Running Head: BACKUP PLAN PARADOX

# First Evidence for "The Backup Plan Paradox"

Christopher M. Napolitano[1] and Alexandra M. Freund[1,2]

[1]University of Zurich, Dept. of Psychology, [2] University Research Priority Program Dynamics of Healthy Aging, University of Zurich

**WORD COUNT: 11,377**

## Author Note

Christopher M. Napolitano and Alexandra M. Freund, Department of Psychology, University of Zurich, Switzerland. Alexandra M. Freund is also affiliated with the University Research Priority Program Dynamics of Healthy Aging, University of Zurich.

This research is supported by the Swiss National Science Foundation Grant 100014-138045 "Readiness to Disengage from Motivational Conflict: Adult Age Differences and Implications for Emotional Experiences" (PI: A. M. Freund) and the Suzanne and Hans Biäsch for Foundation for Applied Psychology Grant 2014/14 "Time for Plan B? Testing for age-related differences in the use, utility, and efficiency of backup plans" (PI: C. M. Napolitano). We thank the members of the Life-Management lab for input into this research, and Zita Mayer, Michaela Wirth and Enrique Wintsch for programming and data collection. Studies 3 and 5 were presented in preliminary form at the June 2016 University of Cologne Self-Regulation Meeting. The second author presented this talk, titled "The Potentially Undermining Effects of Backup Plans."

Correspondence concerning this article should be addressed to either author at Department of Psychology, University of Zurich, Binzmuehlestrasse 14/11, 8050 Zurich, Switzerland. E-mail: napolitano@psychologie.uzh.ch or freund@psychologie.uzh.ch

## Abstract

This research is a first test of the *backup plan paradox.* We hypothesized that investing in a backup plan may facilitate the conditions that it was developed to address: Plan A's insufficiency. Five studies provide initial, primarily correlative support for the undermining effect of investing in a backup plan. Study 1 ($N = 160$) demonstrated that the more participants perceived they had invested in developing a backup plan (preparing a "crib sheet"), the more likely they were to use it, although greater investments were unrelated to backup plan utility. Studies 2 – 4 used a simulated negotiation task. Study 2 ($N = 247$) demonstrated that when goal-relevant resources are limited, investing in developing backup plans and perceiving them as highly instrumental can decrease goal performance through the indirect effect of increased means replacing. Study 3 ($N = 248$) replicated this effect when goal-relevant resources were plentiful. Study 4 ($N = 204$) used an experimental variant of the simulated negotiation task, and demonstrated that simply having a backup plan is not detrimental, but perceiving backup plans to be highly instrumental decreased goal performance, again through the indirect effect of increased means replacing. Study 5 ($N = 160$) replicated findings from Studies 1 – 4 using a lab-based motor task (throwing a ball). Together, these results provide a preliminary indication that backup plans can introduce costs that may jeopardize goal performance.



Keywords: achievement; motivation; self-regulation; backup plans; goal pursuit

Words: 227

First Evidence for the Backup Plan Paradox

Imagine a post-doc at a lectern before she gives her "job talk." She has spent weeks preparing, and hopes to present her research extemporaneously. If her nerves fail, she can use her backup plan, a stack of presentation notes that she spent all week highlighting.

All goal pursuits, including job talks, are marked by uncertain relations between means and ends. For some goals, people attempt to manage this uncertainty by developing a *backup plan.*[1] The research presented here tests a position presented by Napolitano and Freund (2016): developing backup plans can introduce additional practical and motivational costs that may hamper performance. This research is the first to empirically examine this hypothesis that investing in developing a backup plan may facilitate precisely the conditions that the backup plan was developed to potentially address; that is, Plan A's (perceived or actual) insufficiency. We term this effect the *backup plan paradox*. More specifically, we hypothesize that greater objective or perceived investment in developing backup plans and higher perception of the instrumentality of backup plans can *indirectly* compromise performance through decreasing the instrumentality of and commitment to continue using Plan A.

## Backup Plans: A Gap in Understanding Self-regulation

Backup plans can be defined as alternative means to an end that are intentionally developed but not initially (or ever) used (Napolitano & Freund, 2016). Backup planning involves three processes: A person (1) *develops* backup plans, selecting them from a set of means that equifinally lead to a goal (see Kruglanski, Chernikova, Babush, Dugas, & Schumpe, 2015), and investing resources in these alternative means to make them potentially useful. Next, backup plans are (2) *reserved,* initially unused but remaining available to potentially later (3) *replace* a Plan A when it is underperforming (for a theoretical discussion of backup plans, see Napolitano

& Freund, 2016). In the example of the above-described post-doc: She *develops* her backup plan by investing time and effort in preparing the notes. While extemporaneously speaking, she *reserves* the backup plan by keeping the notes available at the lectern without using them. If she struggles, she might *replace* with her backup plan and switch to giving the talk from her notes. To illustrate the backup plan paradox, we speculate that if the post-doc extensively invests in developing her notes, she may quickly decide to present from her notes after initial jitters, perhaps resulting in a poor presentation.

**Locating backup plans in the literature**

Backup plans are widely used yet not adequately described by the existing literature. There are two key distinctions that differentiate backup plans from other constructs in the field of self-regulation (i.e., implementation intentions, coping planning, multiple concurrent means) and serve to frame this investigation. First, developing, reserving, and replacing backup plans are *intentional* processes aimed at increasing the likelihood of achieving a goal (Napolitano & Freund, 2016). In the case of the job talk, the applicant deliberately *chooses* to develop a backup plan, to reserve it, and to use it (or not). In contrast, implementation intentions (i.e., "If challenge X occurs, then I will respond with Y;" Gollwitzer, 1999) and the related concept of coping planning (Sniehotta, Schwarzer, Scholz, & Schüz, 2005) are described as being automatically and efficiently activated by environmental cues (Bargh, Schwader, Hailey, Dyer, & Boothby, 2012).

Intentionally developing, reserving, and potentially replacing with backup plans introduces costs into goal pursuit. Developing backup plans requires some degree of resource expense (e.g., time, effort), that, like opportunity costs, could have been allocated towards one's (now potentially underdeveloped) Plan A. From a motivational perspective, having backup plans

available could later decrease commitment to one's Plan A (Kruglanski, Pierro, & Sheveland, 2011).

Reserving backup plans can also incur costs. Similar to goal shielding (Shah, Friedman, & Kruglanski, 2002), a person should limit the distracting possibility of replacing Plan A while maintaining Plan B's availability for potential use. Reserving backup plans may be especially difficult when Plan B is highly instrumental (Shah & Kruglanski, 2002), or when backup plans are developed at great cost, which might motivate people to keep investing into them (similar to the "sunk costs fallacy," e.g., Garland & Newport, 1991). For example, if the post-doc falters at the podium, it may become increasingly difficult to resist switching to presenting via her notes, especially if she thinks her notes are excellent or if she spent considerable effort developing them.

Intentionally replacing Plan A with a backup plan also introduces costs. One must recalibrate goal pursuit to align with the new means. For example, if the post-doc switches to her notes, she must find her place, and remember to periodically make eye contact while presenting. Consistent with research on affective transfer in goal systems (Fishbach, Shah, & Kruglanski, 2004), replacing Plan A may involve acknowledging the limitations of those means, an affectively-aversive experience akin to admitting failure (Iles, Judge, & Wagner, 2010).

A second key way in which backup plans are distinct from concurrently available multiple means is that they are *reserved* for potential later use. Backup plans are not concurrently used alongside Plan A. Despite this difference, the literature on pursuing goals with multiple concurrent means is informative. For example, when one begins to pursue a goal, having concurrent means available supports motivation by offering several potential paths to success (Etkin & Ratner, 2012, 2013). However, as the goal nears completion, using multiple means

concurrently can be inefficient (Huang & Zhang, 2013). Relatedly, pursuing a goal with multiple means concurrently increases overall commitment to the goal, but can come at the cost of decreasing commitment to continue using any one means (Kruglanski et al, 2011). Concurrently using several equifinal means introduces a dilution effect, where adding more equifinal means decreases perceived instrumentality and motivation to use those means (Bélanger, Schori-Eyal, Pica, Kruglanski, & Lafrenière, 2015; Zhang, Fishbach, & Kruglanski, 2007).

**The present research**

This research examines whether highly investing (or perceiving to invest) in developing backup plans can negatively affect performance through decreasing the instrumentality of and commitment to continue using Plan A. More specifically, the studies reported here test four hypotheses:

1. The more a person invests or perceives to invest into developing backup plans, the more likely s/he is to later use that backup plan, either through the decreased instrumentality of a less developed Plan A, the decreased motivation to continue using Plan A, or both.
2. The more instrumental a person perceives a backup plan to be, the more likely s/he is to later use that backup plan.
3. Because the replacement of Plan A with Plan B incurs practical and motivational costs, backup plan use leads to decreased performance.
4. Highly investing in developing backup plans and perceiving backup plans to be instrumental *indirectly* affects performance through the effect of increased backup plan use.

**Overview of Studies**

We tested these hypotheses across five studies and three tasks. Using a logical deduction task, Study 1 tested the hypothesis that high perceived investment in developing a backup plan

increases the likelihood of its use, regardless of skill or performance level, and even if investment itself was unrelated to backup plan utility. [2] Studies 2 to 4 used variations of a simulated, hypothetical negotiation. Study 2 focused on assessing the practical costs of developing backup plans that are similar to opportunity costs, whereas Study 3 assessed the motivational costs of developing backup plans. Study 4 assessed the motivational costs of developing backup plans through an experimental variant of the task. Study 5 used a new ball-throwing task to further assess the practical and motivational costs of developing backup plans in a third goal domain.

## Study 1

Study 1 tests Hypothesis 1 by examining if the perceived cost of making a backup plan affects the likelihood of using it. Importantly, we ensured that investing in developing the backup plan would not increase its utility (or it would be trivial if people used the backup plan more often). Similarly, we ensured that investing in developing the backup plan would not influence a participant's skill with Plan A. Finally, to investigate the effect of investment independent of prior skill or performance in goal pursuit, we created a new task that was not related to participants' prior knowledge or skill.

## Methods

### Participants[2]

Power analyses (G*Power; Faul, Erdfelder, Buchner, & Lang, 2009) indicated that 325 participants are sufficient to detect a small-to-medium effect ($w = .20$) of group assignment on backup plan use (.95 power, .05 error). To plan for potential dropouts, we recruited an age-diverse community sample of $N = 351$ participants via Amazon's MTurk. After random assignment, the high development cost group included $n = 174$ participants, the low development

cost group $n$ = 177 participants. The two groups did not differ significantly in terms of their demographics or individual characteristics (e.g., performance in the digit symbol substitution task, Wechsler, 1958; hereafter: DSST). More detailed demographic information for all studies can be found in Table 1. Participants were compensated with $2, as well as performance-based compensation for correct answers (see below). Average total participant compensation was $3.61 (min = $2.30, max = $4.00). All participants in this research were limited to participating in a single study.

**Procedure**

Figure 1 shows a simplified version of the procedure. Participants were randomly assigned to a high or a low development cost group, provided demographic information, and completed a 90-second DSST as a measure of fluid intelligence as a control variable. Participants in the low-cost group completed an additional two-minute DSST, to align participation times and relative cognitive effort across groups. This duration was based on two pilot studies.

The central experimental manipulation in this study concerned the perceived and actual cost of developing a backup plan by having the low-cost group spend relatively little time and the high-cost group relatively more time on constructing their backup plan. In this study, the backup plan was a "crib sheet" that participants could elect to use to solve the task. The crib sheet was identical across groups, but the effort required to construct it differed across groups.

*Developing phase*

After the DSST, participants read this text, which introduced the task's developing and trial phases:

"The main task of this study is to find the correct names for shapes. The shapes will be shown to you one by one and you have to find out the correct name in a multiple-choice

task. (…) you can find out the correct name for a given shape only by applying some logic. In some cases, that may be challenging. For this reason, we will give you the opportunity to use a backup plan.

Your backup plan will be a grid that shows all the possible shapes and their names. In order to be able to use the grid, however, you have to sort the shapes.

Once you have correctly sorted the grid, you may look at your backup plan during the main task. However, given that your "Plan A" is to solve the task through logic, there will be a penalty for using the backup plan."

The developing phase involved unscrambling a 5 (shape) x 8 (pattern) grid of symbols to develop a backup plan crib sheet. Participants could not learn the names of the shapes by unscrambling the grid as it did not contain the names. Participants in the high-cost condition had a fully scrambled grid, with no symbols in the correct location. Participants in this group were given only written instructions for the correct vertical and horizontal order of the grid. In contrast, participants in the low-cost condition unscrambled a partially-completed grid. In this condition, the top row and the first three columns were locked in in the correct location. In both conditions, participants could not proceed until they had successfully unscrambled the grid.

*Trial phase*

Participants began the trial phase by viewing three practice items. These items, and all trial items were arrayed as multiple choice questions, with five potential responses: three possible shape names, one "none of the above option" and one "check your backup notes" option. To the left of the response options, a small 3 (shape) by 4 (pattern) grid of shapes was consistently displayed. This grid provided sufficient information for participants to logically deduce a potentially correct response for every trial. Participants were instructed that their "Plan A" was

attempting to provide a correct possible name for the symbol using only deductive logic. Correct responses using only logic were rewarded $0.10. Participants were instructed that their backup plan involved viewing the "crib sheet" that they had unscrambled earlier, which now provided participants with the correct name for each shape. Correct responses using the backup plan were rewarded only $0.07. Incorrect response options resulted in no payment.

The task involved 20 trials displayed in a fixed order. Fourteen trials required participants to use deductive logic (in seven cases, the correct option was "none of the above"). In two trials, the symbol was displayed in the small grid, and the correct name was provided as a response option. In four trials, the symbol was displayed in the small grid, but the correct response was "none of the above." There were four trials for each of the five shapes, and all eight patterns were displayed in either two or three trials.

**Measures**

*Objective and perceived backup plan development costs*

We included objective and perceived manipulation checks to assess whether participants in the high-cost condition invested more in developing their BUP. The objective manipulation check involved comparing time spent unscrambling the grids across groups. Three new self-report items composed the perceived costs manipulation check. An example item was "How much effort did you have to invest into sorting the shapes to make your backup plan?" Responses ranged from 1 (not at all) to 7 (very much). We used a scale comprised of the average of the three items ($M$ =4.64, $SD$ = 1.51), which was acceptably reliable (Cronbach's $\alpha$ = .83).

Participants in the high-cost group took objectively longer to unscramble ($M$ =457.06 seconds, $SD$ = 291.51) than those in the low-cost group ($M$ =171.35, $SD$ = 68.24) $t$ (286) = 12.00, $p$ <.001; $d$ = 1.42, (90% CI = 1.21, 1.64). In addition, participants in the high-cost group

reported a higher perceived cost ($M$ =5.49, $SD$ = 1.20) than those in the low-cost group ($M$ =3.96, $SD$ = 1.38) $t$ (286) = 9.85, $p$ <.001; $d$ = 1.17, (90% CI = .96, 1.38).

*Attrition and dropouts*

Prior to all analyses, we removed 11 people who reported taking screenshots of their backup plan during the trial phase, and two outliers who had extremely high times on developing phase (> 7 SD above the mean). At this point, the two groups did not differ in terms of sex, age, or DSST performance. However, 50 people began but did not finish the study, 43 of whom dropped out during the developing phase; 72 % of the dropouts were from the high-cost group $\chi^2$ (1, 338) = 19.19, $p$ < .001. This disproportionally high drop-out rate in the high-cost group is consistent with results by Zhou and Fishbach (2016) who found that online studies placing high demands on self-regulation lead to a selective drop-out of participants low in self-regulation. In line with this assumption, dropouts had significantly lower DSST performance ($M$ = 37.00, SD = 10.59) than remaining participants ($M$ = 40.81, SD = 8.86), $t$ (336) = 2.72, $p$ = .007; $d$ = .41, (90% CI = .16, .67).

Zhou and Fishbach (2016) argue that selective attrition in demanding tasks can severely impact results in self-regulation studies. Therefore, we excluded the high-cost group as a whole from this study and instead based the analyses on the associations between perceived development cost and backup plan use within the low-cost group only ($N$ = 160). Detailed demographic information for this final sample can be found in Table 1. We report descriptive information for subsequent measures using the final sample of participants in the low-cost group.

*Digit-Symbol Substitution task*

We programmed an online DSST version to assess fluid intelligence, which may be associated with participants' backup plan behaviors.

*Task Performance*

We instructed participants to provide a possible name for shapes using only logical deduction as their Plan A. We operationalized task performance as the proportion of correct answers using logical deduction (Plan A) to the number of answers attempted using logical deduction because using the total number of logical deduction correct answers would be confounded with the number of times participants used their backup plan. On average, participants correctly answered 81% (*SD* = 19%) of the trials they attempted using only logical deduction.

*Replacing with backup plans*

The main dependent variable of Study 1 was participants' decisions to replace Plan A with their backup plan. We operationalized backup plan use as the number of times participants viewed their "crib sheet" during the twenty trials. About half (51.9%) of participants used their backup plan at least once. On average, participants viewed their backup plan more than two times ($M = 2.58$, $SD = 4.20$; min=0, max=20). Replacing behavior was not normally distributed. We recoded backup replacing into a dichotomous non-use/use variable. Sex $\chi^2$ (2, $N = 160$) = .31, $p = .58$, and age $t(158) = -.27$, $p = .78$ were not associated with backup plan use.

## Results

### Replacing with backup plans

We hypothesized that the more people invest (or perceive to invest) in backup plans, the more likely they are to be used. We tested this hypothesis by assessing whether there were significant differences in objective (time) and perceived costs of developing a backup plan across participants who did and did not use backup plans during the task.

Objective development costs did not significantly differ between participants who elected to use Plan A only ($M$ =179.0 seconds, $SD$ = 70.4) and participants who chose to use Plan B ($M$ =163.1 seconds, $SD$ = 65.3; $t$ (158) = -1.47, $p$ =.14). However, in support of our hypothesis, participants' who used Plan B (M = 4.24, SD = 1.33) reported higher perceived development costs than participants who only used Plan A ($M$ =3.66, $SD$ = 1.39), $t$ (158) = -2.70, $p$ = .008, $d$ = -.43 (90% CI = -.93, -.22).

**Testing alternative hypotheses**

A reasonable alternative hypothesis may be that backup plan use is simply explained in terms of a lack of skill in Plan A. In this study, this hypothesis would suggest that participants who correctly named a low proportion of shapes using logic should view their backup plan more often. While plausible, the data do not support this alternative hypothesis, $t$(158) = -1.58, $p$ = .12. Furthermore, while participants' ability to correctly answer items using only logic was concurrently associated with both their performance on the DSST (β = .19, $t$[244] = 2.30, $p$ = .02) and their objective cost (time spent unscrambling; β = -.17, $t$[244] = -2.07, $p$ = .04) neither of these variables – both presumably associated with participants' cognitive abilities – were associated with increased backup plan use (results for DSST: $t$(158) = 1.58, $p$ = .12; results for objective cost: $t$(158) = -1.47, $p$ = .15).

## Discussion

The findings from Study 1 provide initial support of our first hypothesis. Participants who elected to use their backup plan during the task perceived that they invested more in developing this Plan B. This correlational result is consistent with research indicating that perceived (rather than objective) task demands drive effort mobilization (e.g., Gendolla, 2012). Note that higher levels objective levels of costs in developing a backup plan (in this case, time) did not predict its

use. Thus, it seems that, psychologically, the perceived costs – rather than actual time invested – may be important for understanding backup plans' subsequent use.

Study 1 had a key limitation. Our experimental manipulation resulted in a disproportionate number of dropouts from the high-cost group, precluding our ability to conduct experimental analyses. This task is also not suited for assessing the effects of investing in backup plans on overall performance, as participants could elect to answer questions in a more (Plan A) or less (Plan B) difficult way.

## Study 2

## Methods

### Participants

Power analyses indicated that 291 participants are sufficient to detect a small effect ($f^2$ = .06) on performance (with three predictors, .95 power, .05 error). We recruited an age-diverse community sample of $N$ = 286 participants via MTurk. Thirty-nine failed one or more points of the inclusion criteria (age <18 years; USA resident, correct responses to check questions, participation time > 5 minutes). The final sample included $N$ = 247 adults.

Participants were compensated $1 for participation. In addition, for each correct response, $ 0.50 was donated to an environmental charity (1% for the Planet; total donation: $525.50).

### Procedure

Study 2 assessed the practical costs of developing backup plans. The simulated negotiation task (used with adaptations across Studies 2-4) had three phases: introduction, information review, and negotiation. After providing consent and demographic information, participants were introduced to the task, which was couched in the story of hypothetically negotiating with a chemical company to build a cleaner factory. Next, participants read

descriptions of three categories of information they could review: economic benefits of "green" business practices ("economic benefits"); ecological benefits of natural ingredients ("natural ingredients"); and the negative effects of chemical spills ("chemical disasters"). Each category corresponded to a plan they could use in the negotiation. After reading about each plan, participants rated each plan's perceived instrumentality. Each plan was designed to be equally instrumental. Participants next reviewed seven pieces of information (of a possible 21; 3 plans x 7 pieces of information).

In the information review phase, participants began by deciding how many pieces of information, of a possible seven, that they wished to review per plan (e.g., six "chemical disasters" and one "natural ingredients"). Participants then read the seven information pieces in a randomized order. If a participant allocated all seven information pieces to one plan, then s/he adopted a "Plan A only" approach. If a participant allocated investments across more than one plan, then s/he chose to develop a backup plan(s). The most-read category was considered Plan A; participants could have 0-2 backup plans. All information pieces were of similar length and complexity, and combined facts with fictional content that participants would not be aware of if they had not reviewed the information (e.g., factual information about the 1989 *Exxon Valdez* disaster, which was harmful to avian life, and fictional information about the imaginary chemical company's owner having a bird photography hobby).

The simulated negotiation phase involved three steps, repeated over seven rounds: (1) participants' choosing the plan upon which the next negotiation round would be based; (2) participants' answering a question by selecting one statement out of four alternatives; and (3) feedback on whether or not the chosen statement was correct. Each question had one correct answer, and was based on the content presented in one information piece.

After seven rounds, participants were informed about their performance and debriefed. There were no missing data as the instrument required answers for each item. Study 2's design is presented in Figure 2.

## Measures

*Backup plan developing costs*

Participants' choices of how many pieces of information they wanted to review per plan provided an index of backup plan development costs. More than half of the sample (63.2%) chose to invest in backup plans to some extent. Of these, 72.4% invested in one backup plan. Of the seven information pieces they could review, participants read an average of 5.27 (*SD* = 1.45) for their Plan A, and 1.72 (*SD* = 1.45) for their backup plan(s). Therefore, the average ratio of information reviewed for backup plans to overall information reviewed (hereafter "backup plan development ratio") was .25 (*SD* = .21). Sex ($r$ = -.02, $p$ = .71) and age ($r$ = .08, $p$ = .23) were not associated with backup plan development ratio and were not included in subsequent analyses.

*Perception of backup plans' instrumentality*

Before reviewing the information pieces, participants answered the following question for all three plans: "How effective do you think using the [*plan name*] will be in the negotiation?" Responses were provided on a 0 ("not at all") to 7 ("very much") scale. We calculated the perception of backup plans' instrumentality as the highest rating for a plan other than a participant's Plan A (*M* = 4.47, *SD* = 1.55). Therefore, all participants provided information on their perceptions of backup plans' efficacy, although only some developed them. Sex ($r$ = -.01, $p$ = .90) and age ($r$ = .04, $p$ = .50) were not associated perceived backup plan instrumentality. Participants perceived their Plan A to be more instrumental than their backup plan (average difference = 1.58, SD = 2.18).

*Replacing with backup plans*

The number of times a person changed plans during the simulated negotiation provided an index of backup plan use. In total, 54.7% of participants changed plans at least once. There was no limit to the number of times a person could change plans. On average, participants changed plans between one and two times ($M$ = 1.47, $SD$ = 1.64). Sex ($r$ = .07, $p$ = .25) and age ($r$ = .00, $p$ = .98) were not associated with backup plan use.

*Performance*

We operationalized performance as the number of correct choices across the simulated negotiation. On average, participants chose more than four of seven statements correctly ($M$ = 4.26, $SD$ = 1.61). Sex ($r$ = .01, $p$ = .91) and age ($r$ = .00, $p$ = .99) were not associated with performance.

## Results

### Effects of backup plans on commitment to Plan A

Taken together, perceived instrumentality and backup plan development ratio were associated with the number of times participants changed plans during the task, $F$[2, 244] = 44.80, $p$ < .001, $R^2$ = .27 [90% CI = .19, .35]. A higher backup plan development ratio (β = .46, $t$[244] = 8.15, $p$ < .001) and a higher perceived instrumentality (β = .18, $t$[244] = 3.12, $p$ =.002) were associated with an increased number of plan changes.

### Effects of backup plans on performance

Consistent with our expectations, participants who developed at least one backup plan performed worse (i.e., selected fewer correct statements; $M$ =3.92, $SD$ = 1.42) than participants who only developed Plan A ($M$ =4.82, $SD$ = 1.77), $t$(245) = 4.39, $p$ < .001, $d$ = .58 (90% CI = .36, .80). The next analyses explored the process underlying this effect. We hypothesized that high backup plan development ratio and high perceived instrumentality negatively affect

performance indirectly, as these factors lead to a less-well-developed, less instrumental Plan A, and thus increased levels of replacing. Taken together, perception of a backup plan's instrumentality, backup plan development ratio, and the number of changes between different plans significantly predicted performance, $F[3, 243] = 18.19$, $p < .001$, $R^2 = .18$ [90% CI = .11, .25]. Consistent with hypotheses, only the number of plan changes was a significant and negative ($\beta = -.34$, $p < .001$) predictor of performance.

Given these relations and our hypotheses, we next tested whether increased costs in developing backup plans and higher perceptions of backup plan instrumentality were significantly associated with performance through the indirect effect of increased plan changes. We estimated a bootstrapped indirect effects (Hayes, 2009) path model in *MPlus* 7 (Muthеń & Muthеń, 2015). This simultaneous indirect effects model is displayed in Figure 3.

Across 10,000 bootstrapped samples, the standardized indirect effect of backup plan development ratio on performance was -.16 ($p < .001$; 95% CI = -.22, -.09). Simultaneously, the standardized indirect effect of backup plan perceived instrumentality on performance was -.06 ($p = .01$; 95% CI = -.10, -.02). Effect size for indirect effects is often calculated as $\kappa^2$ (Preacher & Kelly, 2011) and interpreted in the same way as $R^2$ (small > .01, medium > .09, or strong effect >.25). The size of the indirect effects of backup plan development ratio and backup plan perceived instrumentality on performance were medium ($\kappa^2 = .14$) and small ($\kappa^2 = .06$), respectively.

## Discussion

Results from Study 2 provide correlative support for our hypotheses and illustrate the practical costs of developing backup plans for performance. Participants who invested in developing backup plans chose fewer hypothetically-correct statements in a simulated

negotiation than participants who only developed a Plan A. In exploring the process underlying this effect, both a higher backup plan development ratio and perceived backup plan instrumentality were associated with increased replacing with backup plans during the task. Second, increased backup plan development ratio and perceived instrumentality were associated with decreased performance through the indirect effect of increased replacing with backup plans. In sum, for goals where investments in backup plans come at the direct expense of investments in one's Plan A, performance may suffer. People who invest scarce resources in backup plans may have fewer resources left to develop their Plan A, thereby potentially decreasing Plan A's instrumentality and the commitment to continue using it.

The design and programming of this task introduces some caveats for interpreting the results. The first limitation involved developing costs. As the choice of correct statements in the simulated negotiation was based on the information pieces, people who developed a single means and continued using that means throughout the task were always presented statements that they had previously reviewed. In contrast, participants who developed backup plans decreased the likelihood of encountering a statement they had previously reviewed. Although this design choice was intentional, and although people often must selectively invest their limited resources to adaptively pursue a goal (Freund, 2008), restricting resources in this way is not generalizable to all goal pursuits. Second, although participants perceived that their Plan B was less effective than their Plan A, we only identified participants' backup plans using an inferential approach. A third limitation was based in the task's programming, which did not provide information on the sequence of questions a participant answered. This precluded our ability to analyze the data hierarchically. We attempted to address these shortcomings in Study 3.

## Study 3

## Methods

### Participants

Sample size was determined using the same power analysis parameters as Study 2. An age-diverse community sample of $N = 303$ participants was recruited via MTurk. Of these, $n = 55$ participants failed the inclusion criteria or had corrupted data. The final sample included $N = 248$ adults. Participants were compensated $1 for participation. In addition, for each correct response, $ 0.50 was donated to 1% for the Planet (total donation: $556).

### Procedure

The procedure for Study 3 was identical to Study 2 except for three changes. First, instead of being limited to reviewing seven pieces of information, participants could now choose to review as little as one and as many as 21 information pieces (which represented the entire set). Second, participants nominated which plan represented their Plan A, and which plan, if any, was a backup plan(s) before they reviewed information. Third, we re-programmed the task to provide sequential response data.

### Measures

*Backup plan developing costs*

The measures in Study 3 were identical to Study 2. Consistent with Study 2, almost two thirds of the sample (68.1%) chose to invest in backup plans. Of the "backup planners," 42.0% chose to invest in one backup plan, and 58% to invest in two backup plans. Of 21 possible pieces of information, participants averaged reading 7.65 ($SD = 4.83$). Participants again invested more in developing Plan A, on average reading 4.93 ($SD = 2.45$) pieces of Plan A information, and reading 2.70 ($SD = 3.37$) for their backup plan(s). Similar to Study 1, the average backup plan development ratio was .26 ($SD = .25$). Sex and age were not associated with backup plan

development ratio ($r$ = -.01, $p$ = .90 and $r$ = .06, $p$ = .36, respectively), nor overall amount of information reviewed ($r$ = .07, $p$ = .25 and $r$ = .08, $p$ = .20, respectively).

*Perception of backup plans' instrumentality*

We calculated the perception of backup plan instrumentality in the same way as Study 1. Average levels of perceived instrumentality were similar across studies ($M$ = 4.39, $SD$ = 1.55). Sex ($r$ = .03, $p$ = .60) and age ($r$ = .05, $p$ = .47) were not associated with perceived instrumentality. Consistent with Study 2, participants reported that they perceived their Plan A to be more instrumental than their backup plan ($M_{\text{difference}}$ = 1.91, $SD$ = 1.82).

*Replacing with backup plans*

Number of plan changes served as an index of replacing. Similar to Study 2, 52.0% of the participants changed plans at least once; on average, participants changed plans between one and two times ($M$ = 1.54, $SD$ = 1.79). Sex ($r$ = -.06, $p$ = .34) and age ($r$ = .07, $p$ = .29) were not associated with number of plan changes.

*Performance*

Performance was assessed in the same way as Study 2. Consistent with Study 2, participants averaged between 4 and 5 correctly chosen statements across the seven rounds of the simulated negotiation ($M$ = 4.35, $SD$ = 1.67). Sex ($r$ = .06, $p$ = .35) and age ($r$ = .01, $p$ = .84) were not significantly associated with performance.

## Results

### Effects of backup plans on commitment to Plan A

We first tested whether increased backup plan development ratio and higher perceived backup plan instrumentality were again associated with more plan changes. In order to compare these results with those of Study 2, and to better assess the motivational costs of backup plans,

we controlled for the overall amount of information participants reviewed. Perceived instrumentality, backup plan development ratio, and the overall amount of information reviewed were together associated with the number of times a person changed plans during the simulated negotiation, $F[3, 244] = 21.15$, $p < .001$, $R^2 = .21$ [90% CI = .14, .28]. A higher backup plan development ratio ($\beta = .45$, $t[244] = 6.38$, $p < .001$) and a higher perceived instrumentality ($\beta = .21$, $t[244] = 3.63$, $p < .001$) were again associated with an increased number of plan changes. The more information a person reviewed was associated with fewer plan changes ($\beta = -.38$, $t[244] = -5.25$, $p < .001$).

**Effects of backup plans on performance**

We next assessed whether participants who elected to develop only a Plan A differed in their performance compared to participants who developed at least one backup plan. Consistent with expectations, participants who developed at least one backup plan selected significantly fewer correct statements ($M$ =4.21, $SD$ = 1.68) than participants who only developed a Plan A ($M$ =4.67, $SD$ = 1.59), $t(246) = 2.06$, $p = .04$, $d = .28$, (90% CI = .05, .50).

Our next analysis tested the process underlying this effect. We expected that high backup plan development ratio and perceived instrumentality were negatively associated with performance through the indirect effect of increased replacing, controlling for participants' chosen levels of information review. Taken together, these predictors were significantly associated with performance ($F[4, 243] = 16.64$, $p <.001$, $R^2 = .22$ [90% CI = .15, .29]). Backup plan development ratio was not significantly associated with performance; however, perception of backup plan instrumentality ($\beta = -.21$, $p = .001$), overall amount of information reviewed ($\beta = .17$, $p = .02$) and number of plan changes ($\beta = -.30$, $p < .001$) were associated with the number of correctly chosen statements.

We next tested whether the indirect effects of Study 2 were replicated in Study 3, with the addition of overall information review as a covariate for both category changes and performance. Figure 4 illustrates these results. Across 10,000 bootstrapped samples, the standardized indirect effect of backup plan development ratio on performance was -.14 ($p < .001$; 95% CI = -.21, -.07). Simultaneously, the standardized indirect effect of backup plan perceived efficacy on performance was -.06 ($p = .01$; 95% CI = -.11, -.02). The size of the indirect effects of backup plan development ratio and backup plan perceived instrumentality on performance were medium ($\kappa^2$=.14) and small ($\kappa^2$=.08), respectively.

The indirect effects of backup plan development ratio (-.08 [$p = .001$; 95% CI = -.13, -.03]) and perceived instrumentality (-.09 [$p = .002$; 95% CI = -.14, -.04]) remain significant without the inclusion of overall information reviewed as a covariate. Without the covariate, the size of the indirect effects of backup plan development ratio and backup plan perceived instrumentality on performance were small ($\kappa^2$=.08) and medium ($\kappa^2$=.09), respectively.

**Testing alternative hypotheses**

An alternative explanation for these findings is that replacements were based on performance in the prior negotiation round. In other words, people are more likely to change their plans when they selected an incorrect statement in the previous round. This explanation is reasonable but not mutually exclusive with our backup plan-based hypotheses. Instead, we expect that higher backup plan development ratio and higher backup plan perceived instrumentality predict a greater likelihood to replace one's active plan even while controlling for prior performance.

To test this position, we used logistic hierarchical linear modeling (HLM) conducted using the *R* package "lme4." An HLM approach accounts for the non-independence of

participants' repeated opportunities to replace plans before negotiation questions. In this model, the likelihood of replacing was predicted by the fixed effects of prior performance, backup plan development ratio, backup plan perceived instrumentality, and overall information review. Table 2 summarizes the results of these and subsequent HLM analyses. As hypothesized, participants were more likely to replace when they invested highly in the development of their backup plans, as well as when they perceived their backup plans to be potentially instrumental, even when controlling for prior performance and overall information

## Discussion

Study 3 extends Study 2 and provides support for our hypothesis that investing in backup plans that are perceived to be instrumental introduces motivational costs that can be detrimental to performance. Participants who invested in developing backup plans performed worse in the negotiation than participants who invested in developing only Plan A. Higher costs of developing backup plans and higher perceptions of the instrumentality of backup plans were again associated with more plan changes during the task. These effects remained significant when controlling for overall information review and prior performance. In turn, increased costs of developing backup plans and higher perceptions of the instrumentality of backup plans were indirectly associated with decreased performance through the effect of increased category changes. It should be noted that, in this study, higher levels of overall investment were associated with increased performance. This suggests that it may not be the raw amount of resources one spends on backup plans that decreases motivation to continue using Plan A, but instead the relative scale of backup plan costs. Taken together, the results indicate that for goals for which investing in backup plans *does not necessarily* come at the direct expense of developing a Plan

A, increased backup plan investment may introduce motivational costs that could decrease commitment to continue using Plan A, which in turn could result in worse performance.

In Studies 2 and 3, participants chose whether or not to develop backup plans, and the extent to which they developed these means relative to their Plan A. Thus, it is uncertain whether the results derive from the goal behaviors themselves or interindividual differences. We addressed this limitation in Study 4.

## Study 4

## Methods

### Participants

Studies 2 and 3 demonstrated that choosing to invest in backup plans can have a small-to-medium negative effect on performance. Study 4 used random assignment into a backup plan or control group to experimentally test this effect. Based on a power analysis, we recruited an age-diverse community sample of $N = 240$ participants via Amazon's MTurk. Of these, $n = 36$ participants failed one or more points on the inclusion criteria or had corrupted data, resulting in a final sample of $N = 204$ (control group: $n = 105$, backup plan group: $n = 99$). The two groups did not differ significantly in terms of their demographics or individual characteristics. Participants received $2.25 for participation, and each correct response resulted in $0.50 donation to 1% for the Planet (total donation: $496.50).

### Procedure

Study 4 used an experimental version of Study 3's simulated negotiation task with the following adjustments: All participants read 14 pieces of information, but the control-group was assigned as Plan A "economic benefits" by reading the seven information pieces associated with this information category that would be used during the negotiation. They also read seven

information pieces describing economic benefits that would not be used in the negotiation to hold the number of reviewed information pieces constant across groups. Control group participants could not change categories during the task. The backup plan group read seven pieces of information about economic benefits as their Plan A and seven pieces on chemical disasters as their backup plan. They could change plans across the seven rounds of the simulated negotiation.

**Measures**

*Perception of backup plan instrumentality*

Only participants in the backup plan group had two categories of information to review. In this study, perception of backup plan instrumentality was calculated as the perceived instrumentality of the chemical spills category because this was Plan B for all participants. Mean levels of perceived backup plan instrumentality were similar to those in Studies 2 and 3 ($M$ = 4.25, $SD$ = 1.49). Sex was not associated with perceived backup plan instrumentality ($r$ = -.09, $p$ = .41); however, age was negatively associated with higher perceived backup plan instrumentality ($r$ = -.23 $p$ = .02). Consistent with Studies 1 & 2, participants reported Plan A to be more instrumental than their backup plan (average difference = 2.67, $SD$ = 2.12).

*Replacing with backup plans*

Similar to prior studies, 57.6% of participants in the backup plan group changed plans at least once during the task. On average, participants changed plans between one and two times ($M$ = 1.34, $SD$ = 1.79). Sex ($r$ = -.05, $p$ = .51) and age ($r$ = -.08, $p$ = .25) were not associated with number of plan changes.

*Performance*

Performance was measured in the same way as Studies 2 and 3. Consistent with levels of performance in those studies, participants in this study averaged between 4 and 5 correct answers ($M$ = 4.55, $SD$ = 1.57). Sex ($r$ = -.12, $p$ = .10) and age ($r$ = .06, $p$ = .39) were not significantly associated with performance.

## Results

### Effects of backup plans on performance

There were no significant differences in performance between the control ($M$ =4.66, $SD$ = 1.67) and backup-plan groups ($M$ =4.42, $SD$ = 1.33), $t(202) = 1.11$, $p$ = .27, $d$ = .16 (90% CI = .07, .39). This result indicates that for this simulated negotiation task, simply having a backup plan available (without intentionally choosing to develop one and with no additional development costs) is not associated with decrements in performance.

Did backup plan group participants who changed plans during the task perform worse than those who did not (and worse than the control group)? A between-groups ANOVA indicated significant differences in performance, $F(2, 201) = 7.79$, $p$ = .001, η= .07 (90% CI = .02, .13). We used a Games-Howell post-hoc comparison to test for group differences because the variances were not homogenous. Participants in the backup condition who replaced Plan A performed significantly worse ($M$ = 3.93, $SD$ = 1.51) than both participants in the backup condition who did not replace ($M$ = 5.10, $SD$ = 1.03, $p$ > .001) as well as control group participants ($M$ = 4.67, $SD$ = 1.67, $p$ = .01).

### Effects of perceived backup plan instrumentality on Plan A commitment and performance

Why did some participants in the backup plan group replace their Plan A, whereas others continued using their Plan A? Consistent with Studies 2 and 3, we expected that participants replaced Plan A more often when they perceived their backup plan to be potentially instrumental.

This expectation was supported: those who replaced Plan A perceived the backup plan to be significantly more instrumental ($M$ =4.54, $SD$ = 1.44) than those who chose to use only Plan A ($M$ =3.86, $SD$ = 1.48) $t$ (97) = 2.32, $p$ = .02; $d$ = .47, (90% CI = .13, .81). Consistently, higher perceived instrumentality of the backup plan was also associated with more replacing, β = .27, $p$ = .006; $F$[1, 97] = 7.84, $p$ =.01, $R^2$ = .08 (90% CI = .00, .16). However, perceived backup plan instrumentality was not associated with performance, $F$[1, 97] = .61, $p$ = .44, $R^2$ = .01.

We next tested whether the indirect effect of perceived backup plan instrumentality on performance found in Studies 2 and 3 could be replicated for the backup plan group participants of Study 4. This path model is presented in Figure 5. Across 10,000 bootstrapped samples, the standardized indirect effect of backup plan perceived instrumentality (controlling for the effect of age) on performance through the effect of increased category replacing was -.10 ($p$ = .01; 95% CI = -.18, -.03). The effect size was medium ($\kappa^2$=.10).

**Testing Alternate Hypotheses**

We again tested the alternate hypothesis that choosing to change plans was simply a function of prior question performance. This HLM analysis differed from the similar analysis in Study 3 in two ways. First, we only analyzed data from participants in the backup plan condition, as they were the only participants able to replace categories during the task. Second, we did not assess the influence of backup plan development ratio on plan replacement, because all backup plan group participants read identical information on Plan A and Plan B (at a 1:1 ratio). We expected that participants were more likely to replace their plans if they perceived their backup plan to be instrumental, even after accounting for the effect of performance in the prior round. The results for this analysis are summarized in Table 2. As hypothesized, the more instrumental participants perceived their backup plans to be, the more likely they were to change plans.

Consistent with Study 3, these effects were significant even controlling for the effects of prior performance.

## Discussion

Study 4 replicates and extends findings from Studies 1, 2, and 3. Consistent with our hypotheses, even when the developing costs of a backup plan were held constant, reserving a backup plan that is perceived to be instrumental can harm performance through decreasing a person's commitment to continue using Plan A. This effect remained significant even when accounting for prior performance. However, the control group and backup plan group did not significantly differ in their overall performance. This suggests that simply being provided a backup plan may not (in and of itself) affect performance. Instead, when a person reserves a backup plan that he or she perceives to be particularly useful, this can undermine motivation to continue using Plan A, and ultimately hamper performance.

One limitation of these findings was the unexpected relation between age and perceived backup plan instrumentality. It may be that, for this sample, older adults recalled the chemical disasters described in the information pieces and perceived this plan to be particularly useful for the task. A second limitation involves the shared sampling technique across Studies 1-4, although recent work suggests that the quality of data collected MTurk participants is at least as reliable as those obtained via more common methods (Buhrmeister, Kwang, & Gosling, 2011; Casler, Bickel, & Hackett, 2013), and draws from larger participant pools than university student samples (Stewart et al, 2015). We attempt to address this limitation using a new task and sample drawn from a different population in Study 5.

## Study 5

### Methods

**Participants**

Consistent with a power analysis based on prior findings, we recruited an age-diverse community sample of $N = 160$ participants. The sample included University of Zurich students and staff, as well as older adults who attended local physical activity centers. The control group included $n = 41$ participants, and the backup plan group included $n = 119$. We randomly assigned using a 3:1 backup plan-to-control-group ratio to facilitate later tests of group differences between backup-plan group participants who did or did not elect to use backup plans during the task. Sex and age did not significantly differ across groups. Because Switzerland is largely ethnically homogenous, we did not collect data on participant ethnicity. Participants were compensated 5 Swiss Francs (approximately $5 USD).

**Procedure**

The aim of Study 5 was to replicate and extend the findings from Studies 1-4 using a different task, a different goal domain, and a sample drawn from a different population. After providing consent and completing demographic information, participants were instructed that the task involved throwing balls into a bucket three meters away. Control group participants were provided with 15 Ping-Pong balls. In a first "practice" phase, participants threw five balls. In the next "trial" phase, participants threw the remaining 10 balls. Participants were instructed to try to land the balls in the bucket as often as possible. All throws were made seated, underhanded, and using participants' dominant hand.

Participants assigned to the backup plan condition were provided with 15 Ping-Pong and 15 tennis balls. A research assistant informed participants that for this task, Ping-Pong balls are the Plan A, and tennis balls were the backup plan. The practice phase still involved attempting five throws; however, participants could allocate practice attempts across one or both balls (e.g.,

four Ping-Pong balls and one tennis ball). The trial phase involved 10 throws; however, consistent with its status as the "Plan A ball," participants were required to use the Ping-Pong ball for their first throw. For all other throws, participants chose which type of ball to use. After the trial phase, all participants were debriefed and compensated. There were no missing data as all participants attempted 15 throws.

**Measures**

*Backup plan developing costs*

Only participants in the backup plan condition could invest in developing backup plans, which we defined as the number of tennis balls thrown during the practice phase. About two thirds of the participants in the backup plan group (67.2%) chose to practice at least once with a tennis ball. On average, participants threw about 2 (of a possible 5) tennis balls during the practice phase (M = 1.89, SD = 1.27), for an average backup plan development ratio of .37 (*SD* = .25). Sex ($r$ = -.05, $p$ = .63) and age ($r$ = -.06, $p$ = .54) were not associated with backup plan development ratio.

*Perceived throwing ability*

Participants who perceived that they were not skilled in ball throwing may invest more in their backup plans. Therefore, we assessed perceived ability with a single item ("How good are your ball throwing skills?") with response options ranging very bad (0) to very good (6; *M* = 2.82, *SD* = 1.15).

*Replacing with backup plans*

The number of times participants changed balls during the task served as an index of replacing. More than two-thirds (67.2%) of participants changed balls at least once, and on

average, participants changed balls about once ($M$ = 1.15, $SD$ = 1.38). Sex ($r$ = -.08, $p$ = .38) and age ($r$ = .03, $p$ = .75) were not associated with number of ball changes.

*Performance*

Performance was measured as the number of balls successfully thrown into the bucket during the trial phase (min = 0, max = 10).[3] Overall, participants were successful in about a third of their throws ($M$ = 3.25, $SD$ = 1.98). Age was not significantly associated with performance ($r$ = -.14, $p$ = .08); however, men ($M$ = 3.89, $SD$ = 2.16) performed significantly better than women ($M$ = 2.89, $SD$ = 1.78), $t(158) = 3.16$, $p < .01$, $d$ = .52 (90% CI = .24, .80).

## Results

### Effects of backup plans on Plan A commitment and performance

There were no significant differences in performance between the control ($M$ =3.49, $SD$ = 1.99) and backup-plan groups ($M$ =3.17, $SD$ = 1.97), $t(158) = -.89$, $p$ = .37, $d$ = .16 (90% CI = .12, .45). Therefore, simply having a backup plan available was not associated with decreased ball-throwing performance. We next tested whether backup plan group participants who changed balls during the task performed significantly worse than those who did not change balls, as well as worse than participants in the control group. A between-groups ANOVA indicated significant differences in performance, $F(2, 157) = 3.96$, $p$ = .02, $\eta$= .05 (90% CI =.004, .11). Post-hoc Tukey tests indicated that participants in the backup condition who switched balls performed significantly worse ($M$ = 2.84, $SD$ = 1.97) than participants in the backup condition who did not ($M$ = 3.85, $SD$ = 1.83). Although control group participants had a higher average score ($M$ = 3.49, $SD$ = 1.99) than participants in the backup condition who switched balls, this difference was not statistically significant.

Consistent with prior findings, we expected that investing more in backup plans was associated with decreased commitment to Plan A, and thus more ball switches during the task. This expectation was supported. Investing more in practicing tennis balls was associated with more ball switching ($\beta = .20$, $p = .01$, $R^2 = .04$).

We next tested whether highly investing in developing a backup plan was indirectly associated with performance, through the indirect effect of decreased commitment to Plan A (i.e., switching from throwing with Ping-Pong to tennis balls). This model controlled for the effect of sex on throwing performance and is presented in Figure 6. Across 10,000 bootstrapped samples, standardized indirect effect of backup plan development ratio on trial throw score (controlling for the effect of sex), through the effect of increased ball switching was -.07, p = .01; 95% CI = -.13, -.02. Consistent with earlier studies, the effect size was small-to-medium ($\kappa^2$=.08).[4]

**Testing alternative hypotheses**

Prior trial throw performance is a reasonable but not mutually exclusive alternative hypothesis for why participants replace their Plan A. We assessed this possibility through an HLM analysis. We expected that participants were more likely switch to tennis balls if they had highly invested in practicing them, even after accounting for the effect prior throw performance. The results for this analysis are summarized in Table 2. As hypothesized, participants were more likely to change balls when they invested more in practicing backup tennis ball throws, even when controlling for the effects of prior performance. In addition, we also examined the alternative hypothesis that perhaps participants who perceive that they are unskilled in throwing invest more in developing their tennis ball backup plan, undermining their ping-pong throwing performance. Although participants' perceptions of their throwing abilities were relatively

accurate predictions of their actual performance (β = .20, $p$ = .01, $R^2$ = .04), we did not find support for this alternative hypothesis in these data. Perceived throwing ability was neither associated backup plan development ($r$ = -.01, $p$ = .56) nor use ($r$ = .01, $p$ = .93).

## Discussion

Study 5 replicates findings from Studies 1-4 and extends them into a third goal domain. Consistent with our hypotheses, highly investing in developing backup plans may harm performance through decreasing a person's commitment to continue using Plan A. The control group and backup-plan group participants did not significantly differ in their performance. This suggests that, like Study 3, simply having a backup plan available may not affect performance. Instead, intentionally heavily investing in a backup plan may lead to performance declines through the effect of increased Plan A replacing.

One limitation of these findings was that they could overemphasize the recalibration costs associated with replacing Plan A with a backup plan. Although recalibration costs exist when replacing many backup plans (e.g., if the earlier-described post-doc switches to her notes, she must pause to find her place), adjusting from throwing a light Ping-Pong ball to a heavier tennis ball could result in initial tennis throws being especially off the mark. A second limitation was that we did not collect data on participants perceived ball instrumentality. Future work should directly examine whether the indirect effect of perceived backup plan instrumentality on performance also extends into this task.

## General Discussion

Can making backup plans subvert goal pursuits rather than support them? If so, what processes underlie this effect? This set of five studies provides initial and primarily correlative evidence that highly investing in developing backup plans can negatively affect performance

through decreasing the perceived instrumentality of and commitment to continue using one's Plan A. Consistent with our first hypothesis, the more people invested into developing a backup plan, the more likely they were to use it. This result was consistent regardless of whether the development resources were restricted (Studies 2 & 5) or unrestricted (Study 3), or when greater costs were subjectively perceived rather than objectively reflective of time investment (Study 1). This consistent effect raises the question of whether participants might have simply been accurate in predicting their need for a backup plan; or alternatively, as we are suggesting, that their greater costs of developing backup plans lead them to use them at some point during the goal pursuit (maybe with thoughts along the lines of "now that I have spent time and effort into Plan B, I'd better use it"). In some ways, our interpretation resembles Popper's *Oedipus effect* (1950), which asserts that a person's *inaccurate* predictions can sufficiently shape their behavior such that, over time, the once-false outcome is made manifest.

The results pertaining to Hypothesis 1 provide initial correlative evidence for the *backup plan paradox*, a situation where investing (or perceiving to invest) in developing a backup plan facilitates precisely the conditions that the backup plan was developed to potentially address; that is, Plan A's (perceived or actual) insufficiency. This effect, in turn, is associated with increased backup plan use. The backup plan paradox may be most prominent when resources are restricted and investments in developing backup plans come at the direct expense of maximally developing Plan A. However, the paradox may also apply for goals where resources are less restricted, when, given the availability of a well-developed Plan B, a person's commitment to continue using Plan A decreases after setbacks. Finally, Study 1 suggests that a person's subjective perceptions of high backup plan development investment, even when objective investments are held constant, may also increase the likelihood of backup plan use.

Through what motivational processes might highly-invested backup plans decrease the commitment to continue using Plan A? One possibility involves the costs of keeping a backup plan in reserve. This process of "shielding" (Shah et al, 2002) oneself from the temptation of premature backup plan use can introduce costs into goal pursuit (Napolitano & Freund, 2016). Reserving a backup plan that one perceives to be instrumental may be especially challenging. For example, recalling the concept of an *action crisis* (Brandstätter, Herrmann, & Schüler, 2013), commitment to continue using Plan A may be compromised when, at a crossroads, a person actively compares their current meager returns with the potentially-better returns from their reserved backup plan. The current studies provide provisional support for this idea. In line with our second hypothesis, we found initial evidence (in Studies 2 – 4) that the more instrumental participants perceived their backup plans to be, the more likely the backup plan would later be used. Truly testing the effects of perceived backup plan instrumentality on reserving costs requires future research to first link developing costs to instrumentality perceptions, and then to dynamically assess change in perceived instrumentality for one's backup plan(s) as well as Plan A.

Consistent with our third hypothesis, we found initial support that the more participants elected to replace Plan A with backup plans, the worse they performed. Our interpretation of these findings combines the practical and motivational costs of replacing with backup plans. Practically, changing means during goal pursuit involves some degree of recalibration. Motivationally, replacing Plan A with a backup plan may be akin to admitting failure, and the resulting negative affect may subvert Plan B's instrumentality (Fishbach. Shah, & Kruglanski, 2004). Taken together, these findings illustrate first support for our fourth hypothesis: that highly

investing in developing backup plans and perceiving backup plans to be instrumental could, in some situations, *indirectly* affect performance through the effect of increased backup plan use.

**Beyond the dilution model: The unique costs (and benefits) of backup plans**

In some cases, the results from research on concurrent means mirror those reported here. For example, concurrently using multiple means decreases commitment to any one means (Kruglanski et al, 2011) and has been shown to lower perceptions of other means' instrumentality (Bélanger et al, 2015). These results are typically described in terms of a *dilution model.* The model's rationale is zero-sum in nature: there is a fixed, constant level of "connection strength" in means-ends relations. Adding more concurrent means into a goal pursuit dilutes connection strength across means, and decreases perceptions of the instrumentality of all means.

The dilution model is a useful first step in understanding the effects of backup plans on goal pursuit. It is a reasonable framework for explaining Study 2's results, where investments in backup plans came at the direct cost of investments in Plan A. However, the dilution model does not adequately explain all of the results presented here. For example, consider two hypothetical participants in Study 3. Linda reads the full seven information pieces for her Plan A, and reads six for her backup plan. Marc reads the full seven information pieces for his Plan A, three for his backup plan, and one for his "Plan C". By the logic of the dilution model, Marc should be less committed to his goal – he has more available means than Linda. However, our results suggest that Linda is less committed to her Plan A. She has intentionally invested in more robustly developing her backup plan, and if Plan A faces setbacks, she may be more likely to switch to Plan B.

**Interindividual and intraindividual differences in the usefulness of backup plans**

This research represents the first attempt to systematically study the potentially negative effects of backup plans. Its focus was limited to describing behavioral and motivational processes that underlie these potential negative effects. There are likely interindividual and intraindividual differences that moderate the relations between investments in developing backup plans, commitment to one's Plan A, and performance. For example, people who are more conscientious (e.g., Roberts, Jackson, Fayard, Edmonds, & Meints, 2009) may better regulate the costs of developing backup plans. Those who are more ruminative (Schultheiss, Jones, Davis, & Kley, 2008) may further fall behind in goal pursuits if they deliberate longer about whether or not to replace their Plan A. Those who engage in defensive pessimism (Norem & Cantor, 1986) may invest more in backup plans, and could disengage from Plan A quickly after initial setbacks to protect self-esteem, but perhaps at the cost of decreased performance. In contrast, people construing their goals at a high level, rather than focusing on the low-level details of Plan A's temporary shortcomings, may delay backup plan use (Fujita & Carnevale, 2012). Older adults with great goal-relevant practical intelligence (e.g., Wagner & Sternberg, 1985) and motivational competence (Freund, Nikitin, & Riediger, 2012) might adeptly use backup plans, sidestepping the backup plan paradox introduced earlier.

**Intentionality**

In our conceptualization, backup plans are intentionally developed and assigned their Plan B status. However, in some of the current studies, we instructed participants to develop a backup plan and, in others, we assigned which means were to be regarded as Plan A and/or Plan B. This was necessary to keep experimental control over different aspects of the paradigms we used, and did not allow us to fully test the role of intentionality for the development and use of backup plans. However, in all studies, participants with backup plans intentionally chose to keep

them reserved to varying degrees, and intentionally chose to replace Plan A with the backup plan (or not). In addition, results from Study 5 suggest that it is not the mere affordance of a backup plan that drives performance decrements, but rather instead participants' intentional investments into developing them that could threaten performance.

**Limitations**

This research is a first step in empirically investigating the potentially-undermining effects of backup plans. Beyond the key study-specific limitations described in the respective study discussions, an overarching limitation of this research was that we did not successfully implement an experimental design in all studies. Thus, our interpretations are based, in large part, on correlational analyses. This constrains causal interpretations of the results. We cannot conclusively argue that backup plan investments, themselves, account for the effects we described here.

Future work using robust experimental designs must incorporate between- and within-person characteristics that may moderate the relations between back up plan investment and use, while also navigating the challenges of intentionality and random assignment. Adapting the methodology presented in Study 1 to limit selective attrition may prove viable; however, future experimental work that incorporates "real world" goal pursuits will likely provide ideal balance of internal and external validity necessary to properly explore the potentially-harmful effects of backup plans.

**Conclusions**

Despite the limitations noted above, this research addresses a gap in the self-regulation literature, and provides initial correlative evidence for the backup plan paradox. The more

participants invested in developing backup plans, the more likely they were to later use those backup plans. Using backup plans was, in turn, associated with decreased performance.

We do not suggest that backup plans are a universally-negative influence on a person's goal striving. In fact, we assume that there are instances backup plans improve one's chances of efficiently achieving a goal, or are essential for prudent goal striving. The interesting question is under which conditions backup planning is beneficial and under which it is maladaptive. Future experimental work needs to discern for which people, at which periods in the life span, pursuing which goals, in which contexts, is investing in a backup plan an adaptive self-regulatory action.

References

Bargh, J. A., Schwader, K. L., Hailey, S. E., Dyer, R. L., & Boothby, E. J. (2012). Automaticity in social-cognitive processes. *Trends in Cognitive Sciences*, *16*, 593–605. doi:10.1016/j.tics.2012.10.002

Bélanger, J. J., Schori-Eyal, N., Pica, G., Kruglanski, A. W., & Lafrenière, M.-A. (2015). The "more is less" effect in equifinal structures: Alternative means reduce the intensity and quality of motivation. *Journal of Experimental Social Psychology*, *60*, 93–102. dx.doi.org/10.1016/j.jesp.2015.05.005

Brandstätter, V., Herrmann, M., & Schüler, J. (2013). The struggle of giving up personal goals: Affective, physiological, and cognitive consequences of an action crisis. *Personality and Social Psychology Bulletin*, *39*, 1668–1682. doi:10.1177/0146167213500151

Buhrmeister, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science, 6,* 3-5. doi: 10.1177/1745691610393980

Casler, K**.,** Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior, 29,* 2156-2160. doi:10.1016/j.chb.2013.05.009

Etkin, J., & Ratner, R. K. (2013). Goal pursuit, now and later: Temporal compatibility of different versus similar means. *Journal of Consumer Research*, *39*, 1085–1099. http://doi.org/10.1086/667203

Etkin, J., & Ratner, R. K. (2012). The dynamic impact of variety among means on motivation. *Journal of Consumer Research*, *38*, 1076–1092. http://doi.org/10.1086/661229

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160. dx.doi.org/10.3758/brm.41.4.1149

Fishbach, A., Shah, J. Y., & Kruglanski, A. W. (2004). Emotional transfer in goal systems. *Journal of Experimental Social Psychology*, *40*, 723–738. http://dx.doi.org/10.1016/j.jesp.2004.04.001

Freund, A. M. (2008). Successful aging as management of resources: The role of selection, optimization, and compensation. *Research in Human Development*, *5*, 94-106. doi:10.1080/15427600802034827

Freund, A. M., Nikitin, J., & Riediger, M. (2012). Successful aging. In R. M. Lerner, M. A. Easterbrooks, & J. Mistry (Eds.), *Handbook of Psychology, Vol. 6. Developmental Psychology* (2nd ed.; pp. 615–638). New York, NY: Wiley. doi:10.1002/0471264385.wei0625

Fujita, K., & Carnevale, J. J. (2012). Transcending temptation through abstraction: The role of construal level in self-control. *Current Directions in Psychological Science, 21*, 248-252. doi:10.1177/0963721412449169

Garland, H., & Newport, S. (1991). Effects of absolute and relative sunk costs on the decision to persist with a course of action. *Organizational Behavior and Human Decision Processes*, *48*, 55–69. dx.doi.org/10.1016/0749-5978(91)90005-e

Gendolla, G. H. E. (2012). Implicit affect primes effort: Theory and research on cardiovascular response, *International Journal of Psychophysiology*, 86, 123-135. doi: 10.1016/j.ijpsycho.2012.05.003

Gollwitzer, P. M. (1999). Implementation intentions: strong effects of simple plans. *American Psychologist*, *54*, 493. doi:10.1037/0003-066X.54.7.493

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*, 408-420. doi: 10.1080/03637750903310360

Huang, S. C., & Zhang, Y. (2013). All roads lead to Rome: The impact of multiple attainment means on motivation. *Journal of Personality and Social Psychology*, *104*, 236-248. doi:10.1037/a0031039

Iles, R., Judge, T. A., & Wagner, D. T. (2010). The influence of cognitive and affective reactions to feedback on subsequent goals: Role of behavioral inhibition/activation. *European Psychologist, 15*, 121-131. doi:10.1027/1016-9040/a000011

Kruglanski, A. W., Pierro, A., & Sheveland, A. (2011). How many roads lead to Rome? Equifinality set-size and commitment to goals and means. *European Journal of Social Psychology*, *41*, 344-352. doi:10.1002/ejsp.780

Kruglanski, A. W., Chernikova, M., Babush, M., Dugas, M., & Schumpe, B. M. (2015). The architecture of goal systems. *Advances in Motivation Science*, *2*, 69–98. http://dx.doi.org/10.1016/bs.adms.2015.04.001

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., Ly, A., Gronau, Q. F., Smira, M., Epskamp, S., Matzke, D., Wild, A., Rouder, J. N., Morey, R. D. & Wagenmakers, E.-J. (2015). JASP (Version 0.7) [Computer software].

Muthén L., & Muthén B. (1998-2015). Mplus user's guide. 7th ed. Los Angeles, CA: Muthén & Muthén.

Napolitano, C. M., & Freund, A. M. (2016). On the use and usefulness of backup plans. *Perspectives on Psychological Science, 11*, 56-73. dx.doi.org/10.1177/1745691615596991

Norem, J. K., & Cantor, N. (1986). Anticipatory and post hoc cushioning strategies: Optimism and defensive pessimism in "risky" situations. *Cognitive Therapy and Research*, *10*, 347–362. http://dx.doi.org/10.1007/bf01173471

Popper, K. (1950). Indeterminism in quantum physics and in classical physics, Part I. *The British Journal for the Philosophy of Science, 1,* 117-133. dx.doi.org/10.1093/bjps/i.3.173

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, *16*, 93–115. http://doi.org/10.1037/a0022658

Roberts, B.W., Jackson, J.J., Fayard, J.V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M.R. Leary & R.H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 369-381). New York: Guilford.

Schultheiss, O. C., Jones, N. M., Davis, A. Q., & Kley, C. (2008). The role of implicit motivation in hot and cold goal pursuit: Effects on goal progress, goal rumination, and emotional well-being. *Journal of Research in Personality*, *42*, 971-987. doi:10.1016/j.jrp.2007.12.009

Shah, J. Y., Friedman, R., & Kruglanski, A. W. (2002). Forgetting all else: On the antecedents and consequences of goal shielding. *Journal of Personality and Social Psychology*, *83*, 1261–1280. doi:10.1037/0022-3514.83.6.1261

Shah, J. Y., & Kruglanski, A. W. (2002). Priming against your will: How accessible alternatives affect goal pursuit. *Journal of Experimental Social Psychology*, *38*, 368–383. dx.doi.org/10.1016/s0022-1031(02)00005-7

Sniehotta, F. F., Schwarzer, R., Scholz, U., & Schüz, B. (2005). Action planning and coping planning for long-term lifestyle change: theory and assessment. *European Journal of Social Psychology*, *35*(4), 565–576. doi:10.1002/ejsp.258

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16*, 27-44. doi:10.1016/0030-5073(76)90005-2

Stewart, N., Ungemach, C., Harris, A. J., Bartels, D. M., Newell, B. R., Paolacci, G., & Chandler, J. (2015). The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers. *Judgment and Decision Making*, *10*(5), 479–491.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology*, *49*, 436. doi: 10.1037/0022-3514.49.2.436

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkens

Zhang, Y., Fishbach, A., & Kruglanski, A. W. (2007). The dilution model: How additional goals undermine the perceived instrumentality of a shared path. *Journal of Personality and Social Psychology*, *92*(3), 389–401. dx.doi.org/10.1037/0022-3514.92.3.389

Zhou, H., & Fishbach, A. (2016). The pitfall of experimenting on the web: How unattended selective attrition leads to surprising (yet false) research conclusions. *Journal of Personality and Social Psychology, 111*, 493-504.

## Footnotes

1. We interchangeably use “Plan B” and “backup plan.”

2. For all studies, we present results from a subsample of the data collected for this research. Excluded from this report are exploratory data regarding participants’ demographic and other individual characteristics (e.g., number of children, life satisfaction).

3. Due to practical limitations, we did not include a control condition where participants only threw tennis balls. This group would allow for direct comparison of the instrumentality of tennis and Ping-Pong balls, and its exclusion is a limitation of this research. However, there is evidence that tennis and Ping-Pong balls were similarly instrumental for this task. Participants made 27.63% of Ping-Pong trial throws. Among backup plan group participants who elected to use tennis balls during the trial phase (n = 80), 28.39% of throws were successful.

4. Including participants’ test-phase throwing performance as a proxy control variable for “throwing ability” did not produce meaningfully different results. Although test performance was significantly associated with trial performance ($\beta = .30$, $p < .001$), the standardized indirect effect of backup plan development ratio on trial throw score (controlling for the effect of sex and practice phase throw score) through the effect of increased ball switching was similar: -.07, p = .01; 95% CI = -.13, -.02. The effect size was small-to-medium ($\kappa^2$=.08).

Table 1

*Demographic and Descriptive Information on Participants in Studies 1-5*

| *Study* | *N* | *Age M, SD* | *Age range* | *% Female* | *% White* | *% Latino* | *% Black* | *% Asian* | *% other ethnicity* |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 160 | 36.62, 11.17 | 21-71 | 42.5% | 82.5% | 5.6% | 6.3% | 7.5% | 1.9* |
| 2 | 247 | 42.74, 15.52 | 18–81 | 55.0 | 80.6 | 3.6 | 7.3 | 4.9 | 3.6 |
| 3 | 248 | 43.25, 15.30 | 18–80 | 56.0 | 82.3 | 2.8 | 6.9 | 3.2 | 4.8 |
| 4 | 204 | 32.25, 9.74 | 19-62 | 58.3 | 78.9 | 3.9 | 7.4 | 7.8 | 2.0 |
| 5 | 160 | 50.32, 24.43 | 18-93 | 64.4 | ** | ** | ** | ** | ** |

*Note.* * = Participants in Study 1 could select multiple ethnicities

** = Participants in Study 5, collected in Switzerland, did not provide data on their ethnicity

Table 2

*Results of Hierarchical Linear Models used in Studies 3, 4, and 5 Predicting the Likelihood of Replacing Plan A with a Backup Plan*

| Variable | Study 3 Likelihood of Replacing Plan A | | | | Study 4 Likelihood of Replacing Plan A | | | | Study 5 Likelihood of Replacing Plan A | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *b* | *SE* | *p* | 95% CI | *b* | *SE* | *p* | 95% CI | *b* | *SE* | *p* | 95% CI |
| Intercept | -1.40 | 0.20 | .000 | 1.79, -1.01 | -0.75 | 0.22 | .000 | -1.18, -0.32 | -2.44 | 0.22 | .000 | -2.87, -2.01 |
| Backup plan development ratio | 4.60 | 0.75 | .000 | 3.13, 6.07 | | | | | 0.23 | 0.09 | .009 | 0.05, 0.41 |
| Overall development | -0.17 | 0.04 | .000 | -0.25, -0.09 | | | | | | | | |
| Backup plan's perc. instrum. | 0.32 | 0.10 | .007 | 0.12, 0.52 | 0.31 | 0.12 | .001 | 0.07, 0.55 | | | | |
| Prior performance | -0.83 | 0.17 | .000 | -1.16, -0.50 | -1.54 | 0.25 | .000 | -2.03, -1.05 | -0.64 | 0.21 | .003 | -1.05, -0.23 |

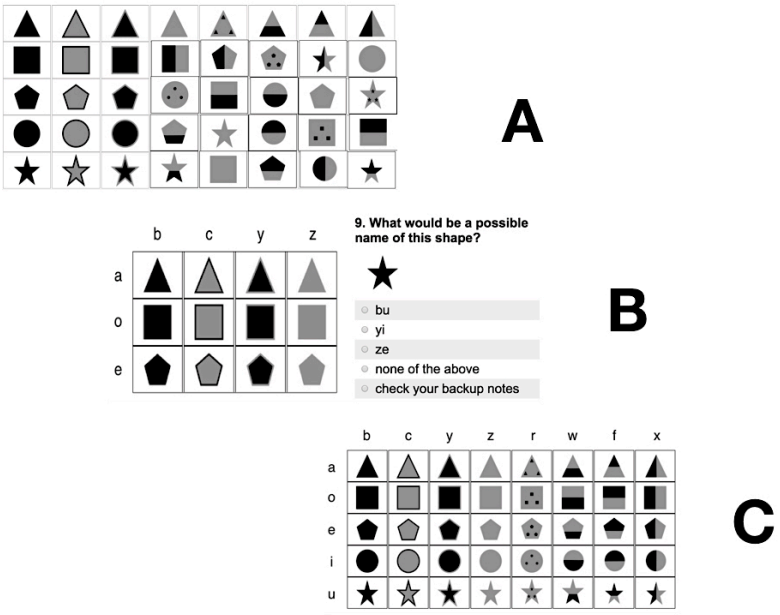


*Figure 1*. Simplified depiction of Study 1's procedure. **A** = *Developing phase*. Participants unscramble the grid to develop their backup plan; the top row and leftmost three columns are locked in the correct position. **B** = *Trial phase, Plan A*. Participants attempt to provide possible name for shape. The small grid to the left provides sufficient information to deduce the correct answer. The correct response is "bu:" the pattern is all black, and the all black pattern must begin with "b." Correct responses using logic earn participants $.10. **C** = *Backup plan use*. If participants select "Check your backup notes," they view the backup plan grid. Correct responses after using the backup plan earn participants $.07.

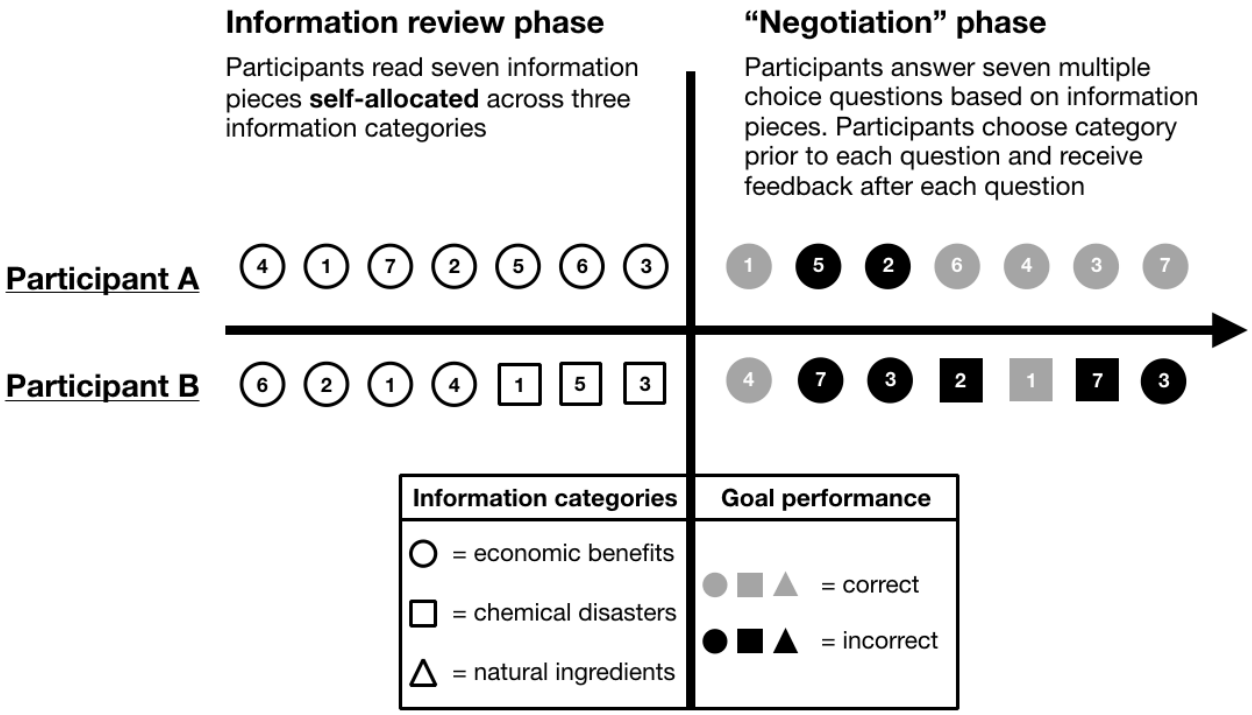


*Figure 2*. Illustration of Study 2's simulated negotiation procedure, with example data. Numbers indicate randomly-ordered information piece or simulated negotiation question, per category. Participant A prototypically maintains commitment to Plan A despite initial setbacks, whereas Participant B invests in developing backup plans and later prototypically replaces with backup plans.

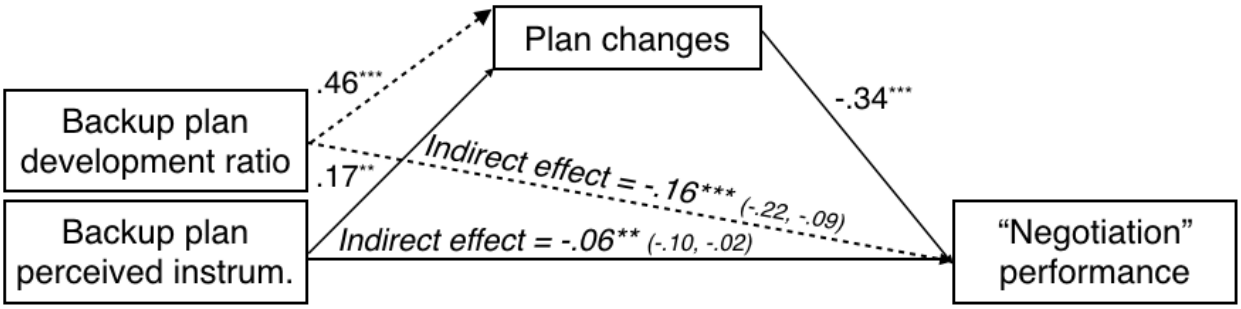


*Figure 3.* Study 2 simultaneous standardized indirect effects of backup plan perceived instrumentality and backup plan development ratio on performance. Direct effects are *ns*.

*** = $p$ <.001; ** = $p$ <.01

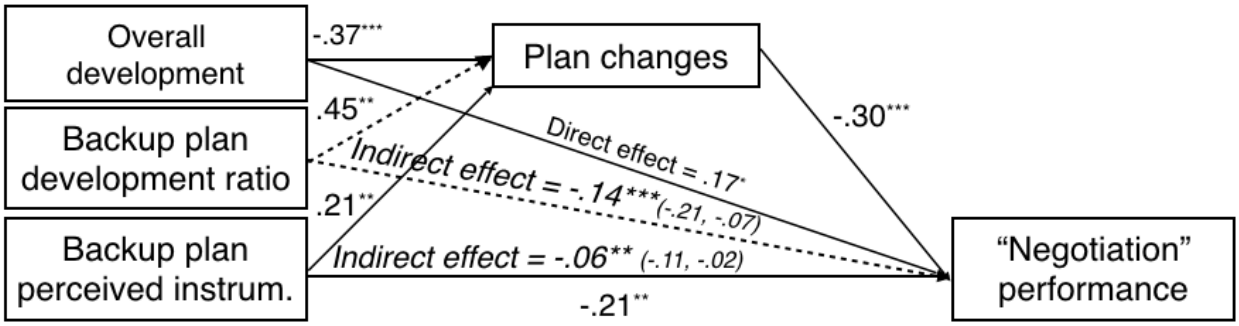


*Figure 4.* Study 3 simultaneous standardized indirect effects of backup plan perceived instrumentality and backup plan development ratio on performance, controlling for overall development. Direct effect of backup plan development ratio is *ns*.

*** = $p < .001$; ** = $p < .01$; ** = $p < .05$

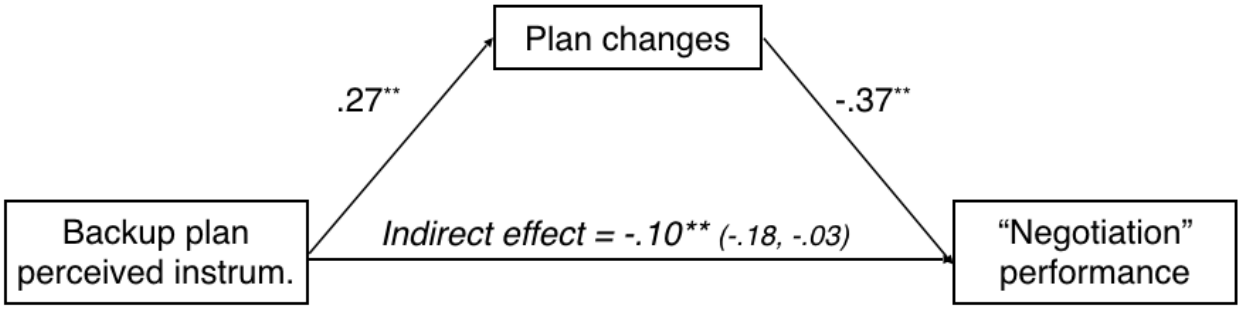


*Figure 5.* Study 4 standardized indirect effect of backup plan perceived instrumentality on performance. Direct effect is *ns.* Included but not displayed is the significant association of age on backup plan perceived instrumentality (β = -.23, $p$ = .02).

** = $p$ <. 01

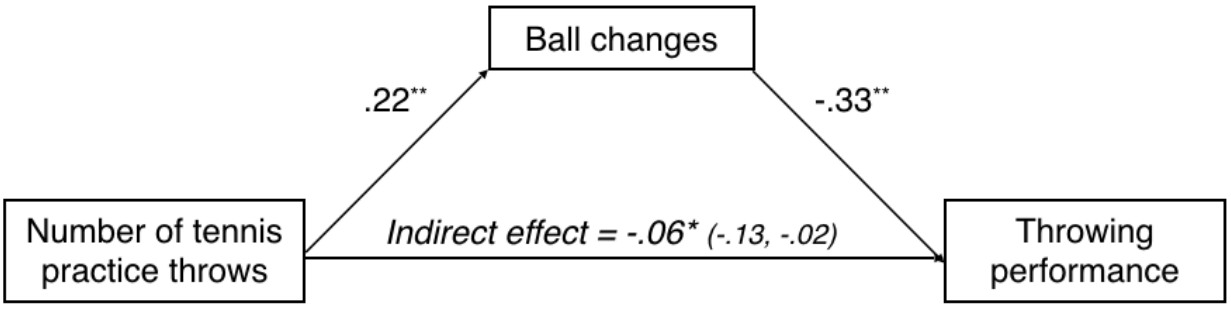


*Figure 6*: Study 5 standardized indirect effect of number of tennis practice throws on throwing performance, through the effect of increased ball changes. Direct effect is *ns.* Included but not displayed is the significant association of sex on throwing performance (β = -.23, $p$ = .02). **= $p$ <. 01, **= $p$ <. 05.