27

Metacognitive Control of Memory Reporting

Morris Goldsmith, Ainat Pansky, and Asher Koriat

Two eyewitnesses are asked to pick the perpetrator of a crime out of a police lineup. One witness is told to keep in mind that the actual perpetrator may not be present in the lineup, and that it is perfectly acceptable to respond “don’t know.” The other witness is simply told to indicate whether one of the persons standing in the lineup is the perpetrator. If you were a falsely accused suspect, in which of these lineups would you prefer to appear? If you are a judge, faced with two different suspects picked out in the different line-ups, which of the two witnesses’ judgments would you put more faith in?

Two students take a 5-altemative multiple-choice exam under formula scoring, in which a *V4* point penalty is paid for each wrong answer. One student answers all 100 questions but 20 of those answers are wrong. The other answers only 75 questions, all of which are correct. Both students, then, will receive the same score on the exam (75). Nevertheless, is there perhaps a substantial difference between the abilities of these two students that the equivalent test score might be hiding? Two other students take a law exam. One writes long and relatively detailed answers, but many of the added details are wrong. The other is more careful in monitoring the correctness of her answers, providing less detailed responses, but all of the provided information is correct. Which student would you prefer to hire as a law clerk in your firm?

These examples illustrate the important role played by metacognitive processes that intervene between the retrieval of information from memory and the decision to volunteer that information and perhaps act on it (for a review of metacognitive contributions to retrieval itself, see Koriat, Goldsmith, & Halamish, 2008). After reading this chapter, it should become clear that the amount and accuracy of the knowledge and information that people convey from memory depend not only on memory processes per se, but also on the operation of metacognitive monitoring and control processes that are used in the strategic regulation of memory reporting.

In what follows, we present an overview of a metacognitive framework that was developed for investigating this regulation. We then use this framework as a backdrop for a selective review of experimental work on the control of memory reporting - its mechanisms and performance consequences - in both theoretical and applied research con- texts, with a special focus on the area of eye- witness memory. Finally, we point to some further directions in which the framework might be extended and applied.

A METACOGNITIVE FRAMEWORK OF MEMORY REPORTING

The storehouse metaphor of memory, which has guided much of traditional memory research (Koriat & Goldsmtih, 1996a), implies a clear goal for the rememberer: to reproduce as much of the originally stored information as possible. This is the essence of the instructions provided to participants in typical list-learning experiments. Growing interest in real-life memory phenomena over the past few decades, however, has led to a greater emphasis on the functions of memory in real-life contexts and on the active role of the rememberer in putting memory to use in the service of personal goals (e.g., Neisser, 1988, 1996). The goals of remembering in everyday life are complex and varied, and may be partially or wholly conflicting. Hence, a great deal of skill and sophistication may be required of the rememberer in negotiating between the different goals and in finding an expedient compromise.

Two prominent and generally conflicting memory goals are informativeness (to provide as much information as possible) and accuracy (to avoid providing wrong information). Consider, for example, a courtroom witness who has sworn "to tell the whole truth and nothing but the truth." To avoid false testimony, the witness may choose to refrain from providing information that she feels unsure about. This, however, will tend to reduce the amount of information that she provides the court, thereby com- promising the oath to tell the "whole" truth. Alternatively, she may choose to phrase her answers at a level of generality at which they are unlikely to be wrong (Neisser, 1988). Once again, however, the increased accuracy will come at the expense of informative- ness. This example illustrates two general means by which rememberers regulate their memory reporting in the wake of generally competing demands for accuracy and informativeness (Goldsmith & Koriat, 2008): The first, *control of report option*, involves the decision to volunteer or withhold particular items of information, with the option to respond "don't know" to specific questions. The second, *control of grain size,* involves choosing the level of precision or coarseness of an answer when it is provided. Each of these will be considered in turn.

Control of report option

Koriat and Goldsmith (1996b) put forward a simple model of how metamemory processes are used to regulate memory accuracy and quantity performance under free-report conditions, that is, when one is free to choose which items of information to report and which to withhold (see Figure 27.1). The model is deliberately schematic, focusing on the manner in which metacognitive processes at the reporting stage affect the ultimate memory performance. Thus, in addition to an unspecified retrieval mechanism, a monitoring mechanism is used to subjectively assess the correctness of potential memory responses, and a control mechanism then determines whether or not to volunteer the best available candidate answer. The control mechanism operates by setting a report criterion on the monitoring output: The answer is volunteered if its assessed probability of being correct passes the criterion, but is withheld otherwise. The criterion is set on the basis of implicit or explicit payoffs, that is, the perceived gain for providing correct information relative to the cost of providing wrong information.

The basic implication of the model is that when given the opportunity to do so, rememberers can enhance the accuracy of the information that they report by withholding answers that are likely to be wrong. Such enhancement, however, is subject to a *quantity-accuracy trade-off* In general, raising the report criterion should result in fewer volunteered answers, a higher *percentage* of which are correct (increased accuracy), but a lower *number* of which are correct (decreased quantity). Because of this trade-off, the strategic control of memory performance

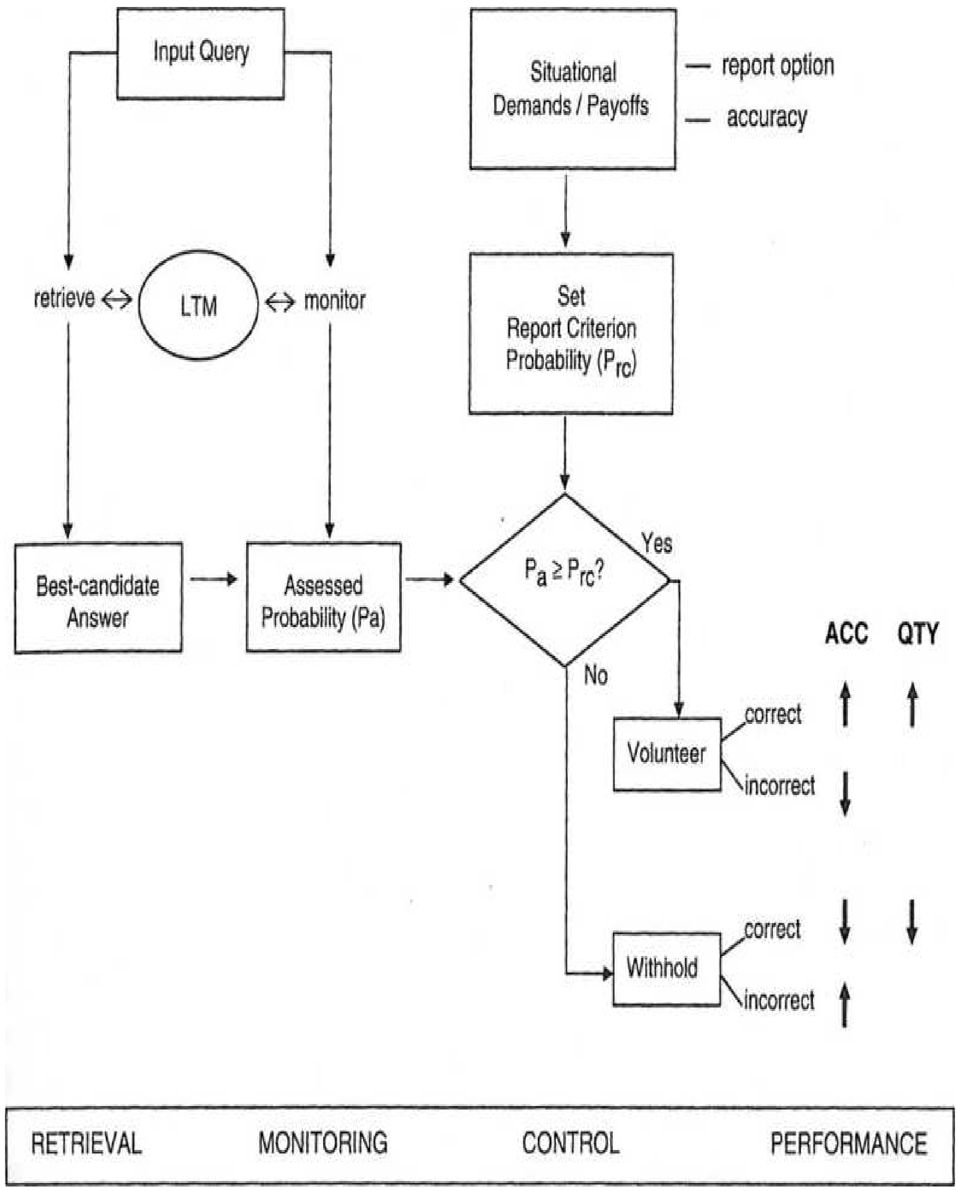

Figure 27.1 A schematic model of the strategic regulation of memory accuracy and memory quantity performance, utilizing the option of free report. The upward and downward pointing arrows on the right of the figure signify positive and negative performance outcomes. Adapted from Koriat and Goldsmith, 1996b

requires the rememberer to weigh the relative payoffs for accuracy and quantity in reaching an appropriate criterion setting.

Although the model is quite simple, its implications for the determinants of free-report memory performance are not.

According to the model, such performance should depend not only on "memory" per se, but also on the operation and effective- ness of the metacognitive monitoring and control processes that mediate between the retrieval of information on the one hand, and the reporting (or withholding) of that information on the other. Before we can flesh out these implications and the empirical evidence to support them, we first briefly describe the basic research paradigm and assessment methodology that was developed to examine the individual cognitive and metacognitive components of the model.

Quantity-Accuracy Profile (QAP) methodology

The Koriat and Goldsmith (1996b) frame- work was developed together with a special experimental paradigm and procedure that combines free and forced reporting with the elicitation of confidence judgments to isolate and assess the cognitive and metacognitive components postulated by the model. In a typical experiment using this paradigm, participants are presented with a series of questions (or retrieval cues) in either a recall or recognition format, and for each question they are asked: (1) to answer the question (forced report), (2) to assess the likelihood that their answer is correct (confidence judgment), and finally, (3) to decide whether or not to report the answer under either an implicit or explicit "payoff' schedule. For example, participants might be told that they will receive one point for each correct reported answer, but lose one point for each wrong reported answer, with points neither gained nor lost for withheld answers.

Rather than evaluating memory performance in terms of a single overall measure (e.g., percent correct), this basic procedure and its variants (see Goldsmith & Koriat, 2008; Higham, 2007) yields a rich profile of measures, including the joint levels of free- report quantity and accuracy performance, and the underlying determinants of this performance: memory retrieval, metacognitive monitoring, and report control. *Memory retrieval* is indexed by the percentage of correct answers under forced report instructions. Metacognitive *monitoring effective- ness* is indexed in terms of both *calibration bias* (over/under-confidence) and monitoring *resolution* (or discrimination accuracy) - the correlation between confidence in one's answers and the actual correctness of those answers. *Control sensitivity* - the extent to which a person's reporting behavior is guided by the output of his or her monitoring process - is indexed by the correlation between confidence in an answer and the decision to report it. Finally, *control policy* (report criterion level) can be estimated by identifying the confidence level above which the participant reports her answers, and below which she withholds them.

A similar methodological approach, based on the same experimental paradigm but using a Type-2 signal-detection framework to conceptualize and measure the monitoring and control aspects, has been put forward by Higham and colleagues (e.g., Higham, 2002, 2007; Higham, Perfect, & Bruno, 2009). A recent exchange on the advantages and disadvantages of each method can be found in Goldsmith (2011) and Higham (2011)

Empirical findings and conclusions: Report option

We now summarize and discuss some of the main findings and conclusions that have emerged with regard to the control of report option.

Rememberers are reasonably successful in monitoring the correctness of their best-candidate answers

A great deal of work has been conducted on the accuracy of metacognitive monitoring from various different perspectives (see Dunlosky & Metcalfe, 2009; Koriat, 2007) For the purpose of choosing which answers to report and which to withhold, the aspect of monitoring that is most crucial for the effective exercise of report option is *monitoring resolution* - the extent to which a person's confidence judgments successfully discriminate correct from incorrect answers. Under experimental conditions in which there has been no deliberate attempt to impair memory monitoring (e.g., by providing post-event misinformation or using misleading/ deceptive/ unanswerable questions; see section on "critical role of monitoring" below), monitoring resolution is generally moderate to high, as indexed by the within-participant Kruskal-Goodman gamma correlation between confidence and actual correctness. This correlation is generally somewhat higher for recall testing than for recognition testing (e.g., Koriat & Goldsmith, 1996b; Robinson, Johnson, & Herndon, 1997). Presumably, recall testing provides the rememberer with an additional effective cue that is diagnostic of accuracy - retrieval fluency. Another reason is that answers held with low confidence may often be correct simply because of the baseline probability of guessing the right answer (see Schwartz & Metcalfe, 1994 for analysis and discussion).

A second aspect of monitoring effectiveness is *calibration bias* (over/under-confidence; Lichtenstein, Fischhoff, & Phillips, 1982) or *absolute* monitoring accuracy (Nelson, 1996). The ubiquitous finding is that remembers are generally overconfident, with the mean assessed probabilities of people's answers substantially higher than the actual proportion correct (see Hoffrage, 2004). The con- sequences of this for report option are that although answers held with high confidence are more likely to be correct than those held with low confidence (reflecting good monitoring resolution), the high-confidence answers may still be more likely to be wrong than the person realizes, thereby lowering the accuracy of the answers that are ultimately reported.

Memory monitoring guides the report control decisions

Consistent with the model, rememberers rely heavily on their subjective confidence in deciding whether to volunteer or withhold an answer. Control sensitivity - the relation- ship between one's level of confidence and the report decision - has been found to be very strong, with within-participant gamma correlations typically on the order of .95 or higher (e.g., Koriat & Goldsmith, 1996b; Mmtzer, Kleykamp, & Griffiths, 2010; Pan- sky, Goldsmith, Koriat, & Pearlman-Avnion, 2009) . Interestingly, this relationship continues to be strong even in cases in which the effectiveness of memory monitoring is poor (e.g., Kelly & Sahakyan, 2003; Koriat, 2011; Konat & Goldsmith, 1996b, Experiment 2; Rhodes & Kelley, 2005; and see section on "critical role of monitoring" below). Rememberers rely blindly on their confidence possibly because they have no access to the diagnosticity of their monitoring or because they have no better alternative.

At the same time, however, control sensitivity has been found to be systematically lower in specific populations such as older adults (Pansky et al., 2009; but see Kelly & ahakyan, 2003) and people with schizophrenia (Damon, Gokalsing, Robert, Massin- auss, & Bacon, 2001; Koren, Seidman, Goldsmith, & Harvey, 2006). Interestingly, control sensitivity was found to correlate with measures of executive functioning (Pansky et al., 2009) and measures of clinical awareness and competence to consent (Koren et al., 2006), suggesting a link between control sensitivity and overall levels of meta- cognitive and executive functioning. The inclusion of control sensitivity as a theoretical component underlying free-report performance, and the examination of potential group and situational differences in such sensitivity, distinguishes the metacognitive approach from the related signal-detection approach (Type-1 or Type-2), in which use of the subjective evidence continuum as the basis for the response decision is axiomatic, and possible variance in control sensitivity is essentially ignored.

Performance consequences: By regulating their own reporting; rememberers substantially enhance the accuracy of the information that they report. This accuracy increase is often achieved at a relatively small cost in the amount of correct reported in formation (quantity-accuracy trade-off).

Perhaps the most basic and robust finding with regard to the consequences of report option for memory performance is that rememberers can in fact enhance their free-report accuracy substantially, relative to forced report (e.g., Higham, 2002; Koriat & Goldsmith, 1994,1996b; Kelley & Sahakyan, 2003). They do so by withholding candidate answers that are likely to be wrong. The potential accuracy gain is particularly high when forced-report accuracy is low. Thus, for example, Koriat and Goldsmith (1996b), using general-knowledge tests of differing difficulty, observed accuracy increases of 29 percentage points (from 47 percent to 76 percent) and 47 percentage points (from 28 percent to 75 percent) in the recall conditions of their Experiments 1 and 2, respectively.

The improved accuracy generally comes at a relatively small decrease in the quantity of correct reported information (e.g., 9 percentage points and 6 percentage points, respectively, in the two recall conditions just mentioned). However, both simulation analyses (Higham, 2011; Koriat & Goldsmith, 1996b) and empirical results indicate that for typical levels of monitoring effectiveness, enhancing one's accuracy becomes relatively costly in terms of quantity performance as the criterion level is raised. Thus, simply giving a person the option of free report may allow a fairly large accuracy improvement to be achieved without much loss of quantity (e.g., Koriat & Goldsmith, 1994, 1996b; Perfect & Weber, 2012; Roebers, Moga, & Schneider, 2001), but placing a larger premium on accuracy leads to a more serious quantity reduction relative to the increased gain in accuracy (e.g., Koriat & Goldsmith, 1994, 1996b; Koriat, Goldsmith, Schneider, & Nakash-Dura, 2001).

The use of report option in the control of memory reporting is strategic

The basic dynamic of a quantity-accuracy trade-off requires rememberers to weigh the potential gain of reporting correct information against the potential penalty for providing wrong information in arriving at an appropriate report criterion for the specific reporting context. In experimental contexts, these incentives are often manipulated in terms of explicit payoffs and penalties for correct and incorrect reported answers. Several studies have found that participants do indeed adjust their report criterion according to the opera- five payoff matrix, setting a more strict report criterion when the motivation for accuracy is higher (e.g., Higham, 2007; Koriat & Gold- smith, 1996b; Koriat et al., 2001).

The critical role of monitoring: Both the accuracy benefits and the quantity costs of self-regulated reporting depend greatly on monitoring effectiveness

Another key implication of the metacognitive model of report control concerns the crucial role of monitoring effectiveness in determining the joint levels of free-report memory accuracy and quantity performance. Clearly some ability to distinguish between correct and incorrect candidate answers is necessary for the control of memory reporting to yield any benefits at all. Moreover, as this ability improves, greater increases in accuracy can be achieved at lower costs in quantity, so that at the extreme, when monitoring effectiveness is perfect, there is no quantity-accuracy trade-off at all. On the other hand, when monitoring ability is poor, the exercise of report option may yield little or no benefit in accuracy, and merely reduce the quantity of correct reported information. In fact, there may even be situations in which participants' monitoring is counter diagnostic, with a negative correlation between subjective confidence and actual accuracy (see Benjamin, Bjork, & Schwartz, 1998, Koriat, 2012). Though presumably rare, in such cases the reporting and withholding of answers on the basis of subjective confidence would be expected to lower both quantity and accuracy.

The crucial role of monitoring effective- ness for the effective use of report option has been elucidated both in simulation analyses (Higham, 2011; Koriat & Goldsmith, 1996b) and in empirical results. For example, Koriat and Goldsmith (1996b, Experiment 2) manipulated participants' monitoring effectiveness by using two different sets of general knowledge recall questions: One set consisted of standard items for which the participants' monitoring was expected to be effective, whereas the other set consisted of "deceptive" items for which the participants' ability to monitor the correctness of their answers was expected to be poor (e.g., "Who composed the Unfinished Symphony"?; see Koriat, 1995). In fact, monitoring resolution for the two sets of items averaged .90 for the standard items versus .26 for the deceptive items. Overconfidence was also much greater for the deceptive items (20 percentage points) than for the standard items (3 percentage points). Because of these differences in monitoring effectiveness, the option of free report allowed participants to increase their accuracy from 28 percent to 75 percent for the standard items, whereas for the deceptive items there was only a negligible increase, from 12 percent to 21 percent. Note that for the latter set, about 80 percent of the participants' freely reported answers were wrong! Even when the overall difficulty of the standard items was matched to that of the deceptive items, the accuracy increase was still about five times greater (from 11 percent to 63 percent) than the respective increase for the deceptive items (for a similar pattern using an associate interference manipulation, see Kelly & Sahakyan, 2003; Rhodes & Kelly, 2005).

Of particular importance is the demonstration that monitoring effectiveness can affect memory performance independent of memory "retrieval." Even when retrieval, as indexed by forced-report performance, is equated, free-report accuracy is far superior when monitoring resolution is high than when it is low. Clearly, then, free-report memory performance depends on the effective operation of metacognitive processes that are simply not tapped by forced-report performance.

Control of grain size

The basic theoretical model and results dis- cussed so far have focused on how people regulate their memory performance when given the option to withhold individual items of information or entire answers about which they are unsure. Control of report option, however, is just one means by which people can regulate their memory reporting. In most real-life memory situations, people do not just have the choice of either volunteering a substantive answer or else responding "don't know." They can provide an answer but indicate that they are not entirely sure about it They can also control the "graininess" or level of precision or coarseness of the information that they provide (e.g., describing the assailant's height as "around 6 feet" or "fairly tall" rather than "5 feet 11 inches"). In attempting to explain the surprisingly superior accuracy of recall over recognition in his naturalistic study, Neisser (1988), for example, noted that the recall participants tended to choose "a level of generality at which they were not mistaken" (p. 553).

The considerations and mechanisms underlying the choice of grain size in memory reporting appear to be similar to, though somewhat more complex than, those underlying the exercise of report option. Let us return to the earlier example of a witness who wants to fulfill her vow to "tell the whole truth and nothing but the truth." How should she proceed? On the one hand, a very coarsely grained response (e.g., "between noon and midnight") will always be the wiser choice if accuracy is the sole consideration. However, such a response may not be very informative, falling short of the goal to tell the whole truth. On the other hand, whereas a very fine-grained answer (e.g., 5.23 p.m.) would be much more informative, it is also much more likely to be wrong. A similar conflict is often faced by students taking open-ended essay exams: Should one attempt to provide a very precise informative answer, but risk being wrong, or try to "hedge one's bet" by providing a coarser, less informative answer, and risk being penalized for vague- ness? In both of these examples, control over grain size can be seen to involve an accuracy-informativeness trade-off (see Yaniv & Foster, 1995, 1997) similar to the accuracy-quantity trade-off observed with regard to the control of report option.

How does one find an appropriate compromise between accuracy and informativeness in choosing a grain size for one's answers? A simple strategy is to provide the most finely grained (precise) answer that passes some preset confidence criterion. Thus, for example, a witness might try to answer the question to the nearest minute, to the near- est 5 minutes, 10 minutes, 15 minutes, and so forth, until she is, say, at least 90 percent sure that the specified answer is correct. Goldsmith, Koriat, and Weinberg-Eliezer

(2002) called this the *satisficing model* of the control of grain size: The rememberer strives to provide as precise-informative an answer as possible (without being overly precise; cf. Grice, 1975), as long as its assessed probability of being correct satisfies some reasonable minimum level. Note that this model is similar to the one presented earlier with regard to report option: The assessed probability correct of each answer that is volunteered must pass a report criterion, and the setting of the criterion level should depend on the relative incentives for accuracy and informativeness in each particular situation.

Research paradigm and methodology: Control of grain size

As in the study of report option, the challenge in the study of the control of grain size is to find a way to allow participants to control the grain size of their answers while also obtaining information about the underlying metacognitive mechanisms and performance consequences. A productive approach has been to adapt the "free-forced" paradigm used for report option.

In the basic paradigm, participants are exposed to a stimulus event and later asked to answer a set of questions that pertain to quantitative values, such as the time of an accident, the speed of a car, the height of an assailant, and so forth (for an extension to non-quantitative values, see Weber & Brewer, 2008). Alternatively, participants might be asked to answer a set of general-knowledge questions that pertain to quantitative-numeric information: date, age, height, distance, and so forth. For each question, participants are required to provide an answer at both a fine-precise grain size (e.g., to the nearest minute, mile per hour, inch, etc.) and a coarse grain size (e.g., to the nearest half hour, 20-mph interval, 10-yard, etc.). Confidence judgments are also elicited for each answer at each grain size. Finally, participants are given the option to choose which of the two alternative answers for each item (i.e., which of the two grain sizes) they prefer to actually provide under an implicit or explicit incentive for accuracy and informativeness. For example, participants might be offered five points for each correct precise answer, one point for each correct coarse answer, and penalized one point for each incorrect • answer (Goldsmith et al., 2002, Experiment 3). More naturalistically, participants might be instructed to choose the answer they would prefer to provide, assuming that "you are the only eyewitness of this crime and the police need information that is very likely to be correct'' (high accuracy - low informative- ness incentive), or that "there were several eyewitnesses to this crime and the police are in the initial stages of the investigation looking for leads" (high informativeness - low accuracy incentive; Higham, Luna, & Bloomfield, 2011).

In an adaptation of this basic paradigm, participants may be allowed to determine for themselves the grain size of the answer that they provide, rather than selecting from the grain sizes specified in advance by the experimenter (e.g., Evans. & Fisher, 2010; Goldsmith, Koriat, & Pansky, 2005, Experiment 2; Pansky, 2012; Pansky & Nemets, 2012). This additional freedom, however, requires a method for quantifying the grain size (or informativeness) of the provided answers, either as a function of the width of the provided answer interval (for quantitative information; see Goldsmith et al., 2005, Experiment 2; Yaniv & Foster, 1995), or on the basis of subjective ratings by independent judges of the precision or informativeness of the answer (e.g., Evans & Fisher, 2010; Pansky, 2012; Pansky & Nemets, 2012).

Empirical findings and conclusions: Control of grain size

We now summarize and discuss some of the main findings and conclusions that have emerged with regard to the control of grain size in memory reporting.

Rememberers are able to monitor the correctness of their best- candidate answers at different grain sizes, but differences in grain size appear to be tied to systematic differences in monitoring effectiveness Depending on the complexity of the theoretical model that is assumed (see following point), the effective control of memory grain size requires that people be able to monitor the correctness of their candidate answers at different grain sizes. In general, monitoring resolution has been found to be moderately high for both fine-grained and coarse-grained answers, with a tendency for lower resolution for the coarse-grained answers (e.g., Goldsmith et al., 2005; Luna, Higham, & Martin-Luengo, 2011). In addition, although the general finding of over- confidence applies to the monitoring of precise answers, people tend to be much less overconfident and sometimes even under- confident in the correctness of their course-grained answers (e.g., Goldsmith et al., 2002; Luna et al., 2011; Weber & Brewer, 2008). In a recent generalization of grain control to multiple-choice testing (called “plurality option”; see below), Luna et al. (2011) found that although reported coarse-grained answers were more likely to be correct than reported fine-grained answers, confidence in the former answers was lower than in the latter, yielding an apparent dissociation between confidence and accuracy across grain sizes.

One account of this pattern is that participants do not sufficiently adjust their subjective probability assessments to accommodate differences in the baseline probabilities that an answer will be correct at the different grain sizes (cf. Tversky & Kahneman, 1974). This is suggested by the typical shape of the calibration plots for fine-grained and coarse- grained answers (e.g., Figure 1 in Goldsmith et al., 2002; Figure 2 in Luna et al., 2011), in which the actual proportions correct for the coarse-grained answers are higher than the corresponding proportions correct for the fine-grained answers across the range of subjective probability categories. Thus, when participants feel that they possess precise knowledge regarding the question, they may tend to underrate the likelihood that a volunteered fine-grained answer is nevertheless wrong, and conversely, when they feel that they lack precise knowledge, they may tend to underrate the likelihood that a chosen coarse-grained answer is nevertheless correct. Such a tendency would be expected to hinder the effectiveness of the grain control process, biasing it toward the choice of fine- grained answers.

Memory monitoring guides the grain control decisions

As with the control of report option, the basic assumption of the metacognitive model of grain control is that the choice of grain size is based on subjective confidence in the correctness of one’s candidate answers. Indeed, in line with the satisficing model described earlier, high correlations have been found between confidence in one’s best-candidate fine-grained answer and the decision to provide that answer rather than a more coarse-grained answer (e.g., mean gamma correlations ranging between .74 and .85; Goldsmith et al., 2002, 2005; Weber & Brewer, 2008). Of course the grain control decision could conceivably be based not only. On confidence in one’s fine-grained candidate answer but also on confidence in alternative coarse-grained candidate answers, or perhaps on the relative gain in confidence when moving from the fine-grained to a more coarse- grained answer. Results indicate that when given the opportunity to provide either a fine-grained or a more coarse-grained answer, the grain control decision is based primarily on confidence in the fine-grained answer, in line with the simple satisficing model (Gold- smith et al., 2002, 2005; Weber & Brewer, 2008). It is possible, however, that when the choice of grain size is less constrained, the control process will be more complex.

Control over grain size enhances the accuracy of reported information, at a cost in the informativeness of the reported information (accuracy-informativeness trade-off)

Perhaps the most basic finding is that when given the option to choose the appropriate grain size for their answers, participants are not guided solely by the desire to be correct - in which case they would always choose to provide a coarse-grained answer, nor solely by the desire to be informative - in which case they would have always choose to provide a precise/fine-grained answer. Instead, participants tend to choose the coarse-grained answer when the more precise answer is unlikely to be correct. By sacrificing informativeness in this calculated manner, participants generally improve their accuracy substantially compared with what they would have achieved by providing the fine-grained answers throughout (e.g., Goldsmith et al., 2002, 2005; Luna et al., 2011; Pansky & Nemets, 2012; Weber & Brewer, 2008). The control of grain size is far from optimal, however, apparently because of imperfect monitoring: There are still many cases in which fine-grained answers are provided even though they are wrong, and coarse-grained answers are provided even though the fine-grained answer is correct.

The control of grain size in memory reporting is strategic

A key assumption of the metacognitive model of grain control is that the grain size of reported information is determined not only by the grain size of the information that is available and accessible in memory, but also on strategic control: Holding the quality of the accessible information constant, people may choose to report the information either more precisely or more coarsely, and they do so based on their subjective assessment of the likely correctness of the information and in light of implicit or explicit incentives for accuracy and informativeness. In support of this idea, studies manipulating the incentives for accuracy and informativeness have found that participants do in fact strategically adjust their grain control criterion, requiring lower levels of confidence for reporting fine- grained answers and providing more of such answers when a premium is placed on informativeness, and vice versa when the premium is placed on accuracy (e.g., Goldsmith et al., 2002, 2005; Higham et al., 2011).

The control of grain size is constrained by a minimum-informativeness criterion: When respondents are unable to provide an answer that is both sufficiently accurate and sufficiently informative, they prefer to withhold the answer entirely (if a report option is available), or violate the confidence criterion, if necessary, to provide a reasonably informative answer

According to social and pragmatic norms of communication, people are expected not only to be accurate in what they report, but also to be reasonably informative (Grice, 1975). What, then, should a rememberer do if achieving the desired level of likely correct- ness requires her to provide a ridiculously coarse answer such as “the assailant was between 5 and 7 feet tall” or “the French revolution occurred sometime between the years 1000 and 2000”? Ackerman and Gold- smith (2008) examined the control of grain size in answering either easy or very difficult general-knowledge questions, and found that when knowledge of the answer was very poor, such that a minimum-confidence criterion and a minimum-informativeness criterion could not be jointly satisfied, respondents tended to violate the minimum-confidence criterion, choosing to provide relatively precise but low-confidence answers. Relatedly, Yaniv and Foster (1995) observed that recipients of quantitative information often prefer an estimate that is precise but somewhat inaccurate to one that is completely accurate but so coarse as to be uninformative. How- ever, in the Ackerman and Goldsmith (2008, Experiment 4) study, when rememberers were allowed simultaneous control over both grain size and report option, they utilized the “don’t know” option to avoid violating either the minimum-confidence or minimum-informativeness criterion, though some precise low-confidence answers were still reported.

Ackerman and Goldsmith (2008) speculated that there may also be social-pragmatic norms that prohibit overuse of the “don’t know” option, because this too may be seen as being uninformative or uncooperative. Several other studies have also examined the joint control of grain size and report option, and the division of labor between them (e.g., Evans & Fisher, 2010; Weber & Brewer, 2008).

Applied research contexts: Focus on eyewitness memory

The strategic control of memory reporting is an important topic of research in its own right, but is also of interest because of the role that such control plays in a variety of memory research domains and topics. The application of the metacognitive framework to examine how rememberers regulate their memory reporting, as well as the performance consequences of such regulation, has yielded new insights with regard to several important memory topics and phenomena, such as (a) the effectiveness of different questioning and testing procedures in eliciting accurate memory reports (e.g., Koriat & Goldsmith, 1994; Evans & Fisher, 2010; Luna et al., 2011; Pansky & Nemets, 2012; Perfect & Weber, 2012; Weber & Perfect, 2012), (b) the credibility of children’s witness testimony (e.g., Koriat et al., 2001; Roebers & Fernandez, 2002; Roebers & Schneider, 2005; Waterman & Blades, 2011), (c) memory decline in old [a]ge (e.g., Huff, Meade, & Hutchison, 2011; Kelley & Sahakyan, 2003; Pansky et al., 2009, Rhodes & Kelley, 2005), (d) cognitive and metacognitive impairments related to schizophrenia and psychoactive medication (e.g., Danion et al., 2001; Koren et al., 2006; Mintzer et al., 2010), (e) encoding-retrieval interactions and the encoding specificity principle (e.g., Higham, 2002; Higham & Tam, 2005), and (f) psychometric and scholastic testing (e.g., Higham, 2007; Higham & Arnold, 2007; Notea-Koren, 2006).

A systematic review of the work in each of these areas is beyond the scope of this chapter (for a summary review, see Goldsmith & Koriat, 2008). Instead, we will focus here on some illustrative applications and extensions of the metacognitive report-control frame- work in the study of eyewitness memory.

Perhaps nowhere is the potential importance of metacognitive report control more clear than in the domain of eyewitness research, in which there has been enormous interest in the effects of different questioning formats and procedures on the amount and accuracy of information that can be elicited from witnesses to a crime. Thus, for example, it is established wisdom in eyewitness research that witnesses should first be allowed to tell their story in their own words (i.e., in a free-narrative format) before being subjected to more directed questioning, and that, even then, greater faith should be placed in the accuracy of the former type of testimony (e.g., Milne & Bull, 1999; Neisser, 1988). This wisdom has been incorporated, for instance, into the Cognitive Interview technique (Fisher, Schreiber Compo, Rivard, & Him, Chapter 31, this volume; Fisher & Geiselman, 1992), and into various government documents concerning the proper way to interview witnesses such as the Memorandum of Good Practice (1992), the National Institute of Child Health and Human Development (NICHD) protocol for interviewing children (Orbach, Hershkowitz, Lamb, Sternberg, Esplin, & Horowitz, 2000), and others (see Wells, Malpass, Lindsay, Fisher, Turtle, & Fulero, 2000). An important component of the Cognitive Interview and other structured interview protocols is establishing clear “communication rules”

to enable the witness to regulate his or her responses in an appropriate manner - for example, clarifying the level of detail that is forensically relevant and emphasizing to the witness that it is perfectly acceptable to respond “don’t know” or “don’t remember” when applicable (see, e.g., Powell, Fisher, & Wright, 2005, Table 2). This appears to be especially crucial with child witnesses (e.g., Roebers and Fernandez, 2002; see below).

In an early attempt to clarify the role of report option in “open-ended” reporting, Koriat and Goldsmith (1994) examined what they called the “recall-recognition paradox”: Whereas the general finding from decades of laboratory research (e.g., Brown, 1976) is that recognition testing is superior to recall testing in eliciting a greater quantity of correct information from memory, the established wisdom in eyewitness research is that recognition is inferior to recall in eliciting accurate information from rememberers (e.g., Milne & Bull, 1999). Koriat and Goldsmith (1994) showed that this seeming inconsistency stems from the common confounding in research practice between test format (recall versus recognition) and report option (free versus forced): Typically, in recognition testing, participants are forced either to choose between several alternatives or to make a yes-no decision regarding each and every item (i.e., forced report), whereas in recall testing participants have the freedom to withhold information that they are unsure about (free report). Comparing performance on a free-recognition test to a free-recall test, Koriat and Goldsmith (1994) found that recognition quantity performance was still superior to recall, but now recognition accuracy was as high as or even higher than recall accuracy. Thus, although the superior memory quantity performance of forced-recognition over free-recall testing does appear to stem from the test-format difference, the generally superior accuracy of free recall over forced recognition appears to stem entirely from report option (for a similar pattern in a developmental study, see Koriat et al., 2001).

The implication is that recognition testing and other forms of directed questioning may yield information that is highly reliable, as long as witnesses are clearly instructed regarding the legitimacy of responding “don’t know” (see also Perfect & Weber, 2012; Weber & Perfect, 2012, discussed below). Nevertheless, free-narrative reporting still appears to yield the most reliable information, as it allows simultaneous control over both grain size and report option (see Evans & Fisher, 2010, discussed below). However, more structured free-report for- mats that include both report option and grain size should be considered as well (e.g., Higham et al., 2011; Hope, Gabbert, & Brewer, 2011; Weber & Brewer, 2008). For example, Luna et al. (2011) have put forward a “plurality option” questioning format that allows rememberers to select more than one response alternatives on a multiple-choice test, thereby increasing the likelihood that the response is correct, but reducing its informativeness (see also Notea-Koren, 2006, who examined this procedure together with the simultaneous control of report option).

A second topic that is of great relevance for eyewitness memory research is the role of report control in regulating the decline of memory accuracy and informativeness over time. Although a decline in the accessibility of memories over time will almost inevitably lead to a decline in the amount of correct information that can be reported, conceivably the effective use of the option of free report and/or control over grain size could allow rememberers to avoid a decline in memory accuracy over time by withholding or coarsening information that is likely to be wrong. Ebbesen and Rienick (1998), for example, found that although the number of correct statements reported about an experienced event decreased dramatically over a four- week period, the accuracy of those statements remained stable (at about 90 percent). Essentially the same pattern was obtained by Flin, Boon, Knox, and Bull (1992) in com- paring the number and accuracy of propositions about a staged event made either one day or five months after the event. These, as well as other studies, found stable accuracy using open-ended questioning procedures that gave participants control over what information to report, and at what grain size to report it.

The idea that rememberers might use control over grain size and/or report option to maintain a stable accuracy rate over time has been examined in several studies. Goldsmith et al. (2005) had participants read a mock crime witness transcript and then asked them to answer specific questions about the described events at either a precise or coarse grain size, "to help the investigator reproduce the facts of the case." As predicted, participants provided more coarse-grained answers after 24 hours than at immediate testing, thereby maintaining a high and stable accuracy rate at the cost of reduced informativeness. After a one-week delay, even more coarse-grained answers were provided, but now there was a drop in accuracy, though much less steep a drop than would have occurred without the use of grain control. Part of the reason for the reduced accuracy was that without report option, participants could not, in some cases, avoid providing coarse-grained answers that were likely - both subjectively and objectively - to be wrong. Recently, Pansky and Nemets (2012) found that allowing participants control both over what information to report, and at what grain size to report it, enabled them to maintain stable accuracy over a retention interval of 48 hours, at the cost of a reduction in both the number and the informativeness of the answers.

The most comprehensive investigation to date was conducted by Evans and Fisher (2010), who questioned participants about details from a mock crime video using one of three questioning formats - free narrative, specific questioning (cued recall), or yes-no recognition - after either ten minutes or one week. The free-narrative and specific-questioning formats allowed participants control over both whether to report an item of information and at what level of precision, whereas the yes-no recognition format allowed control of report option only. For the specific questioning and recognition formats, forced-report answers were elicited after each initial "don't know" response, providing information about the performance that would be observed in the absence of report control. As expected, there was a significant decrease in the amount of correct information and in the precision of the information that was reported at delayed compared with immediate testing. This was also so when the forced-report responses were included, indicating a significant drop in information accessibility. At the same time, however, there was only a negligible (3 percentage points; only approaching significance) drop in free-report accuracy over this same time period. The level of accuracy attained (and maintained) was about 10 points higher for the free-narrative format (94 percent) than for the other two formats (84 percent), which did not differ from each other.

Another interesting finding in the Evans and Fisher (2010) study is that in comparing the use of the "don't know" option between the specific-questioning and yes-no recognition formats, participants tended to respond "don't know" more often to the recognition questions, possibly because the option for control of grain size was unavailable (for related results implying a "division of labor" in employing the joint control of grain size and report option, see Ackerman & Goldsmith, 2008; Weber & Brewer, 2008).

Although the studies discussed so far have focused on verbal witness reports, Weber and Perfect (Weber & Perfect, 2012; Perfect & Weber, 2012) recently demonstrated the importance of report control for eyewitness identifications as well. Weber and Perfect (2012) examined whether single-suspect ("show up") identification accuracy could be improved by providing an explicit "don't know" response option. When the "don't know" option was left implicit, it was rarely used spontaneously, with 98 percent of the participants providing a substantive yes or no response; hence performance in this condition was indistinguishable from the forced-report condition. Making the “don’t know” option explicit increased the rate of “don’t know” responses, thereby increasing the diagnosticity of the responses considerably: Compared with the forced-report condition, including an explicit “don’t know” response option reduced the proportion of false identifications by almost 50 percent, with no reduction in the proportion of correct identifications. This was so both on immediate testing and after a three-week delay.

Similarly, Perfect and Weber (2012) found that including an explicit “don’t know” option increased the overall diagnosticity of both suspect identifications and lineup rejections (suspect-absent responses) in a simultaneous lineup situation, again with no loss of diagnostic information stemming from exclusion of the “don’t know” responses. In this study, report option was manipulated within participants, with the free-report decision elicited either before or after the forced- report decision. The results indicated that a “one-step” free-report procedure yielded the most diagnostic witness decisions (best ratio of hits to false alarms), with no additional diagnostic information provided by forcing the witness to guess after an initial “don’t know” response.

These results have important implications regarding not only the diagnosticity of eyewitness identification responses, but also regarding their *output-bound* accuracy (Koriat & Goldsmith, 1994, 1996a) - the extent to which they can be relied upon to be correct. The conditional probability that an eyewitness identification response is correct, given that it was freely volunteered under conditions in which there is an explicit “don’t know” response option, is much higher than the probability correct under forced-report conditions. This higher conditional probability should be of great interest to law enforcement officials, judges, and juries (cf. Deffenbacher, Bomstein, McGorty, & Penrod, 2008; and see Koriat, Pansky, & Goldsmith, 2011, for a related analysis and discussion of output-bound accuracy).

The increased reliability of the identification responses in the two studies by Perfect and Weber, just described, with no cost in diagnostic information stemming from the exercise of report option, is important practically, but also theoretically: It indicates that, at least in these two studies, the participants made very effective use of their metacognitive monitoring and control processes, using the “don’t know” option only when their ability to identify the suspect, or correctly reject the foil suspects’ lineups, was at chance levels.

Of even greater theoretical importance is the extension of the metacognitive report control model to what is essentially a yes-no recognition memory situation. This situation involves not only the decision about whether one is confident enough that the target is *pre- sent* to volunteer an *identification* response, but also about whether one is confident enough that the target is *absent* to volunteer a suspect or lineup *rejection* response. Perfect and Weber (2012) modeled this situation by assuming a single-threshold (Type-1) signal-detection model for the forced-report (yes- no) decision, and a double-threshold (Type-2) signal-detection model for the free-report (volunteer-withhold) decision. Using this model, they were able to show, for example, that participants adopted a more conservative (Type-1) forced-report criterion, reflecting a tendency to reject the lineup rather than identify a suspect, when the forced-report decision was made first, compared with when this decision followed an initial free-report decision. A plausible explanation is that, ordinarily, a witness who is uncertain will prefer to err by wrongly rejecting a lineup rather than falsely accusing an innocent suspect. However, a witness who has just made a “don’t know” response and is then asked to make a guess may feel less wary of making a false identification, having already indicated his or her uncertainly in the preceding "don't know" response. Interestingly, such "pragmatic" social-communicative considerations (see further discussion later) appear to have influenced the placement of the Type-1, yes-no

criterion, with no effect on the placement of the double-threshold Type-2, volunteer- withhold criteria.

The final area of application to be mentioned is the regulation of responses to misleading or unanswerable questions. In a developmental study, Roebers and Fernandez (2002) (followed up by Roebers & Schneider, 2005) had children (six- to eight-year-olds) and adult participants view a short video and then answer a set of questions that included "answerable" questions in either an unbiased (e.g., "What did the girl hold in her hand?") or biased-misleading (e.g., "The girl held a bouquet of flowers in her hand, didn't she?") format, as well as questions that were "unanswerable" because the pertinent information was not contained in the video, again in either an unbiased or biased format. Report option was manipulated either with or with- out an additional accuracy incentive. With regard to the answerable questions, adults utilized the option of free report to increase report accuracy for both biased and unbiased questions, regardless of accuracy incentive. For the children, however, only the free-report option combined with incentives was effective in causing them to utilize the "don't know" option to increase report accuracy for the misleading questions (for similar results in a lineup study, see Brewer, Keast, & Sauer, 2010) . Roebers and Fernandez (2002) speculate that in the free-report condition without explicit accuracy incentives, the children may have treated the biased questioning format as similar to forced report - presenting implicit pressure to provide a substantive response (cf. Koriat et al., 2001).

With regard to the unanswerable questions, when these were asked in an unbiased manner, children and adults were both able to appropriately admit their lack of knowledge by utilizing the "don't know" option, some- what more often in the incentives condition (77 percent) than in the no-incentives condition (64 percent). When such answers were asked in a biased format, however, children in the no-incentives condition utilized the "don't know" option much less adequately (30 percent) than did adults (57 percent). Note that in the case of unanswerable questions, "don't know" is treated by researchers as a "correct" response, whereas for answer- able questions, it is treated as an "omission" that presumably reflects subjective lack of knowledge.

Pointing to this difference, Scoboria, Mazzoni, and Kirsch (2008) (see also Waterman & Blades, 2011) argue that "don't know" responses are essentially ambiguous, and that in addition to the possibility that they conceal no information at all, any given "don't know" response might in fact reflect an unstated assertion that (a) the question is unanswerable (i.e., it refers to something that did not occur or was not originally witnessed) or (b) the question is answerable (i.e., the event did occur and was witnessed) but the respondent's memory of the solicited details is insufficient to provide a confident answer. Using a post-report procedure to clarify the intended meanings of participants' "don't know" responses to a mixed set of answer- able and unanswerable questions, Scoboria et al. (2008) found that a substantial number of initial "don't know" responses could in fact be recoded as substantive assertions about "presence" or "absence," which could then be scored as either correct or wrong.

The consideration of how witnesses deal with unanswerable questions poses challenges to the metacognitive model of report control that resemble those mentioned earlier with respect to report regulation in suspect lineup identifications (e.g., Perfect & Weber, 2012). Faced with a question about a witnessed event, a witness may have to decide not only whether she is confident enough to provide her best-candidate answer, but if no such candidate arises, she may also have to decide whether she is confident enough to assert that such an event never occurred or was not witnessed. Insufficient confidence in either decision would then lead to a “don’t know” response, the intended meaning of which might later need to be clarified. Of course, one would expect that the rememberer’s monitoring processes and

control decisions with respect to potentially unanswerable questions would be strongly guided by pragmatic assumptions regarding the state of knowledge of the questioner, and the likelihood that he or she would be asking an unanswerable question in a particular research or real-world context.

CONCLUSION

Interest in “real-life” remembering over the past few decades has brought with it a myriad of challenging metatheoretical, theoretical, and methodological issues (e.g., Koriat & Goldsmith, 1996a; Koriat, Goldsmith, *&* Pansky, 2000; Cohen & Conway, 2008), including a functional approach that views memory as a multifaceted tool used in the service of achieving personal and social goals (e.g., Neisser, 1988, 1996). As Neisser has eloquently argued, remembering is like “doing” (Neisser, 1996), and hence, any complete theory of memory “retrieval” will need to deal with “the reason for retrieval ... with persons, motives, and social situations” (Neisser, 1988, p. 553).

The metacognitive framework and associated research presented in this chapter has focused on situations in which the rememberer’s goals are presumably served by providing both informative and accurate memory reports. However, as should be clear by now, depending on the effectiveness of memory monitoring, it is generally not possible to be both maximally informative and completely accurate - to tell the whole truth and nothing but the truth. Thus, rememberers are generally faced with a dilemma: Should they attempt to provide more information, taking a chance that it is wrong, or provide less information but increase the likelihood that it is correct? Much theoretical work has been directed at understanding how rememberers attempt to resolve this dilemma, and the mechanisms that they use to regulate the accuracy and quantity of the information that they report. This understanding has then been applied to examine the consequences of such regulation, as well as potential situational and group differences, in a variety of different domains, some of which were mentioned earlier (see also Goldsmith & Koriat, 2008)

Of course there is still much work that remains to be done to reach a more complete understanding of the metacognitive regulation of memory reporting and its theoretical and applied implications. Perhaps most fundamentally, it should be worthwhile to try to extend the metacognitive framework to encompass a greater range of potential goals and means of control, such as those studied by researchers taking a social-communication approach to memory (see Ackerman & Goldsmith, 2008). This approach emphasizes the cooperative pragmatic principles and assumptions involved in the explicit and implicit communication that is served by memory reporting in specific social contexts. Relevant studies have shown that people’s answers to questions are guided by pragmatic considerations and tacit assumptions relating to the background and existing knowledge of the questioner, his or her purpose in asking the question, personal goals, self-expectations, and so forth (e.g., Gibbs & Bryant, 2008; Smith & Clark, 1993). Thus, for example, people have been found to adjust the detail of the information they convey according to their perception of how much the listener needs to know (Gibbs & Bryant, 2008), to focus more on story details and narrative structure in recalling a story to an experimenter than when conveying it to a peer (Hyman, 1994), to include fewer details and verbatim quotes in recounting events when the goal was to entertain than when accuracy was emphasized (Dudukovic, Marsh, & Tversky, 2004), and to convey less detailed information to inattentive than to attentive listeners (Pasupathi, Stallworth, & Murdoch, 1998).

Another general direction for future theoretical development concerns the division of labor and potential interactions between metacognitive processes involved in report regulation and those involved in controlling

the retrieval process itself. A useful way of conceptualizing this interaction is the “manufacturing” metaphor promoted by Jacoby and colleagues (e.g., Jacoby, Shimizu, Daniels, & Rhodes, 2005): Quality control in manufacturing can be achieved either by a post-production screening process, which identifies and screens out defective products at the “back end,” or by improving the production techniques at the “front end,” so that fewer defective products are produced in the first place. Likewise, the processes used by rememberers in controlling the quality of their memory outputs presumably involve a complex interplay of cognitive and metacognitive processes that operate both to guide retrieval (front-end) and to guide reporting (back-end).

To isolate and examine both front-end and back-end components, Halamish, Goldsmith, and Jacoby (2011) developed a refined version of the QAP methodology described earlier, in which recall participants record the candidate answers that come to mind in response to a recall cue, choose from among these a best-candidate answer, rate their confidence in that answer, and, finally, decide whether to report the answer for points. Using this procedure, Halamish et al. (2011) found that rememberers use metacognitive knowledge about source encoding conditions not only to monitor the accuracy of the retrieved candidate answers, but also to control the mode of retrieval itself, mentally reinstating the relevant encoding operation to enhance the quality of produced answer-candidates (“source-constrained recall”; cf. Jacoby et al., 2005). A similar approach and methodology was used by Thomas and McDaniel (2012) to examine the effects of testing and error feedback on retrieval and report processes. A recent discussion and organizing framework for research on front-end and back-end control processes in remembering can be found in Koriat et al. (2008).

Finally, there are many further applied topics for which the metacognitive frame- work for the study of monitoring and control of performance could be useful. For exam- pie, returning to the anecdotal examples that appeared in the introduction to this chapter, do law exams, medical exams, business exams, and so forth evaluate the metacognitive as well as cognitive abilities of the examinees? Should they?

Higham and colleagues (Higham, 2007; Higham & Arnold, 2007; Higham & Gerrard, 2005) have examined the contribution of metacognitive monitoring and report control processes to performance on free-report scholastic tests (in which examinees decide which questions to answer and which to skip), pointing out that the ed, 14,

is in fact an amalgamation of cognitive and metacognitive abilities. Taking a similar approach, Notea-Koren (2006) applied the QAP procedure to separately assess the cognitive and metacognitive contributions to performance on a free-report multiple-choice aptitude test, finding that a component measure of metacognitive ability, monitoring resolution, contributed unique variance in predicting first-year university grades, beyond the predictive power of the free-report formula score (or the forced-report performance score) alone. Such results emphasize the need to consider carefully the potential contributions of metacognitive monitoring and control processes to performance not only on scholastic and psychometric tests, but also on the “real- world” criterion tasks, and attempt to devise ways to isolate and assess those contributions in an effective manner.

REFERENCES

Ackerman, R., & Goldsmith, M. (2008). Control over grain size in memory reporting - with and without satisficing knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1224-1245.

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127,* 55-68.

Brewer, N., Keast, A., & Sauer, J. D. (2010). Children's eyewitness identification performance: Effects of a Not Sure response option and accuracy motivation. *Legal and Criminological Psychology, 15,* 261-277.

Brown, J., (Ed.) (1976). *Recall and recognition.* London: Wiley.

Cohen, G., & Conway, M. A. (2008). *Memory in the real world* (3rd ed.). Hove, UK: Psychology Press.

Danion, J. M., Gokalsing, E., Robert, P., Massin-Krauss, M., & Bacon, E. (2001). Defective relationship between subjective experience and behavior in schizophrenia. *American Journal of Psychiatry, 158,* 2064-2066.

Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., & Penrod, S. D. (2008). Forgetting the once-seen face: Estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied, 14,* 139-150.

Dudukovic, N. M., Marsh, E. J., *&* Tversky, B. (2004). Telling a story or telling it straight: the effects of entertaining versus accurate retellings on memory. *Applied Cognitive Psychology, 18,* 125-143.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition.* Thousand Oaks, CA: Sage.

Ebbesen, E. B., & Rienick, C. B. (1998). Retention interval and eyewitness memory for events and personal identifying attributes. *Journal of Applied Psychology, 83,* 745-762.

Evans, J. R., & Fisher, R. P. (2010). Eyewitness memory: Balancing the accuracy, precision, and quantity of information. *Applied Cognitive Psychology, 25,* 501-508.

Fisher, R. P., & Geiselman, R. E. (1992). *Memory enhancing techniques for investigative interviewing: The Cognitive Interview.* Springfield, IL: Charles C. Thomas.

Flin, R., Boon, J., Knox, A., & Bull, R. (1992). The effect of a five-month delay on children's and adults' eyewitness memory. *British Journal of Psychology, 83,* 323-336.

Gibbs, R. W., & Bryant, G. A. (2008). Striving for optimal relevance when answering questions. *Cognition,* 106, 345-369.

Goldsmith, M. (2011). Quantity-accuracy profiles or type-2 signal detection measures? Similar methods toward a common goal. In P. A. Higham and J. P. Leboe (Eds.), *Constructions of remembering and metacognition: Essays in honor of Bruce Whittlesea* (pp. 128-136). Basingstoke: Palgrave-MacMillan.

Goldsmith, M., & Koriat, A. (2008). The strategic regulation of memory accuracy and informativeness. In A. Benjamin and B. Ross (Eds.), *Psychology of learning and motivation,* Vol. 48: Memory use as skilled cognition (1-60). San Diego, CA: Elsevier.

Goldsmith, M., Koriat, A., & Pansky, A. (2005). Strategic regulation of grain size in memory reporting over time. *Journal of Memory and Language, 52,* 505-525.

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A (2002). Strategic regulation of grain size memory reporting. *Journal of Experimental Psychology General, 131,* 73-95.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41-58). New York: Academic Press.

Halamish, V., Goldsmith, M., & Jacoby, L. L. (2011) Source-constrained recall front-end and back-end control of retrieval quality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38* 1-15.

Higham, P. A. (2002). Strong cues are not necessarily weak: Thomson and Tulving (1970) and the encoding specificity principle revisited. *Memory & Cognition, 30,* 67-80.

Higham, P. A. (2007). No special K! A signal detection framework for the strategic regulation of memory accuracy. *Journal of Experimental Psychology General, 136,* 1-22.

Higham, P.A. (2011). Accuracy discrimination and type-2 signal detection theory: Clarifications, extensions, and an analysis of bias. In P. A. Higham and J. P Leboe (Eds.), Constructions of remembering and metacognition. *Essays in honour of Bruce Whittlesea* (pp. 109-127). Basingstoke, UK: Palgrave Macmillan

Higham, P. A., & Arnold, M. M. (2007). How many questions should I answer? Using bias profiles to estimate optimal bias and maximum score on formula-scored tests. *European Journal of Cognitive Psychology, 19*(4-5), 718-742.

Higham, P. A., & Gerrard, C. (2005). Not all errors are created equal: Metacognition and changing answers on multiple-choice tests. *Canadian Journal of Experimental Psychology, 59,* 28-34.

Higham, P. A., Luna, K., & Bloomfield, J. (2011). Trace-strength and source-monitoring accounts of accuracy and metacognitive resolution in the misinformation paradigm. *Applied Cognitive Psychology, 25,* 324-335.

Higham, P. A., Perfect, T. J., & Bruno, D. (2009) Investigating strength and frequency effects in recognition memory using Type-2 signal detection theory. *Journal of Experimental Psychology Learning, Memory and Cognition, 35,* 57-80

Higham, P. A., & Tam, H. (2005). Generation failure estimating metacognition in cued recall. *Journal of Memory and Language, 52,* 595-617.

Hoffrage, U. (2004). Overconfidence. In R. F. Pohl (Ed.), *Cognitive illusions: Fallacies and biases in thinking, judgment, and memory* (pp. 235-254). Hove, UK: Psychology Press.

Hope, L., Gabbert, F., & Brewer, N. (2011). *Regulating memory: Increasing output quantity and diagnosing accuracy with a Grain Size approach.* Paper presented at the 5th International Conference on Memory, York, UK.

Huff, M. J., Meade, M. L., & Hutchison, K. A. (2011). Age-related differences in guessing on free and forced recall tests. *Memory, 19,* 317-330.

Hyman, I. E. (1994). Conversational remembering: Story recall with a peer versus for an experimenter. *Applied Cognitive Psychology, 8,* 49-66.

Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. (2005) . Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin and Review, 12,* 852-857.

Kelley, C. M., & Sahakyan, L. (2003). Memory, monitoring, and control in the attainment of memory accuracy. *Journal of Memory and Language, 48,* 704-721.

Koren, D., Seidman, L. J., Goldsmith, M.·& Harvey, P. D. (2006) . Real-world cognitive - and metacognitive - dysfunction in schizophrenia: A new approach for measuring (and remediating) more "right stuff". *Schizophrenia Bulletin, 32,* 310-326.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124,* 311-333.

Koriat, A. (2007). Remembering: Metacognitive monitoring and control processes. In H. L. Roediger, III, Y. Dudai, and S. M. Fitzpatrick (Eds.), *Science of memory: Concepts* (pp. 243-246). New York: Oxford University press.

Koriat, A. (2011). Subjective confidence in perceptual judgments: A test of the self-consistency model. *Journal of Experimental Psychology: General, 140,* 117-139.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review, 119,* 80-113.

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General, 123,* 297-316.

Koriat, A., & Goldsmith, M. (1996a). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences, 19,*167-188.

Koriat, A., & Goldsmith, M. (1996b). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review, 103,*490-517.

Koriat, A., Goldsmith, M.·& Halamish, V. (2008). Controlled processes in voluntary remembering. In J. H. Byrne, H. Eichenbaum, R. Menzel, H. Roediger and D. Sweatt (Eds.), *Learning and memory: A comprehensive reference* (pp. 307-324). Oxford: Academic Press.

Koriat, A., Goldsmith, M., & Pansky, A. (2000). Toward a psychology of memory accuracy. *Annual Review of Psychology, 51,* 481-537.

Koriat, A., Goldsmith, M., Schneider, W., & Nakash Dura, M. (2001). The credibility of children's testimony: Can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology, 79,* 405-437.

Koriat, A., Pansky, A., & Goldsmith, M. (2011). An output-bound perspective on false memories: The case of the Deese-Roediger-McDermott (DRM) paradigm. In A. Benjamin (Ed.), *Successful remembering and successful forgetting* (pp. 302-332). New York: Psychology Press.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306-334). New York: Cambridge University Press.

Luna, K., Higham, P. A., & Martin-Luengo, B. (2011). Regulation of memory accuracy with multiple answers: The plurality option. *Journal of Experimental Psychology: Applied, 17,* 148-158.

Memorandum of good practice. (1992). London: Her Majesty's Stationary Office.

Milne, R., & Bull, R. (1999). *Investigative interviewing: Psychology and practice.* Chichester, UK: Wiley.

Mintzer, M. Z., Kleykamp, B. A., & Griffiths, R. R. (2010). Dose effects of triazolam and scopolamine on metamemory. *Experimental and Clinical Psychopharmacology,* 75(1), 17-31.

Neisser, U. (1988). Time present and time past. In M. M. Gruneberg, P. Morris, and R. Sykes (Eds.), *Practical aspects of memory: Current research and issues {Vo|.* 2, pp. 545-560). Chichester, UK: Wiley.

Neisser, U. (1996). Remembering as doing. *Behavioral and Brain Sciences, 19,* 203-204.

Nelson, T. 0. (1996). Gamma is a measure of the accuracy of predicting performance on one item relative to another item, not of the absolute performance on an individual item. Comments on Schraw (1995). *Applied Cognitive Psychology, 10,* 257-260.

Notea-Koren, E. (2006). *Performance accuracy and quantity in psychometric testing: An examination and assessment of cognitive and metacognitive components.* Unpublished doctoral dissertation, University of Haifa, Israel.

Orbach, Y., Hershkowitz, I., Lamb, M. E., Sternberg, K. J., Esplin, P. W., & Horowitz, D. (2000). Assessing the value of structured protocols for forensic interviews of alleged child abuse victims. *Child Abuse & Neglect, 24,* 733-752.

Pansky, A. (2012). Inoculation against forgetting: Advantages of immediate versus delayed initial testing due to superior verbatim accessibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1792-1800.

Pansky, A., Goldsmith, M., Koriat, A., & Pearlman- Avnion,

S. (2009). Memory accuracy in old age: Cognitive, metacognitive, and neurocognitive determinants. *European Journal of Cognitive Psychology, 21,* 303-329.

Pansky, A., & Nemets, E. (2012). Enhancing the quantity and accuracy of eyewitness memory via initial memory testing. *Journal of Applied Research in Memory and Cognition, 1*, 2-11.

Pasupathi, M., Stallworth, L. M., & Murdoch, K. (1998). How what we tell becomes what we know: Listener effects on speakers' long-term memory for events. *Discourse Processes, 26,* 1-25.

Perfect, T. J., & Weber, N. (2012, May 7). How should witnesses regulate the accuracy of their identification decisions: One step forward, two steps back? *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication, doi: 10.1037/a0028461

Powell, M. B" Fisher. R. P., & Wright, R. (2005). Investigative interviewing. In N. Brewer and K. Williams (Eds.), *Psychology and law: An empirical perspective* (pp. 11-42). New York: Guilford.

Rhodes, M. G., & Kelley, C. M. (2005). Executive processes, memory accuracy, and memory monitoring: An aging and individual difference analysis. *Journal of Memory and Language, 52,* 578-594.

Robinson, M. D., Johnson, J. T., & Herndon, F. (1997). Reaction time and assessments of cognitive effort as predictors of eyewitness memory accuracy and confidence. *Journal of Applied Psychology, 82,* 416-425.

Roebers, C. M., & Fernandez, O. (2002). The effects of accuracy motivation and children's and adults' event recall, suggestibility, and their answers to unanswerable questions. *Journal of Cognition and Development, 3,* 415-443.

Roebers, C. M., Moga, N., & Schneider, W. (2001). The role of accuracy motivation on children's and adults' event recall. *Journal of Experimental Child Psychology, 78, 313-329.*

Roebers, C. M., & Schneider, W. (2005). The strategic regulation of children's memory performance and suggestibility. *Journal of Experimental Child Psychology, 91,* 24-44.

Schwartz, B. L., & Metcalfe, J. (1994). Methodological problems and pitfalls in the study of human metacognition. In J. Metcalfe and A.P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 93-113). Cambridge, MA: MIT Press.

Scoboria, A., Mazzoni, G., & Kirsch, I. (2008). "Don't know" responding to answerable and unanswerable questions during misleading and hypnotic interviews *Journal of Experimental Psychology: Applied, 14,*
' 255-265.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language, 32,*.25-38.

Thomas, R. C., & McDaniel, M. A. (2012, June 25) Testing and feedback effects on front-end control over later retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication, doi: 10.1037/a0028886.

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

Waterman, A. H., & Blades, M. (2011). Helping children correctly say "I don't know" to unanswerable questions. *Journal of Experimental Psychology. Applied, 17,* 396-405.

Weber, N., & Brewer, N. (2008). Eyewitness recall: Regulation of grain size and the role of confidence. *Journal of Experimental Psychology: Applied, 14,* 50-60.

Weber, N., & Perfect, T. J. (2012). Improving eyewitness identification accuracy by screening out those who say they don't know. *Law and Human Behavior, 36,* 28-36.

Wells, G. L., Malpass, R. S., Lindsay, R. C. L., Fisher, R. P., Turtle, J. W., & Fulero, S. M. (2000). From the lab to the police station: A successful application of eyewitness research. *American Psychologist,* 55, 581-598.

Yaniv, I., & Foster, D. P. (1995). Graininess of judgment under uncertainty: An accuracy-informativeness trade-off. *Journal of Experimental Psychology General, 124,* 424-432.

Yaniv, I., & Foster, D. P. (1997). Precision and accuracy of judgmental estimation. *Journal of Behavioral Decision Making, 10,* 21-32.

The SAGE Handbook of

Applied Memory

Edited by

Timothy J. Perfect and .
D. Stephen Lindsay

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

®SAGE

Los Angeles | London | New Delhi Singapore | Washington DC

SAGE Publications Ltd 1 Oliver’s Yard 55 City Road London EC1Y ISP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd B 1/1 1 Mohan Cooperative Industrial Area Mathura Road New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd 3 Church Street # 10-04 Samsung Hub Singapore 049483

Editor: Michael Carmichael
Assistant editor: Keri Dickens
Production manager: Cenveo Publisher Services
Marketing manager: Alison Borg
Cover design: Wendy Scott
Typeset by: Cenveo Publisher Services
Printed in Great Britain by Henry Ling Limited,
at the Dorset Press, Dorchester, DTI 1HD

Library of Congress Control Number: 2013938425

British Library Cataloguing in Publication data
A catalogue record for this book is available from the British Library

ISBN 978-1-4462-0842-7

FSC
SA-COC-1565