Real-Life Decision Making in College Students I: Consistency Across Specific Decisions

KATHLEEN M. GALOTTI, HILLARY J. D. WIENER, and JANE M. TANDLER
Carleton College

First-year undergraduates participated in a short-term longitudinal study of real-life decision making over their first 14 months of college. They were surveyed about 7 different decisions: choosing courses for an upcoming term (3 different terms), choosing an academic major (twice), planning for the upcoming summer, and planning for sophomore-year housing. Participants showed moderate levels of consistency in the options they considered and in the criteria they used to decide between options, with about half of the options or criteria being used at 2 different points on the decision repeatedly studied. Participants varied somewhat in structural consistency, the tendency to consider the same number of options or criteria across decisions. They also varied in the way they integrated information across decision-making tasks. We suggest that people attempt to keep the information demands of the task within workable limits, sometimes sacrificing consistency as a result.

Katy had no idea which college she wanted to attend. She carefully investigated 15 liberal arts colleges, comparing them on everything from financial aid to the quality of the cafeteria food. Does this suggest that she will be equally methodical in choosing a major or choosing a summer job? Previous research is not enlightening on this subject; little research has examined how consistent people are across different real-life decisions. Our work explores the issue of cross-situational consistency (Diener & Larsen, 2009; Mischel, 1968) in real-life decision making by contrasting the adaptive decision maker hypothesis (Bettman, Johnson, Luce, & Payne, 1993; Payne, Bettman, & Johnson, 1988, 1993) with predictions derived from the literature on decision-making styles (Cacioppo & Petty, 1982; Rayner & Riding, 1997; Sternberg, 1997).

Real-Life Decision Making

Much of the literature on decision making concerns decisions made hypothetically on self-contained laboratory tasks (e.g., Beach, 1993, 1998; Payne, 1976; Payne, Bettman, Coupey, & Johnson, 1992; Tversky & Kahneman, 1974, 1981). These studies assume that findings from the laboratory will translate directly to real-world decision making; however, there is good reason to be suspicious of this claim (Galotti, 1989, 2002, 2007). Simulated, hypothetical decision making

American Journal of Psychology
Spring 2014, Vol. 127, No. 1 pp. 19–31 • © 2014 by the Board of Trustees of the University of Illinois

as investigated in laboratories probably excludes activities central to real-life decision making such as clarifying goals, gathering information, and weighing the relative importance of different criteria. As Rozin (2006) points out, neglecting real-life contexts of mental processes such as decision making leaves an important gap in the literature.

Existing studies of real-life decision making often center on a single decision. Ellen Peters and her colleagues have studied consumer decision making as well as dieting and other health-related decisions (Cole et al., 2008; Hibbard & Peters, 2003; Lipkus, Peters, Kimmick, Litocheva, & Marcom, 2010; Peters, 2009). Kmett, Arkes, and Jones (1999) studied high school students choosing a college. Choice of an educational program has also been studied by Schindler and Tomasik (2010) and by Svenson and Jakobsson (2010). Perneger, Charvet-Bérard, and Perrier (2008) studied retrospective reports of medical decision making. Pennington and Hastie (1990) studied influences on juror decision making (often with mock jurors). Lenton and Francesconi (2010) examined decisions on dating made by actual speed daters. However, almost all these studies examined only one decision, leaving unanswered the question of whether the decision makers would behave similarly while making other important decisions.

However, there is a body of existing research on *repeated* real-life decision making involving the study of experts. For example, the Recognition-Primed Decision Making (RPDM) model has been constructed to describe decision making by people with many years of experience, making online decisions under stressful conditions (Klein, 1998, 2011; Zsambok & Klein, 1997). This model shows that experts such as fire commanders, nurses in neonatal intensive care units, airline pilots, and military commanders are unlikely to consider multiple options at a time. Instead, they seem to quickly categorize a situation, even a novel one, as an example of a pattern or prototype. They tend to recognize the kind of situation they are dealing with and implement the applicable solution from memory.

It is risky to apply the RPDM model to other, more common types of decision making because RPDM is based on the premise that the decision maker has a stored mental library of examples from which to draw when they face an instance of decision making. By definition, nonexperts lack extensive experience encountering and cataloging examples.

Past research in our lab has examined how nonexpert real-life decision making changes over time for a single important and possibly life-framing choice, such as how high schoolers' college choice process changes from the end of their junior year to the end of their senior year (Galotti, 1995a, 1995b; Galotti & Kozberg, 1996; Galotti & Mark, 1994). We have also studied college students selecting a major (Galotti, 1999; Galotti et al., 2006), pregnant women choosing a birth attendant and birthing options (Galotti, Pierce, Reimer, & Luckner, 2000), and parents settling on a school placement for their child (Galotti & Tinkelenberg, 2009).

Although the specific findings differed slightly from study to study, several common patterns have emerged (Galotti, 2007). Participants are consistently observed to constrain the amount of information they consider to a few options (e.g., possible choices) and to a somewhat larger set of criteria (factors or considerations used to decide between options). Interestingly, although the specific decisions we have studied varied, the range of options and criteria reported stayed within a fairly narrow range: about two to five options and about three to nine criteria. In the longitudinal studies, the number of options considered over time shrank; the number of criteria used did not. People's intuitive "calibration" with the predictions of normative linear models was surprisingly good (average correlation around .70; range, .54 to .90). There were slight effects of education and ability on the amount of information considered. However, each of these studies again looked only at consistency over time for a single decision rather than consistency across different decisions, and it is therefore unknown how well these patterns of behavior would apply to other consequential decisions. To learn this, we would have to study the same group of decision makers confronting a series of decisions over time.

That was the aim of the current study. Two traditions of decision-making research suggest that there will be differing levels of consistency. One tradition emphasizes individual decision-maker characteristics as a constant influence brought to many different decisions, and it predicts a fair amount of consistency across different important choices and over time. Another tradition highlights the specific circumstances

of a particular decision, predicting little consistency in choice behavior from one judgment to another. In essence, these contrasting predictions recapitulate the question of the cross-situational consistency of behavior or affective responses in personality psychology (Diener & Larsen, 2009).

Decision-Making Styles

Decision-making style research suggests that there is a great deal of variability in decision-making preferences between individuals but relative consistency in these preferences within an individual. This assumption explains the observation that different people make different decisions under the same circumstances by positing that different people possess different decision-making styles. These styles are stable individual differences in the ways people approach decision-making (or other cognitive) tasks, including the way they prefer to acquire information, process it, and evaluate it (Cacioppo & Petty, 1982; Rayner & Riding, 1997; Sternberg, 1997). Decision-making styles constitute a subset of broader cognitive styles, defined generally as the way people deploy their intellectual abilities, or the manner in which they approach cognitive tasks (Loo, 2000; Scott & Bruce, 1995).

Decision-making styles are assessed by self-report measures. That is, respondents are asked to describe how they typically make decisions, instead of being observed as they actually face specific decisions (Scott & Bruce, 1995). Note that using such measures presupposes a certain consistency of approach to a variety of decision types. However, as previously mentioned, research has not examined this assumption of consistency, at least for real-life decision making.

It may be that the connection between styles and behavior is weak or nonexistent, even though people think they have a consistent style that guides their behavior. Research suggests this is the case with respect to learning styles: Although people report having distinct preferences, little evidence has been found to support the premise that those preferences affect educational outcomes or experiences (Pashler, McDaniel, Rohrer, & Bjork, 2008).

Decision-Making Strategies

Other research gives reason to doubt that an individual is consistent in how she or he makes a variety of decisions. The adaptive decision maker hypothesis (Bettman et al., 1993; Payne et al., 1988, 1993) suggests that individuals adopt different strategies when they face different decisions. Essentially, this hypothesis holds that individuals' decision making is affected more by specific aspects of a particular decision than by general, preexisting decision-making styles or proclivities.

For example, in a hypothetical apartment choice task (Payne, 1976), people used different strategies when choosing an option from a large choice set than they did when choosing an option from a very small choice set. When faced with only two options, people compared each option on all or most criteria, making trade-offs such as letting a desirable value of one factor (e.g., low rent) trade off against a less desirable value of another (e.g., less closet space). However, when faced with 6 or 12 options, they eliminated some options on the basis of only one or a few dimensions. For instance, they looked first at rent and immediately eliminated all apartments with high rents, without considering trade-offs with other factors. This noncompensatory elimination is a cognitively efficient strategy used when too much information is available.

Payne and Bettman (2004) argue that noncompensatory elimination is just one of many potential strategies, some of which are complex and take a great deal of cognitive effort and others that rely more heavily on intuition and are less mentally demanding. Strategies are usually acquired through experience and training, although some may be hardwired. Strategies differ in their effectiveness for different decisions, depending on a variety of task aspects, the contexts in which decisions are made, social factors, and individual differences between decision makers. The ability to use different strategies in response to different task environments is one hallmark of good decision making.

The adaptive decision maker hypothesis was created and studied only with hypothetical decision making. Thus, it remains to be seen whether decision makers would show a wide range of differences in behavior across different real-life decisions.

What Is Consistency in Real-Life Decision Making?

To study consistency in decision making, we first need to understand what it is, especially in real-life decision making. Some studies of hypothetical decision making have examined consistency, although the

definition and measurement of it have varied. One operationalization of consistency that has been used is an *R* value in a regression equation predicting a participant's judgments from the available cues given in the hypothetical problems (see Ruscio & Stern, 2006). In a study of nurses' clinical decision making, Hughes and Young (1990) measured consistency by correlating nurses' decisions with recommendations of a decision analytic model on various hypothetical decision scenarios. Finucane et al. (2002) examined similarity in appraisal of different hypothetical options as a function of differing evaluation contexts.

In the present study, we will examine two aspects of consistency: content and structural. The adaptive decision maker model and the decision-making style research suggest different levels of structural consistency. *Structural consistency* involves measuring noncontent aspects of different decisions, such as the number of options and criteria considered and the degree to which information about options and criteria is integrated (we will discuss integration later). These can be compared across different decisions to assess whether people who report considering many options on one decision do so for other decisions, for example, and whether people who weigh many criteria for one decision also do so on other decisions.

Another type of consistency is *content consistency,* or the degree to which participants consider the same or similar criteria and options in making similar decisions. That is, do the kinds of things people think about vary a little or a lot over different decisions? Past research in our laboratory has examined this kind of consistency, and we include these measures to facilitate comparisons between the present study and those of previous investigations.

STUDY

We examined both kinds of consistency through a longitudinal study of first-year college students who were followed in four sessions over a 14–month period. At each session, participants filled out various survey instruments. Included among these were instruments concerning several specific and important decisions: choosing courses for an upcoming term, choosing a major, planning for the upcoming summer after their first year, and planning for student housing in the second year.

For each decision, we counted the number of options the participants reported under active consideration, the number of criteria they reported using to decide between these options, and the calibration of the participants' holistic ratings of the overall goodness of options with the predicted ratings of various linear models (described in more detail later). We then examined the degree to which corresponding measures from different decisions were intercorrelated. We predicted that consistency would be higher for similar decisions (e.g., between the academic major decision as assessed on two different occasions; between the decisions about courses for future terms as assessed on two different occasions) than it would for less similar decisions (e.g., between any other set of decisions).

To establish a baseline measure of consistency, we also recruited and ran a sample of 32 first-year and sophomore students who attended a single session and filled out instruments about two decisions (choosing courses for an upcoming term and declaring a major) twice in the same session.

METHOD

Participants

Participants were first-year students at Carleton College. The first cohort, recruited in the fall of 2009, consisted of 101 first-year students (39 male). Students were recruited through a letter of invitation that was sent to all 520 first-year students. Of these 101 students, 94 (35 male) returned for a second session in winter of 2010, 93 (34 male) of those returned for a third session in the spring of 2010, and 88 (33 male) returned for a fourth and final session in the fall of 2010, for an overall retention rate of 87%. Although specific race and ethnicity data were not collected on the sample, the overall population of first-year Carleton students in the fall of 2009 included 22% who self-reported their racial group as African American, Asian American, Hispanic/Latino/Chicano, or Native American (Lawrence, 2010).

The second cohort of participants, who were recruited in the fall of 2010, consisted of 48 first-year students (12 male). Students were recruited through a letter of invitation that was sent to 200 quasirandomly selected first-year students.[1] Of the original 48, 45 (12 male) returned for the second and third sessions, in the winter and spring of 2011, respectively, and 44

(12 male) returned for the fourth session in the fall of 2011, for an overall retention rate of 92%.

A third cohort of students was used to provide baseline comparison data. They included 16 first-year students (8 male) and 16 sophomores (8 male), none of whom had previously participated in the study. These students were recruited in the fall of 2012 and were paid $8 for participating in a single 45–min session.

Materials

A number of different instruments were administered during the different sessions,[2] but only the ones relevant for this article will be described here.

COURSE SCHEDULE WORKSHEET (CSW).

Students were asked to list their options for course schedules for the next term. "Options" were defined as a set of planned courses; each distinct set, even if differing only by one course, was counted as a separate option. Students filled out a facsimile of a registration card for each option under active consideration.

HOUSING OPTIONS WORKSHEET (HOW).

Students were asked to list their options for housing for the next year. Each option was defined by a unique location (e.g., Davis, Burton), room type (e.g., double, triple), and potential roommates (given by initials). Participants were instructed that any change in any of these elements defined a separate option (thus, "Burton double with A.B." would constitute a different option than "Burton double with M.E.").

FACTORS AND OPTIONS WORKSHEET (FAOW).

This instrument, adapted from previous research (Galotti, 1999, 2007; Galotti & Tinkelenberg, 2009) was used to provide a systematic way for participants to describe the options under active consideration and the criteria they used to evaluate those options. The worksheet consisted of a grid containing 10 columns of blanks. In the second column participants were asked to list the criteria by which they were currently evaluating their options. Each criterion was rated for its importance on a scale from 1 (*Not very important*) to 10 (*Extremely important*), and these weights were placed in the first column. At the top of the third through tenth columns, the options under active consideration were listed. Participants rated the options based on how well they fulfilled each of these criteria using a 10–point scale. Table 1 provides a fictitious example of a filled-out version of this instrument for choosing a major.

OVERALL RATING (OVRAT).

This measure was used only in Sessions 2, 3, and 4. Participants were shown a typed list of the options they had listed in the previous session for a specific decision (i.e., choosing courses, choosing a major, choosing housing, or choosing summer plans). They rated each option on a scale of 0 to 10, where higher numbers indicated greater overall goodness of the option in light of the student's current goals and objectives.

Procedure

Participants were run in small groups by trained undergraduate research assistants. Participants filled

TABLE 1. Fictitious Example of a Completed Factors and Options Worksheet for Choosing a College Major

Factor	Importance weight	Option 1: psychology	Option 2: biology	Option 3: English
Average class size	9	8	2	6
Professor quality	7	7	10	5
Major requirements	10	5	2	8
Career opportunities	6	6	10	2
Laboratory requirements	3	7	3	9
Predicted value calculated from				
Top criterion model		5	2	8
Equally weighted criteria model		33	27	30
Full multiattribute utility theory model		228	177	208

out a CSW and an FAOW for choosing courses for winter term during the first (fall, first year) session. Research assistants read a detailed set of instructions for each instrument and walked participants through the process of filling out the grid. In the second (winter) session, participants filled out an OVRAT for winter term courses and filled out a CSW and an FAOW for spring term courses as well as an FAOW for declaring a college major (at Carleton College, majors are not formally declared before the sixth or spring term in the sophomore year).

In the third (spring term) session, participants filled out an OVRAT for spring term courses and an OVRAT for the academic major decision and filled out a CSW and an FAOW for fall term courses, an HOW, and FAOWs for sophomore year housing and their summer plans (the latter could include internships, jobs, travel, study, down time, or various combinations of these). Finally, in their fourth session (fall term of sophomore year), participants filled out OVRATs for spring term courses, housing, and summer plans, as well as an FAOW and an OVRAT for their academic major declaration decision.

Worksheets to list options were necessary for housing and course selection decisions because the definition of "options" was a little more complex for these. For example, two schedules that both contained French 101 and Economics 253 but differed in the third course were treated as two distinct options, as were housing options of the same type of room in a specific dorm but with different potential roommates. Listing options for summer plans and for the academic major was more straightforward and thus did not require a specific worksheet before filling out the FAOW. We emphasized to participants that all specific options listed should be distinct. Therefore, an option such as "Religion" for the academic major decision was distinct and to be listed separately from an option such as "Religion and Political Science." Similarly, for the summer plans decision, an option such as "volunteer" was to be treated as distinct from "volunteer and travel."

The baseline consistency cohort attended a single session in which they filled out instruments (CSW, FAOW courses, OVRAT courses, FAOW major, OVRAT major) twice in a single session. All instruments for one decision (e.g., CSW, FAOW, OVRAT) were completed before moving on to the second decision. Then, the first decision was repeated, followed by a repetition of the second decision. The order of decisions (courses, major) was counterbalanced across each group of participants (i.e., first-year men, first-year women, sophomore men, sophomore women).

RESULTS

Content Consistency

Content consistency is a measure of the degree to which a person considers the same criteria and options for different decisions. Because content consistency is context dependent, we assessed it only between similar decisions. In particular, we looked at content consistency within two kinds of decisions: the same decision assessed at two points in time (specifically, the academic major decision) and two similar decisions assessed at two points in time (specifically, the class choice decision for two different terms).

We begin with a discussion of the criteria participants used to make a decision. Recall that participants listed criteria (standards for choosing between options) in free response format. We computed the consistency of criteria by counting the number of common criteria listed for two decision instruments (e.g., for academic major as assessed in Session 2 and for academic major as assessed in Session 4) divided by the average number of criteria listed in the two sessions. The resulting number is the percentage of criteria listed twice.

Table 2 presents the mean overlap in content for similar decisions and shows that criteria content consistency is moderate and roughly equivalent for the same decision made at two different times as it is for two similar decisions made at different times. As shown, the average overlap between the criteria listed for academic major decisions was a little over 50%, and the overlap for criteria listed for the course choice decisions averaged about 54%. These values are similar to the content consistency measures computed in a previous longitudinal study for high school students choosing a college (Galotti, 1995a) and another previous longitudinal study of college students choosing a major (Galotti, 1999). In comparison, the figures generated from the baseline comparison control group were 90% for the academic major decision and 92% for the academic courses decision. (Recall that these figures come from the participants who

TABLE 2. Criteria Listed in Common for Different Decisions

Decision	% overlap[a]
Academic major (Sessions 2 and 4)	51.96
Academic major (baseline comparison, single session)	89.85
Courses for winter and courses for spring	54.30
Courses for spring and courses for fall	57.88
Courses for winter and courses for fall (next year)	49.94
Courses for upcoming term (baseline comparison, single session)	92.06
Academic major (Session 2) and courses for winter	27.81
Academic major (Session 2) and courses for spring	30.33
Academic major (Session 2) and courses for fall (next year)	28.21
Academic major (Session 4) and courses for winter (previous year)	27.04
Academic major (Session 4) and courses for spring (previous year)	28.69
Academic major (Session 4) and courses for fall (same year)	31.80

[a]Computing the percentage of criteria listed for both decisions, with the average number listed in the denominator and the number of criteria listed on both occasions or for both decisions in the numerator.

filled out instruments about the same decisions twice in the same session.)

We also looked at criteria consistency between the related but distinct decisions of academic major and course choice at a number of different times. Table 2 shows that these mean percentage overlaps are smaller in magnitude, averaging just over 30%.

In Table 3 we present data on the consistency of options considered. We found a 61.33% overlap between the potential academic majors that participants listed on the two occasions they were asked about their academic majors. Once again, this comports well with results found in previous longitudinal studies of students choosing a college or major

TABLE 3. Options or Similarity Ratings for Options Listed in Different Decisions

Decisions	% overlap
Academic major (Sessions 2 and 4)	61.33[a]
Academic major (baseline comparison, single session)	98.87
Courses for upcoming term (baseline comparison, single session)	80.04
	Rated similarity (1–10 scale)[b]
Courses for winter and courses for spring	5.88
Courses for spring and courses for fall	5.74
Courses for winter and courses for fall (next year)	4.36

[a]Computing the percentage of options listed for both decisions, with the average number listed in the denominator and the number of options listed on both occasions or for both decisions in the numerator.
*[b]Because course decisions for different terms necessarily listed different specific courses, we rated the overall similarity between the options considered at each session on a scale of 0 (*no similarity*) to 10 (*maximum similarity*).*

(Galotti, 1995a, 1999). The corresponding figure from the baseline comparison group is 99%. For the course decisions, we did not expect that students would ever list exactly the same options, because the decisions were in reference to different upcoming terms when different courses were offered. Thus, we could not meaningfully compute a percentage overlap. We did look at the exact overlap from the baseline comparison group, finding it to be 80% (for the number of options listed for the courses decision twice in the same session).

For the main group of participants, trained undergraduate raters rated the mean similarity between the sets of options for different course selection decisions on a 10-point scale (10, *maximum similarity;* 0, *no similarity*). Rated similarity was highest for shorter time periods (surveys performed one term apart) and lower for surveys 6 months apart.

Taken together, the results suggest a moderate degree of content consistency. Across a short (3–8 month) interval, participants list up to about half of the same criteria and slightly more of the same options when making the same or similar decisions. This figure is a little higher when participants are asked twice about the same decision (choosing an academic major) than when they are asked about two similar decisions (choosing an academic major and choosing courses for an upcoming term); however, values of content consistency are always in the moderate range.

Structural Consistency

Before determining structural consistency, we first measured the structure—or noncontent features—of the decision with two sets of structural measures. The first included the number of options considered for each decision, the number of criteria considered for each decision, and the "decision map" size for each decision. This last measure was simply the number of options multiplied by the number of criteria, which indexed the number of pieces of information a person considered when making a decision. We present the means of these measures for each of the seven decisions surveyed in Table 4.

A second set of structural consistency measures assessed how participants integrated their various thoughts about a decision. To do this, we made use of so-called calibration coefficients, correlations of participants' overall ratings of options (from the OVRAT measure) and the predicted values of those options by various linear models of decision making (Dawes, 1982; Dawes & Corrigan, 1974), based on how participants rated each option on the criteria on the corresponding FAOW measure. Three linear

TABLE 4. Options, Criteria, and Decision Map Size for Decisions

Decision	Mean number of options	Mean number of criteria	Decision map size[a]
Choosing courses for 2nd term[b]	4.75	5.87	29.08
Choosing courses for 3rd term[b]	4.44	5.68	25.36
Choosing academic major, 1st year[c]	4.85	4.81	24.21
Choosing courses for 4th term[b]	4.90	5.78	28.20
Choosing housing for 2nd year	4.55	4.77	22.72
Choosing summer plans after 1st year	4.15	4.79	25.05
Choosing academic major, 2nd year[d]	3.75	5.08	19.41
Choosing courses (baseline comparison)	4.16	5.00	20.80
Choosing academic major (baseline comparison)	3.58	5.11	18.29

[a]Equal to the number of options multiplied by the number of criteria.
[b]All course selections were made 1 term before enrollment. This college has 3 terms each academic year. Data were collected approximately 2 weeks before final registration.
[c]Final declaration of academic major is made in the 3rd term of the 2nd year. These data were collected in the 2nd term of the 1st year.
[d]Final declaration of academic major is made in the 3rd term of the 2nd year. These data were collected in the 1st term of the 2nd year.

models, which we have used extensively in previous research (e.g., Galotti, 1995a, 1995b, 1999, 2007; Galotti et al., 2000, 2006; Galotti & Tinkelenberg, 2009), were considered:

> *Top criterion model.* Predicted values under this very simple model were calculated by using only the ratings on the criterion to which a participant had given the highest importance weighting. If she or he had given more than one criterion the highest weight, then the ratings on all these criteria were averaged to calculate the predicted value of that potential option.
>
> *Equally weighted criteria model.* Here, the predicted value for each option was computed by giving each criterion equal weight (i.e., ignoring the participant's own importance weights). The subjective ratings of each option on each criterion were summed and correlated with the participants' overall holistic ratings of each option.
>
> *Multiattribute utility theory.* This model incorporated all the information a participant provided on an FAOW. The importance weight of a given criterion was multiplied by the subjective rating of each option on each criterion, and these products were summed over all criteria listed. Thus, each potential option received a summary score: a predicted value of its goodness of fit with the decision maker's criteria and appraisal of options with respect to each criterion.

The bottom three rows of Table 1 present an example of the predicted value calculations for these three linear models. These scores were then correlated with the participants' overall impression (holistic) ratings of each option, as given on the OVRAT instruments. Positive correlations indicate better calibration with the predictions of the different linear models. In Table 5 we present the means of these calibration coefficients for each of the seven decisions surveyed.

In order to measure overall consistency, we subjected all sets of structural consistency measures to an internal reliability analysis, across the seven decisions. If participants were extremely consistent across decisions in the amount of information considered (number of options, number of criteria, decision map size) or if they integrated that information with their overall holistic impressions (calibration coefficients) similarly across decisions, then the internal reliability analyses should yield high values of coefficient alpha, ideally, higher than .80 (values of coefficient alpha range from 0 to 1.00, with reasonable values for coherent scales thought to be greater than .70; Constantine & Ponterotto, 2006).

Table 6 presents the values of coefficient α for the six structural measures in the first column. Note

TABLE 5. Calibration Coefficients for Decisions

Decision	Top criterion model	Equally weighted criteria model	Multiattribute utility theory model
Choosing courses for 2nd term[a]	.17	.12	.22
Choosing courses for 3rd term[a]	.32	.27	.19
Choosing academic major, 1st year[b]	.44	.39	.41
Choosing courses for 4th term[a]	.20	.23	.35
Choosing housing for 2nd year	.35	.44	.47
Choosing summer plans after 1st year	.31	.32	.36
Choosing academic major, 2nd year[c]	.65	.65	.43
Choosing courses (baseline comparison)[d]	.34	.54	.55
Choosing academic major (baseline comparison)[d]	.68	.67	.74

[a]All course selections were made 1 term before enrollment. The college has 3 terms each academic year. Data were collected 2–3 weeks before final registration.
[b]Final declaration of academic major is made in the 3rd term of the 2nd year. These data were collected in the 2nd term of the 1st year.
[c]Final declaration of academic major is made in the 3rd term of the 2nd year. These data were collected in the 1st term of the 2nd year.
[d]Entries in the row reflect the mean calibration coefficient computed for the 2 repetitions of this decision.

TABLE 6. Mean Internal Reliability Across Performance Measures

Dependent measure	Internal reliability (coefficient α)	Mean interitem correlation
Number of options	.60	.20
Number of criteria	.78	.34
Decision map size	.70	.28
Calibration coefficient, top criterion model	.28	.06
Calibration coefficient, equally weighted criteria model	.21	.03
Calibration coefficient, multiattribute utility theory model	.29	.05

that two of the three measures of the amount of information considered—number of criteria (α = .78) and decision map size (α = .70)—showed reasonable consistency. However, participants did not appear to integrate information similarly across decisions at all (all αs < .30). The third column of this table presents the average correlation between all pairs of corresponding structural measures, and the range is from .06 to .34. This suggests that there is moderate structural consistency in how much information is considered but little or none in how it is integrated.

DISCUSSION

We found evidence for moderate levels of content consistency and varying levels of structural consistency across multiple important real-life decisions. The moderate level of content consistency across both the criteria and the options listed suggests that although people do repeatedly consider some of the same issues when making an important decision (probably because that information is valued by the decision maker), the set of choices under consideration is not static over time. Put more simply, people's thinking about a decision undergoes a fair degree of modification within the span of a few months.

Structural consistency presented a more complex picture; different measures of structural consistency yielded very different impressions of the degree to which people's thinking changes. Calibration coefficients—measuring consistency of people's intuitions with predictions derived from linear models—did not show much consistency across decisions. That is, students who were highly calibrated on one decision were not necessarily highly calibrated on others. However, structural consistency was high for two performance measures: the size of the decision map and the number of criteria. The consistency in decision map size suggests that people may have a general preference for considering either a lot or a little information in making a decision, but a deeper investigation reveals that this is true only for the number of criteria considered. In contrast, people are far less consistent in the number of options they consider across decisions.

One possible explanation for this difference between options and criteria is that for the specific decisions we studied, it may have been more difficult for our undergraduates to maintain consistency in the number of options considered than it would be to maintain consistency in the number of criteria listed. Given that no participants were under time pressure, it would be easy for them to conform to their preferences for listing many or few criteria. However, differences in the specific decisions we investigated made it helpful for students to be able to expand or contract the consideration set size. For example, for the class choice decision, even small changes in a proposed set of courses constituted a new option, making it easy for the number of options to accumulate quickly. However, for the major or summer plans decisions, there are likely to be fewer options or combinations available, restricting the number of options to consider. Students appeared to attempt to maintain some degree of consistency when it was easy to do so, but when it became burdensome they adapted and changed the structure of the decision to fit its unique characteristics. This flexibility in performance from one decision to the next is the hallmark of the adaptive decision maker hypothesis (Payne, Bettman, & Johnson, 1988, 1993), which holds that decision makers are heavily influenced by specific features of the decisions in front of them when they engage in making choices.

It is worth noting that the term *consistency* has several possible meanings, only some of which are captured in our measures. For example, we did not have direct measures of strategy use, or of advice seeking, the kinds of information people sought. Instead,

we focused on measuring content and structural consistency across a series of connected yet distinct decisions that were presumably of at least moderate importance to the decision makers, and that occurred on a known and imminent schedule.

A second issue is whether our task artifactually enhanced consistency. A reviewer of an earlier version of this manuscript wondered whether asking participants to adopt the analytic approach invited by the FAOW instrument led to an unnatural mode of responding that inflated, in particular, the calibration coefficients and possibly caused an overreporting of the number of options and criteria.

Although we agree that the FAOW sheets lead a participant to structure information about their decision making analytically, we do not believe that it fundamentally changes the process. In a previous study of parents making school choice decisions (Galotti & Tinkelenberg, 2009), experimenters first interviewed participants in a very open-ended way, using those interviews to fill out an FAOW instrument for the parent. The ranges of options and criteria listed, along with the magnitude of calibration coefficients from that study, are very much in line with other studies, including this one, in which we ask participants to fill out the FAOW themselves.

We have also looked to see whether filling out an FAOW instrument changes one's approach to decision making. An early study (Galotti, 1995a) compared the decision making of a core group of 90 high school students who filled out an FAOW instrument on four different occasions with that of a control group of high school students who filled out the instrument only once. We found no differences in any of the dependent measures (number of options, criteria, or calibration coefficients). Indeed, we have found (Galotti, 2007) very similar values in all the performance measures across a variety of decisions and populations of decision makers, and those values are consistent with the ones found here.

Another issue raised by the reviewer is whether the specific decisions we included constrained the number of options or criteria (or both) that participants *could* consider. Although at first blush that worry might seem plausible, there is evidence against its use as an explanation for our data. For example, with the major decision, the college offers 37 regular majors, plus a number of special majors available by petition. None of our participants listed even 15 options under consideration. In terms of the course decision, any unique combination of courses counted as a separate option, and there are literally millions of combinations (an average of 250 courses a term are offered, and students typically choose 3), not including physical education or music courses (which many students take). For the housing decision, there are more than 900 rooms available, although an option for this decision was defined as a room choice with a selection of roommates, thus multiplying the number. We are less sure about the number of options for the summer plans decision, or the number of possible criteria one could list for any of the decisions, but note that the participants' responses on these measures were in a very narrow range (and within the range of responses given by other participants to other decisions studied in our lab). And, as a comparison, the maximum number of options listed ranged from 8 to 12, with no apparent relationship to the number of possible options available, and the maximum number of criteria listed by any participant was between 9 and 12.

Interestingly, our research may also shed some new light on what exactly it means to be an adaptive decision maker. In particular it suggests that for many real-life decisions, maintaining consistency can be an important part of making an adaptive decision. It is easier and more efficient to consider information a second time than it is to consider new information, and keeping the *number* of options or criteria under active consideration at any point consistent allows a person to use a similar decision-making strategy across different decisions. However, for situations when an attempt to maintain consistency becomes a hindrance, not a help, people are adaptive and therefore able to drop it and respond to specific features of a particular decision. Of course, this speculative explanation needs further test.

Going forward, we need to understand real-life decision making across an even wider set of decisions and among a more diverse population of decision makers. After all, these decisions were important and in some cases potentially life framing; the decision makers were highly educated and motivated. We note that generating criteria and constructing the short list of options are both tasks that typically are not part of laboratory decision-making tasks. These processes are important parts of what it means to be a good or

efficient real-life decision maker, and any account of real-life decision making must provide good descriptions of these processes before embedding these descriptions in a fully elaborated theoretical context. We hope our work contributes to an initial understanding of these very important steps.

NOTES

Support was provided through a grant from the Howard Hughes Medical Institute to Carleton College (grant no. 52006286) and faculty endowment grants from the Dean of the College office, Carleton College, to the first author. We thank Lacey Dorman, Amy Murdoch, Samantha Drogue, Stephanie Moberg, Christina Veatch, Andrea Rockwood, Kelsey Han, Lauren Hickman, Tasha Rhoads, Kelley Stevens, Johanna Williams, and Naomi Fine for assistance with data collection and analysis and Robert Proctor, John Ruscio, and an anonymous reviewer for their comments on an earlier draft of the manuscript.

Hillary Wiener is now at the Fuqua School of Business at Duke University.

Address correspondence about this article to Kathleen Galotti, Director of Cognitive Science, Carleton College, 1 North College Street, Northfield, MN 55057 (E-mail: kgalotti@carleton.edu).

1. Students who opted to remove their names from the college's public directory were not sent invitations. Moreover, international students were excluded from invitations because their visa status made compensation an issue. Of the remaining students, invitations were sent to a randomly selected subset.

2. For a complete list or copies of each instrument used, contact the first author.

REFERENCES

Beach, L. R. (1993). Broadening the definition of decision making: The role of prechoice screening of options. *Psychological Science, 4,* 215–220.

Beach, L. R. (Ed.). (1998). *Image theory: Theoretical and empirical foundations.* Mahwah, NJ: Erlbaum.

Bettman, J. R., Johnson, E. J., Luce, M. F., & Payne, J. W. (1993). Correlation, conflict, and choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 931–951.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42,* 116–131.

Cole, C., Laurent, G., Drolet, A., Ebert, J., Gutchess, A., Lambert-Pandraud, R., & Peters, E. (2008). Decision making and brand choice by older consumers. *Marketing Letters, 19,* 355–365.

Constantine, M. G., & Ponterotto, J. G. (2006). Evaluating and selecting psychological measures for research purposes. In F. T. L. Leong & J. T. Austin (Eds.), *The psychology research handbook* (pp. 104–113). Thousand Oaks, CA: Sage.

Dawes, R. M. (1982). The robust beauty of improper linear models in decision-making. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: heuristics and biases* (pp. 391–407). New York, NY: Cambridge University Press.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision-making. *Psychological Bulletin, 81,* 95–106.

Diener, E., & Larsen, R. J. (2009). Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses. In E. Diener (Ed.), *The collected works of Ed Diener* (pp. 7–24). New York, NY: Springer.

Finucane, M. L., Slovic, P., Hibbard, J. H., Peters, E., Mertz, C. K., & MacGregor, D. G. (2002). Aging and decision-making competence: An analysis of comprehension and consistency skills in older versus younger adults considering health-plan options. *Journal of Behavioral Decision Making, 15,* 141–164.

Galotti, K. M. (1989). Approaches to studying formal and everyday reasoning. *Psychological Bulletin, 105,* 331–351.

Galotti, K. M. (1995a). A longitudinal study of real-life decision making: Choosing a college. *Applied Cognitive Psychology, 9,* 459–484.

Galotti, K. M. (1995b). Memories of a "decision-map": Recall of a real-life decision. *Applied Cognitive Psychology, 9,* 307–319.

Galotti, K. M. (1999). Making a "major" real life decision: College students choosing an academic major. *Journal of Educational Psychology, 91,* 379–387.

Galotti, K. M. (2002). *Making decisions that matter: How people face important life choices.* Mahwah, NJ: Erlbaum.

Galotti, K. M. (2007). Decision structuring in important real-life choices. *Psychological Science, 18,* 320–325.

Galotti, K. M., Ciner, E., Altenbaumer, H. A., Geerts, H. J., Rupp, A., & Woulfe, J. M. (2006). Making a "major" life-framing decision: Individual differences in performance and affective reactions. *Personality and Individual Differences, 41,* 629–639.

Galotti, K. M., & Kozberg, S. F. (1996). Adolescents' experience of a life-framing decision. *Journal of Youth and Adolescence, 25,* 3–16.

Galotti, K. M., & Mark, M. C. (1994). How do high school students structure an important life decision? A short-term longitudinal study of the college decision-making process. *Research in Higher Education, 35,* 589–607.

Galotti, K. M., Pierce, B., Reimer, R. L., & Luckner, A. E. (2000). Midwife or doctor: A study of pregnant women making childbirth decisions. *Journal of Midwifery & Women's Health, 45,* 320–329.

Galotti, K. M., & Tinkelenberg, C. E. (2009). Real-life decision making: Parents choosing a first-grade placement. *American Journal of Psychology, 122,* 455–468.

Hibbard, J. H., & Peters, E. (2003). Supporting informed

consumer health care decisions: Data presentation approaches that facilitate the use of information in choice. *Annual Review of Public Health, 24,* 413–433.

Hughes, K. K., & Young, W. B. (1990). The relationship between task complexity and decision-making consistency. *Research in Nursing & Health, 13,* 189–197.

Klein, G. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Klein, G. (2011). *Streetlights and shadows: Searching for the keys to adaptive decision making.* Cambridge, MA: MIT Press.

Kmett, C. M., Arkes, H. R., & Jones, S. K. (1999). The influence of decision aids on high school students' satisfaction with their college choice decision. *Personality and Social Psychology Bulletin, 25,* 1293–1301.

Lawrence, J. (2010). *Class of 2013 profile: Highlights*. Retrieved from http://apps.carleton.edu/admissions/apply/requirements/PastProfiles/2013profile/highlights/

Lenton, A. P., & Francesconi, M. (2010). How humans cognitive manage an abundance of mate options. *Psychological Science, 21,* 528–533.

Lipkus, I. M., Peters, E., Kimmick, G., Liotcheva, V., & Marcom, P. (2010). Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: The influence of numeracy. *Medical Decision Making, 30,* 464–473.

Loo, R. (2000). A psychometric evaluation of the General Decision-Making Style Inventory. *Personality and Individual Differences, 29,* 895–905.

Mischel, W. (1968). *Personality and assessment*. New York, NY: Wiley.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9,* 105–119.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16,* 366–387.

Payne, J. W., & Bettman, J. R. (2004). Walking with the scarecrow: The information-processing approach to decision research. In D. K. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 110–132). Malden, MA: Blackwell.

Payne, J. W., Bettman, J. R., Coupey, E., & Johnson, E. J. (1992). A constructive process view of decision making: Multiple strategies in judgement and choice. *Acta Psychologica, 80,* 107–141.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 534–552.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker.* Cambridge, England: Cambridge University Press.

Pennington, N., & Hastie, R. (1990). Practical implications of psychological research on juror and jury decision making. *Personality and Social Psychology Bulletin, 16,* 90–105.

Perneger, T. V., Charvet-Bérard, A., & Perrier, A. (2008). Patient assessments of the most important medical decision during a hospitalization. *Journal of General Internal Medicine, 23,* 1659–1665.

Peters, E. (2009). A perspective on eating behaviors from the field of judgment and decision making. *Annals of Behavioral Medicine, 38,* S81–S87.

Rayner, S., & Riding, R. J. (1997). Towards a categorisation of cognitive styles and learning styles. *Educational Psychology, 17,* 5–27.

Rozin, P. (2006). Domain denigration and process preference in academic psychology. *Perspectives on Psychological Science, 1,* 365–376.

Ruscio, J., & Stern, A. R. (2006). The consistency and accuracy of holistic judgment: Clinical decision making with a minimally complex task. *The Scientific Review of Mental Health Practice, 4,* 52–65.

Schindler, I., & Tomasik, M. J. (2010). Life choices well made: How selective control strategies relate to career and partner decision processes. *Motivation and Emotion, 34,* 168–183.

Scott, S. G., & Bruce, R. A. (1995). Decision-making style: The development and assessment of a new measure. *Educational and Psychological Measurement, 55,* 818–831.

Sternberg, R. J. (1997). *Thinking styles.* Cambridge, England: Cambridge University Press.

Svenson, O., & Jakobsson, M. (2010). Creating coherence in real-life decision processes: Reasons, differentiation and consolidation. *Scandinavian Journal of Psychology, 51,* 93–102.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211,* 453–458.

Zsambok C. E., & Klein, G. (Eds.). (1997). *Naturalistic decision making*. Mahwah, NJ: Erlbaum.