of uncertain return and a cost of 1 million and the utility of nothing. The order is the same under the special assumption that the utility function is linear. For example, Behn and Vaupel's (1982, pp. 229–233) argument to ignore sunk costs is based on expected value theory (i.e., a linear utility function). However, expected utility theory in particular and backward induction analysis in general allows the utility function to be nonlinear. In general, sunk cost effects do not imply violations of consequential consistency because the test of consequential consistency requires identical payoffs for the final and isolated choices.

[6] Although it is true that DFT does not predict any violations of consequential consistency as defined in the present experimental paradigm, this fact does not imply that DFT cannot explain various types of sunk costs effects reported in the literature, and a comprehensive explanation of all of the sunk costs findings goes beyond the purpose of this article. As noted earlier, sunk cost experiments do not satisfy the requirements for the test of consequential consistency (see footnote 5).

More specific predictions can be derived from DFT, contingent on assumptions about the approach and avoidance gradients. If the avoidance gradient is steeper than the approach gradient, then preferences should reverse from plans to final decisions (see Equation 1). During planning, when participants are several stages away from the final consequences, the approach component exceeds the avoidance, and participants prefer to take the gamble. Later, when participants reach the final choice and directly face the consequences, the avoidance component exceeds the approach, and participants prefer to take the sure thing. This case is consistent with the results obtained from Experiments 2 and 3, where the Indiana students generally preferred to take the gamble during the planning stage but then switched and preferred to take the sure thing during the final or isolated choice.

If the approach and avoidance gradients are approximately equal, then DFT predicts that preference strength should be attenuated for planned preferences compared with final preferences (see Equation 3). This case is consistent with the results of Experiment 1, where the Purdue students generally produced less extreme preferences for taking the gamble for planned choices compared with their strong preferences favoring the gamble on final or isolated choices.

In short, the observed differences between the participant populations can be explained in terms of individual differences in approach—avoidance gradients for the two groups. The Purdue engineering and management majors might have found the arithmetic problems less aversive than the Indiana University liberal arts majors. It is interesting to note that a cross-cultural study by Weber and Hsee (1999) found that Chinese participants tended to be less risk averse than American participants, which agrees with our finding that the predominantly Asian students from Purdue University were less risk averse than the predominantly American students from Indiana University.

Alternatively, the individual differences between the Purdue and Indiana students also may be explained in terms of a recent refinement of approach–avoidance conflict theory presented by Förster, Higgins, and Idson (1998), which asserts that individuals adopt either a promotion or a prevention focus. In the former case, the approach gradient increases as distance decreases, and in the latter case the avoidance gradient increases as distance decreases. The promotion focus may have predominated in the Purdue population, and the prevention focus may have predominated in the Indiana population. Interestingly, this distinction between promotion and prevention focus is closely related to Lopes' (1987) distinction between potential-mindedness and security-mindedness, respectively.

In summary, predictions derived from DFT about direction of change from planned to final decisions are contingent on assumptions regarding differences in approach–avoidance gradients. However, predictions derived from DFT about rates of change for dynamic and consequential consistency tests are independent of assumptions regarding differences in approach–avoidance gradients. For this reason, the latter tests are emphasized more strongly as providing empirical support for DFT.

Alternative Explanations

Prospect theory (see Kahneman & Tversky, 1979; Tversky & Kahneman, 1981) makes an explicit statement regarding strategic consistency. According to the isolation principle, there should be no difference between planned and isolated choices, and thus strategic consistency should be satisfied. Although Kahneman and Tversky's own studies supported their isolation principle, the present results and those reported by Cubitt et al. (1998) do not. One obvious difference between the present study and the past study by Kahneman and Tversky is that the present study used five-stage decision trees, whereas the past study used two-stage decisions. According to the goal-gradient hypothesis (see Figure 4), reducing the distance (number of stages) should reduce the rate of strategic inconsistency.

Cubitt et al. (1998) suggested that violations of dynamic consistency result from difficulty that decision makers have predicting future preferences. This idea arises from previous research (Kahneman & Snell, 1992; Loewenstein & Adler, 1995) showing that predicted evaluations are inaccurate indicators of experienced values. If a decision maker cannot predict how he or she will eventually feel about a prospect at a later stage, then the planned decision cannot accurately represent the final choice. Cubitt et al.'s explanation is conceptually consistent with the goal-gradient hypothesis (see Figure 4) in the sense that valences anticipated from a distance feel very different than valences experienced close to the point of commitment.

Liberman and Trope (1998) recently proposed that individuals tend to focus more on the desirability of the end states when they are far from the goal, but they change and increase their attention to the feasibility of reaching the desired end state as they get closer to the goal. Applying this idea to the present context, participants might attend more to the amount to win for planned choices, but during the final choice they might attend more to the probability of the winning. Thus, dynamic inconsistency might be caused by a shift in attention from amount to probability of winning.

Earlier, we mentioned that dynamic inconsistency might be caused by temporal inconsistency (see footnote 2 for a distinction). Perhaps participants view the planned choice as a decision between slightly delayed consequences, but they view the final choice as a decision between imminent consequences. However, temporal discounting cannot explain the results of Experiment 3, where the consequences for planned and final choices were delivered after the same amount of time. Nevertheless, temporal inconsistency may play an important role in other contexts (see Hoch & Loewenstein, 1991). For example, Mischel (1974) investigated a delayed-gratification paradigm, where children were given a choice between an immediate small reward or a large delayed reward. In this paradigm, children initially planned to take the large delayed reward, but after suffering the wait for several minutes, they changed their mind, opting instead for the immediate small reward.

It is unnecessary to argue for a single cause producing violations of dynamic consistency, and it is plausible that some combination of the above explanations may be operating, depending on the situation. More research is needed to start dissecting these initial findings and determine the relative importance of each of these various possible causes of dynamic inconsistency. The main goal of the present research was to firmly establish the basic phenomena and provide some theoretical groundwork for guiding future research on this critical topic.

Dynamic Decision-Making Paradigm

A final comment concerns the use of decision trees to test fundamental principles of planning and decision making. Earlier experiments on planning and dynamic decision making (see Brehmer, 1992; Busemeyer, in press; Kerstholt & Raaijmakers, 1997, for reviews), were not designed to test fundamental principles of backward induction analysis. Instead, they were designed to compare human decision performance with optimal performance in a global manner. For example, Sterman (1989) compared the profits earned by human participants with profits earned by an optimal model in a complex task that was designed to simulate marketplace decision behavior. Optimal performance depends on myriad assumptions, and departures from optimality are confounded with numerous alternative explanations (computational errors, memory failures, insufficient learning). These confounds make it impossible to isolate the source of the deviations from optimality in complex dynamic decision tasks, and so these tasks do not permit direct tests of the principles underlying backward induction analysis. What is needed is an experimental paradigm that is sufficiently complex to include the dynamic features of strategic planning and decision making without confounding all the major issues. Decision trees with a small number of stages are ideal for satisfying these requirements. First, decision trees require strategic decision making and planning of action sequences like the more complex dynamic decision tasks. Second, decision trees permit direct tests of basic principles of decision behavior, like earlier single-stage decision tasks. Third, previous research on dynamic decision making (see Rapoport, 1975) indicates that humans may only be capable of planning 2 or 3 stages ahead, and thus decision trees may be useful for understanding planning in more complex dynamic decisions. The success of the reported experiments using decision trees to provide the direct tests of the three principles of backward induction analysis should encourage other researchers to investigate this useful paradigm.

Conclusion

Violations of dynamic consistency pose a serious challenge for decision and problem-solving theories, especially when emotional outcomes are involved. Violations are likely to arise when the emotional state of the decision maker changes as he or she moves through the decision tree. Planned preferences at the beginning of the tree under one emotional state may disagree with final preferences at the end of the tree under a different emotional state. This paradox within the individual is difficult to resolve because there is no way to know which of the two different emotional states should guide the decision-maker's preferences. Referring back to the example in Figure 1, if the woman facing this decision changes from a planned preference to abort to a final preference not to abort, how can one determine which is the best path?

References

Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes, 35,* 124–140.

Barkan, R., & Busemeyer, J. R. (1999). Changing plans: dynamic inconsistency and the effect of experience on the reference point. *Psychological Bulletin and Review, 6,* 547–555.

Behn, R. D., & Vaupel, J. W. (1982). *Quick analysis for busy decision makers.* New York: Basic Books.

Bertsekas, D. P. (1976). *Dynamic programming and stochastic control.* New York: Academic Press.

Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81,* 211–241.

Busemeyer, J. R. (in press). Dynamic decision making. In A. A. J. Marley (Ed.), *International encyclopedia of the social and behavioral sciences: Mathematical and computer models.* Amsterdam: Pergamon Press.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100,* 432–459.

Christensen-Szalanski, J. J. (1984). Discount functions and the measurement of patients' values. Women's decisions during childbirth. *Medical Decision Making, 4,* 47–58.

Clemen, R. (1996). *Making hard decisions.* Pacific Grove, CA: Duxbury Press.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). Dynamic choice and the common ratio effect: An experimental investigation. *The Economic Journal, 108,* 1362–1380.

DeGroot, M. H. (1970). *Optimal statistical decisions.* New York: McGraw-Hill.

Fenz, W. D., & Epstein, S. (1967). Gradients of physiological arousal of experienced and novice parachutists as a function of an approaching jump. *Psychosomatic Medicine, 29,* 33–51.

Förster, J., Higgins, E. T., & Idson, L. C. (1998). Approach and avoidance strength during goal attainment: Regulatory focus and the "goal looms larger" effect. *Journal of Personality and Social Psychology, 75,* 1115–1131.

Goldstein, W. M., & Weber, E. U. (1995). Content and discontent: Indications and implications of domain specificity in preferential decision making. In J. R. Busemeyer, R. Hastie, & D. L. Medin (Eds.), *Decision making from a cognitive perspective. Psychology of learning and motivation* (pp. 83–126). New York: Academic Press.

Hammond, P. J. (1988). Consequentialist foundations for expected utility. *Theory and Decision, 25,* 25–78.

Hoch, S. J., & Loewenstein, G. F. (1991). Time-inconsistent preferences and consumer self-control. *Journal of Consumer Research, 17,* 492–507.

Kahneman, D., & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making, 5,* 187–200.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–291.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs.* New York: Wiley.

Kerstholt, J., & Raaijmakers, J. G. W. (1997). Decision making in dynamic task environments. In R. Ranyard, R. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations* (pp. 205–217). London: Routledge.

Lewin, K. (1935). *A dynamic theory of personality.* New York: McGraw-Hill.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology, 75,* 5–18.

Loewenstein, G., & Adler, D. (1995). A bias in the prediction of tastes. *The Economic Journal, 105,* 929–937.

Lopes, L. L. (1987). Between hope and fear: The psychology of risk. *Advances in Experimental Social Psychology, 20,* 255–295.

Losco, J., & Epstein, S. (1977). Relative steepness of approach and avoidance gradients as a function of magnitude and valence of incentive. *Journal of Abnormal Psychology, 86,* 360–368.

Luce, R. D. (2000). *Utility of gains and losses: Measurement-theoretical and experimental approaches.* Hillsdale, NJ: Erlbaum.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature, 27,* 1622–1668.
Miller, N. E. (1944). Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders* (Vol. 1, pp. 431–465). New York: The Ronald Press.
Mischel, W. (1974). Processes in delay of gratification. *Advances in Experimental Social Psychology, 7,* 249–292.
Nilsson, N. J. (1980). *Principles of artificial intelligence.* Palo Alto, CA: Tioga.
Raiffa, H. (1968). *Decision analysis.* London: Addison-Wesley.
Rapoport, A. (1975). Research paradigms for studying dynamic decision behavior. In D. Wendt & C. Vlek (Eds.), *Utility, probability, and human decision making* (Vol. 11, pp. 347–369). Dordrecht, Holland: Reidel.
Sarin, R., & Wakker, P. (1998). Dynamic choice and non-expected utility. *Journal of Risk and Uncertainty, 17,* 87–119.
Savage, L. J. (1954). *The foundations of statistics.* New York: Wiley.
Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: Nonconsequential reasoning and choice. *Cognitive Psychology, 24,* 449–474.
Sterman, J. D. (1989). Misperception of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes, 43,* 301–335.
Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies, 23,* 165–180.
Thaler, R. (1980). Toward a positive theory of economic choice. *Journal of Economic Behavior and Organizations, 1,* 39–60.
Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters, 8,* 210–207.
Townsend, J. T., & Busemeyer, J. R. (1989). Approach-avoidance: Return to dynamic decision behavior. In C. Izawa (Ed.), *Current issues in cognitive processes: Tulane flowerree symposium on cognition* (pp. 107–133). Hillsdale, NJ: Erlbaum.
Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211,* 453–458.
Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 3,* 305–309.
Von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research.* New York: Cambridge.
Weber, E. U., & Hsee, C. (1999). Models and mosaics: Investigating cross-cultural differences in risk perception and risk preference. *Psychonomic Bulletin and Review, 6,* 611–617.
Weinstein, M. C., & Fineberg, H. V. (1980). *Clinical decision analysis.* Philadelphia, PA: Saunders.

Appendix

Stimulus Orders Used in Experiments

The order of decision trees for the between-trial design are shown in the array below:

*P*10	*I*30	*F*50
*P*20	*I*40	*F*10
*P*30	*I*50	*F*20
*P*40	*I*10	*F*30
*P*50	*I*20	*F*40

The trials progressed across the rows and down the columns: Trial 1 was a planned choice with punishment = 10; Trial 2 was an isolated choice with punishment = 30; Trial 3 was a final choice with a punishment = 50; Trial 4 was a planned choice with a punishment equal to 20, etc. The above table shows only the first half of the trials, and this order was replicated across the remaining 15 trials. Half of the participants received the order shown above, and half of the participants received the reverse of this order.

We formed the pairs used to test consistency by taking one member from the first half and the second member from the second half of the trials. For example, a planned choice for a five-stage tree with a punishment level of 30 problems was presented on Trial 7; 20 trials later, the final choice for exactly the same tree and punishment level was presented on Trial 27, and this pair of trials was used to form a test of dynamic consistency.

The next array shows the order of the decision trees for the within-trial experiments:

*I*40	*P*10	*I*30	*F*50
*I*20	*P*40	*I*10	*F*30
*I*50	*P*20	*I*40	*F*10
*I*30	*P*50	*I*20	*F*40
*I*10	*P*30	*I*50	*F*20

Once again, the trials progressed across the columns and down the rows. For example, Trial 1 was an isolated choice with punishment = 40; Trial 2 was a plan–plan choice with punishment = 10; Trial 3 was an isolated choice with punishment = 30; Trial 4 was a plan–final choice with punishment = 50, etc. Half of the participants received the order shown above, and half of the participants received the reverse of this order.

Received September 15, 1998
Revision received December 30, 1999
Accepted January 7, 2000 ■