Published in final edited form as:

J Exp Child Psychol. 2013 November ; 116(3): 747–754. doi:10.1016/j.jecp.2013.04.004.

Influences of credibility of testimony and strength of statistical evidence on children's and adolescents' reasoning

Robert V. Kail

Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907, USA

Robert V. Kail: rkail@purdue.edu

Abstract

According to dual-process models that include analytic and heuristic modes of processing, analytic processing is often expected to become more common with development. Consistent with this view, on reasoning problems, adolescents are more likely than children to select alternatives that are backed by statistical evidence. It is shown here that this pattern depends on the quality of the statistical evidence and the quality of the testimonial that is the typical alternative to statistical evidence. In Experiment 1, 9- and 13-year-olds ($N = 64$) were presented with scenarios in which solid statistical evidence was contrasted with casual or expert testimonial evidence. When testimony was casual, children relied on it but adolescents did not; when testimony was expert, both children and adolescents relied on it. In Experiment 2, 9- and 13-year-olds ($N = 83$) were presented with scenarios in which casual testimonial evidence was contrasted with weak or strong statistical evidence. When statistical evidence was weak, children and adolescents relied on both testimonial and statistical evidence; when statistical evidence was strong, most children and adolescents relied on it. Results are discussed in terms of their implications for dual-process accounts of cognitive development.

Keywords

Cognitive development; Dual-process approaches; Heuristics; Statistical reasoning; Information processing; Decision making

Introduction

In dual-process theories of cognitive development, age differences in reasoning and decision making have been attributed to developmental changes in use of different kinds of processing. For example, several theorists (e.g., Evans, 2008; Stanovich, 2011) distinguish two forms of processing. Heuristic processing (also called System 1 or Type 1 processing) is fast and automatic, whereas analytic processing (also called System 2 or Type 2 processing) is slower and controlled. The critical development prediction has been that analytic processing should be more common in older children and adolescents (e.g., Klaczynski, 2004, 2009).

The predicted developmental outcome has been tested on several tasks. Consider, for example, reasoning problems in which participants must choose between two options, with statistical evidence favoring one option but personal experience or testimony favoring the other (Fong, Krantz, & Nisbett, 1986). In one illustrative problem (from Kokis, Macpherson, Toplak, West, & Stanovich, 2002), a child is going to a baseball game and wants to sit where she can catch a fly ball. She learns from team officials that most fly balls are caught in one section of the park but that her friend recently caught two fly balls in another section. Relying on the statistical information is thought to reflect analytic processing, whereas relying on a friend's testimony is thought to reflect heuristic processing.

Consistent with the prediction that analytic processing increases with development, several investigators have reported an age-related increase in reliance on statistical information in problems like these (Jacobs & Potenza, 1991; Klaczynski, 2001; Kokis et al., 2002). This same developmental pattern has emerged on other tasks in which participants choose between solutions thought to reflect either heuristic or analytic processing such as sunk cost decisions and gambler's fallacy problems (Klaczynski, 2001; Klaczynski & Cottrell, 2004). For example, if tossing a coin 10 times has yielded eight heads, 16-year-olds are more likely than 12-year-olds to predict that the probability of a head on the next toss is .50; children are more likely to predict that the toss will yield a tail, reflecting an averaging heuristic.

Nonetheless, it may be premature to draw strong conclusions from these findings for two reasons. The first is empirical. Not all findings are consistent with the prediction that analytic processing increases with age. For example, across four kinds of reasoning and decision-making tasks, Morsanyi and Handley (2008) found that analytic processing actually decreased between 6 and 10 years of age, whereas heuristic processing increased. Similarly, although analytic processing on inductive reasoning problems increased with age in studies by Jacobs and Potenza (1991), Klaczynski (2001) and Kokis et al. (2002), overall levels of performance varied considerably; use of statistical evidence at 11 and 12 years of age was 65%, 18%, and 49% in the three studies, respectively.

A second reason is theoretical. In fact, there are several different dual-process accounts that differ in the names of the two broad categories, the processes subsumed under each category, and specific hypotheses regarding development (Evans, 2011; Stanovich, West, & Toplak, 2011). Most accounts would hold that the different processes that comprise analytic processing typically improve with age, but predictions regarding developmental trajectories for different forms of heuristic processing are more varied. Consequently, many theorists (e.g., Stanovich et al., 2011) have argued for a more nuanced view of dual-process accounts with respect to the processes that comprise the two categories and the predictions concerning developmental change.

Against this backdrop, the research reported here was conducted to reconsider performance on the reasoning problems described earlier in which participants choose options associated with statistical or testimonial evidence. In the problems used previously (e.g., Klaczynski, 2001; Kokis et al., 2002), the statistical evidence was usually credible but the testimonial evidence was not. In the example with the baseball game, for instance, the statistical evidence comes from a credible source (the team itself) and is definitive (most balls are caught in a specific section). In contrast, the testimonial evidence comes from a friend who has no special expertise and attended only one game.

In reality, however, both kinds of evidence can range from worthless to valuable. Statistical evidence, for example, may vary in quality (e.g., based on small or unrepresentative samples) and strength (e.g., "everyone prefers Option A" vs. "many prefer Option A"). Similarly, testimonial evidence may range in value. The individual providing testimony may be someone (a) whose advice should be ignored (e.g., because his or her advice has been wrong previously), (b) who has relevant but limited experience, or (c) who has special expertise that would make his or her advice particularly valuable.

The impact of such variations in the strength of statistical and testimonial evidence was examined in two experiments. In Experiment 1, the expertise of the individual providing personal evidence was varied; in problems like those used previously the individual had limited expertise, but in other problems the individual was a bona fide expert in the problem domain. In Experiment 2, the quality of the statistical evidence was varied by replacing the typical phrase (i.e., "most") with percentages and using values that represented a slight

majority or a vast majority. Thus, the two studies were designed to determine whether the typical developmental pattern—adolescents are more likely than children to rely on statistical evidence—would hold when the strength of the evidence was varied systematically. Specifically, (a) would adolescents continue to rely on statistical evidence when testimonial evidence was compelling and (b) would children continue to rely on testimonial evidence when statistical evidence was overpowering?

Experiment 1: Casual *versus* expert testimony

In the problems used previously, the testimonial evidence was provided by individuals who were familiar to the child in the scenario (e.g., a friend or relative) but lacked special expertise or knowledge in the problem's domain. Children typically find such evidence compelling, perhaps reflecting a heuristic to follow recommendations from familiar trusted people (e.g., Koenig & Woodward, 2010). In contrast, adolescents are more likely to ignore such testimony when it competes with statistical evidence. Experiment 1 was conducted to examine performance when testimony was made more credible by describing the individual providing the testimony as having special expertise with respect to the problem (e.g., in choosing a type of soccer ball, the individual providing testimony was a soccer coach). Children were expected to continue to rely on testimony in this case. Of greater interest was adolescents' performance. Would they continue to rely on statistical evidence in the face of credible testimony, or would such testimony trump statistical evidence?

Method

Participants—The participants were 36 8- to 10-year-olds ($M = 9.52$ years, $SD = 0.83$, 19 girls and 17 boys) and 28 12- to 14-year-olds ($M = 13.33$ years, $SD = 0.85$, 12 girls and 16 boys). Most of the children were of European American descent (88%), and most were from middle-class homes (97% had mothers who had graduated from college). The participants were drawn from a longitudinal study of cognitive development and completed the tasks described here as part of a larger battery of tasks assessing processing speed, working memory, and reasoning (described in Kail, 2007).

Materials—Eight vignettes were created based on those used in prior work (e.g., Klaczynski, 2001; Kokis et al., 2002). In their original form, each described a scenario and participants chose from two alternatives. One was supported by statistical evidence from an authoritative source indicating that most people preferred this option; the alternative was supported by personal testimony from a familiar individual (e.g., relative, friend and neighbor) who had no special expertise in the scenario's domain. The following is an example:

Kyle wants a new bicycle. An article in a bicycle magazine says that most kids think the Zippo bike is best. But his uncle has a Whammo bike and he thinks that a Whammo bike is the best. Which bike should Kyle get, the Zippo or the Whammo?

An alternative form was created for each vignette by adding a phrase establishing that the person providing the testimonial had expertise related to the decision. Thus, in the example listed here, the uncle recommending one model of bike owned a bike shop. (In another vignette in which the choice involved one of two soccer balls, the friend recommending one ball was described as a soccer coach in the alternative vignette.)

Two sets of vignettes were created. Each contained eight vignettes, half in the original form that included casual testimony and half in the altered form that included expert testimony. These eight vignettes were ordered randomly. In addition, two simple base-rate problems were presented (e.g., "A class is organizing a raffle. The winning ticket will be chosen from 20 blue tickets and 10 red tickets. Which color of ticket is most likely to win?"). These were

included to disguise the contrast between the two forms of testimony included in the other vignettes. All vignettes are listed in the online Supplementary material.

Procedure—The 10 vignettes (four vignettes with casual testimony, four with expert testimony, and two base-rate problems) were presented on two pages of $8\frac{1}{2} \times 11$ -inch paper. The experimenter read each one aloud. Participants, who were tested individually, were encouraged to read along silently. At the end of each vignette, the experimenter asked the participant to select an option (e.g., “Which bike should Kyle get, the Zippo or the Whammo?”) and asked whether the person in the vignette should do this “definitely or just probably.”

Results

For each vignette, a child’s response was scored on a scale from 1 to 4, where 4 indicated the choice of the response associated with the testimonial with high confidence (i.e., definitely choose the option supported by the testimonial) and 1 indicated the choice of the response associated with statistical evidence with high confidence (i.e., definitely choose the option supported by the statistical evidence). Analyses of these data revealed main effects for age and type of testimony, $F_s(1, 60) = 17.60, p < .001$, partial $\eta^2 = .23$, that were qualified by an interaction between these variables, $F(1, 60) = 7.68, p < .01$, partial $\eta^2 = .11$. As shown in Fig. 1, on the traditional problems in which testimony was casual, adolescents relied on statistical information to a greater extent than children. Confidence intervals computed around mean ratings indicated that children’s mean ratings were significantly ($p < .05$) greater than the midpoint of 2.5, but adolescents’ mean ratings were significantly ($p < .01$) less than the midpoint. However, when the testimony was expert, age differences were eliminated; both groups relied on testimonial evidence (mean ratings were greater than the midpoint, $p_s < .01$).

This pattern emerged in a second analysis that focused on responses for individuals. Specifically, if a participant responded similarly on three of the four vignettes within a category of testimony (i.e., casual vs. expert), regardless of their confidence, they were classified as relying on that form of evidence (i.e., statistical vs. testimonial). Using this criterion, approximately 71% of participants could be classified as consistently relying on statistical or testimonial evidence (see Table 1). Analyses of these participants showed that when testimony was casual, children were more likely to rely on testimonial evidence, but adolescents were more likely to rely on statistical evidence, $\chi^2(1, N = 46) = 18.10, p < .01$. In contrast, with expert testimony, both children and adolescents relied on testimonial evidence, $\chi^2(1, N = 45) < 1$.

In summary, the typical developmental pattern emerged on conventional problems in which compelling statistical evidence was contrasted with casual testimony. Children typically chose the alternative supported by testimony, but adolescents opted for the alternative backed by statistical evidence. However, when the testimony was provided by an expert, both children and adolescents chose the alternative supported by that testimony over the alternative supported by statistical evidence.

Experiment 2: Weak versus strong evidence

The findings of Experiment 1 show that adolescents abandon statistical evidence in the face of testimony from an expert. The aim of Experiment 2 was to determine whether a complementary effect would emerge with children. Would they abandon testimony in the face of overpowering statistical evidence? Specifically, in the problems used in Experiment 1, statistical evidence was described with verbal labels (e.g., “most”) rather than exact percentages. Many children interpret *most* to include a broad range of values, extending

from 50% to 100% (Papafragou & Schwarz, 2005/2006), and this might lead them to see *most* as providing potentially equivocal evidence. To test this possibility, in Experiment 2, all problems contrasted casual testimony with either strong statistical evidence (86% preference for one option) or weak statistical evidence (51–60% preference). Of particular interest was whether children would be swayed by strong statistical evidence and ignore testimony.

Method

Participants—The participants were 48 8- to 10-year-olds ($M = 9.4$ years, $SD = 0.85$, 28 girls and 20 boys) and 35 12- to 14-year-olds ($M = 13.53$ years, $SD = 0.84$, 16 girls and 19 boys). As in Experiment 1, most children were European American (87%), and most were from middle-class homes (77% had mothers who had graduated from college). The participants were drawn from the same longitudinal study; none had participated in Experiment 1.

Materials and procedure—The eight vignettes from Experiment 1 were used. Each contrasted casual testimony with precise statistical evidence. Specifically, in each vignette, the phrase indicating that “most people” prefer one option was replaced with an exact percentage. For example, in the vignette involving choice of a bicycle, the statistical evidence was presented as “An article in a bicycle magazine says that $X\%$ of kids think the Zippo bike is best,” where X ranged from 51 to 60 in vignettes that presented weak statistical evidence and ranged from 86 to 95 in vignettes that presented strong statistical evidence. In addition, two new vignettes were created in which testimony was compared with statistical evidence that was described in the same general terms used in Experiment 1 (e.g., “most kids”). These were included to disguise the contrast between weak and strong statistical evidence.

Results

Responses were scored as before, with 4 indicating the choice of the response associated with the testimonial with high confidence (i.e., “definitely”) and 1 indicating the choice of the response associated with statistical evidence with high confidence. Analyses revealed main effects for age and strength of evidence, $F(1, 79) = 11.30, p < .01$, partial $\eta^2 = .13$, but no interaction between these variables, $F(1, 79) = 1.46, p > .10$, partial $\eta^2 = .02$. As shown in Fig. 1, adolescents relied on statistical evidence to a greater extent than children regardless of the strength of the evidence. In addition, children and adolescents relied on statistical evidence to a greater extent when that evidence was relatively stronger. When evidence was weak confidence intervals revealed that neither children’s nor adolescents’ ratings differed from the midpoint of 2.5 ($ps > .05$), but when evidence was strong both children’s and adolescents’ ratings were significantly less than the midpoint ($ps < .01$).

Using the criterion described previously, approximately 73% of participants could be classified as consistently relying on statistical or testimonial evidence (see Table 1). Analyses of their responses revealed that age was not associated with use of statistical evidence when that evidence was weak, $\chi^2(1, N = 58) = 1.73, p > .10$, or strong, $\chi^2(1, N = 63) = 2.44, p > .10$. When the statistical evidence was weak, both children and adolescents relied on testimonial and statistical evidence; when the statistical evidence was strong, most children and adolescents relied on it.

In a final analysis, participants’ ratings from the casual testimony condition of Experiment 1, in which the competing statistical evidence was conveyed verbally (i.e., with “most” in seven of eight problems), were compared separately with their ratings on the weak and strong evidence problems used in Experiment 2. For children, when compared with

statistical evidence that was described verbally, ratings of testimony did not differ when statistical evidence was weak, $t < 1$, but were less when statistical evidence was strong, $t(61) = 2.41$, $p < .02$. In contrast, for adolescents, when compared with statistical evidence that was described verbally, ratings of testimony were greater when statistical evidence was weak, $t(61) = 2.15$, $p < .05$, but less when statistical evidence was strong, $t(61) = 2.41$, $p < .02$. In other words, adolescents seem to have a more constrained interpretation of *most* in these problems (61–85%) than do children (51–85%).

General discussion

Prior studies of reasoning contrasting casual testimony with statistical evidence have found that adolescents are more likely than children to rely on statistical evidence (Jacobs & Potenza, 1991; Klaczynski, 2001; Kokis et al., 2002). That pattern was replicated in the casual testimony condition of Experiment 1, where 13-year-olds more often selected the alternative supported by statistical evidence, but 9-year-olds selected the alternative supported by testimony. However, when testimony was made stronger—by claiming that the person providing the testimony had problem-relevant expertise—both children and adolescents relied on testimony over statistical evidence. And when statistical evidence was made stronger—by replacing the nonspecific “most children” to percentages indicating that the vast majority (i.e., 86%) preferred an alternative—both children and adolescents relied on statistical evidence over testimony.

These findings underscore calls (e.g., Stanovich et al., 2011) for a more nuanced approach to dual-process accounts of cognitive development. That is, performance on the reasoning task used here has previously been interpreted as providing support for a dual-process view in which children typically rely on heuristic processing, represented by their reliance on testimony from familiar people, but adolescents favor analytic processing, represented by their reliance on statistical evidence. The current findings challenge that interpretation empirically and theoretically.

On the empirical front, it is striking that the developmental profile varied considerably with relatively modest changes in the wording of scenarios. Adolescents found testimony more compelling when a few words were added that established the credibility of the person providing that testimony. Children found statistical evidence compelling when information representing endorsement by an ill-defined majority (e.g., “most kids”) was replaced by percentages indicating overwhelming endorsement. Such sensitivity to task demands has been noted previously in the literature on heuristic and analytic processing (e.g., Chun & Kruglanski, 2006), but it has not been emphasized in developmental work. More critically, the previously reported developmental shift from relying on testimony to relying on statistical information evidently is but one of several developmental profiles for this task.

On the theoretical front, although these results are likely inconsistent with early dual-process accounts, to gain real theoretical traction research needs to provide better evidence concerning the processes that drive responses on the tasks used to investigate dual-process accounts. To illustrate, consider the finding in Experiment 1 that adolescents favor testimony over statistical evidence when the testimony comes from an expert. The traditional interpretation would be that adolescents have shifted from analytic (Type 2) processing to heuristic (Type 1) processing. An alternative account, drawing on Stanovich and colleagues’ (2011) framework, is that adolescents have shifted from one form of Type 2 processing to another. Specifically, Type 2 processing involves an algorithmic mind that taps into strategies and production systems as well as a reflective mind that taps into beliefs, goals, and general knowledge. Perhaps adolescents’ responses reflect a shift from relying on

the statistical processing associated with the algorithmic mind to relying on beliefs of the reflective mind about the strength of experts' knowledge in their domain of expertise.

For this alternative account to be compelling, we need stronger evidence that the novel response patterns observed here—adolescents' use of experts' testimony and children's use of compelling statistical information—map onto the systems described by Stanovich and colleagues (2011) or, for that matter, the modular processing, experiential learning, and automated skills that make up Evans' (2011) description of Type 1 processing.

The kind of research that is needed would use the distinctive profile of different forms of processing to reveal the underlying nature of children's and adolescents' responses. Specifically, Type 2 processing is more resource demanding than Type 1 processing. In addition, the algorithmic level of Type 2 processing is associated with measures of aptitude, but the reflective level is associated with measures of cognitive style. Consequently, the alternative account of adolescents' responding predicts that, first, adolescents' reliance on expert testimony should be predicted by measures of cognitive style (e.g., need for cognition) because it is due to reflective Type 2 processes; second, adolescents' reliance on statistical evidence should be predicted by measures of cognitive aptitude (e.g., IQ scores) because it is due to algorithmic Type 2 processes; and third, reliance on both types of evidence should be predicted by working memory measures because both reflect resource-demanding Type 2 processing. Evidence of this sort would reveal the processes that drive children's and adolescents' responses on tasks like those used here and would help to fine tune dual-process accounts of cognitive development.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The research described here was supported by a grant from the National Institute of Child Health and Human Development (NICHD, R01 046927). Thanks are due to Sherry Fetterman, Megan Rockwood, Ariel Hughes, and Will Bankston for their help in collecting the data.

References

- Chun WY, Kruglanski AW. The role of task demands and processing resources in the use of base-rate and individuating information. *Journal of Personality and Social Psychology*. 2006; 91:205–217. [PubMed: 16881759]
- St Evans JBT. Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*. 2008; 59:255–278.
- St Evans JBT. Dual-process theories of reasoning: Contemporary issues and developmental applications. *Developmental Review*. 2011; 31:86–102.
- Fong GF, Krantz DH, Nisbett RE. The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*. 1986; 18:253–292.
- Jacobs JE, Potenza M. The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development*. 1991; 62:166–178.
- Kail RV. Longitudinal evidence that increases in processing speed and working memory enhance children's reasoning. *Psychological Science*. 2007; 18:312–313. [PubMed: 17470254]
- Klaczynski PA. Analytic and heuristic processing influences on adolescent reasoning and decision-making. *Child Development*. 2001; 72:844–861. [PubMed: 11405586]
- Klaczynski, PA. A dual-process model of adolescent development: Implications for decision-making, reasoning, and identity. In: Kail, RV., editor. *Advances in child development and behavior*. Vol. 32. Amsterdam: Elsevier; 2004. p. 73-123.

- Klaczynski, PA. Cognitive and social cognitive development: Dual-process research and theory. In: St vns, JBTE.; Frankish, K., editors. *In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press; 2009. p. 265-292.
- Klaczynski PA, Cottrell JM. A dual-process approach to cognitive development: The case of children's understanding of sunk cost decisions. *Thinking & Reasoning*. 2004; 10:147-174.
- Koenig MA, Woodward AL. Twenty-four-month-olds' sensitivity to the prior inaccuracy of the source. *Developmental Psychology*. 2010; 46:815-882. [PubMed: 20604604]
- Kokis JV, Macpherson R, Toplak ME, West RF, Stanovich KE. Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*. 2002; 83:26-52. [PubMed: 12379417]
- Morsanyi K, Handley SJ. How smart do you need to be to get it wrong? The role of cognitive capacity in the development of heuristic-based judgment. *Journal of Experimental Child Psychology*. 2008; 99:18-36. [PubMed: 17920617]
- Papafragou A, Schwarz N. Most wanted. *Language Acquisition*. 2005/2006; 13:207-215.
- Stanovich, KE. *Rationality and the reflective mind*. New York: Oxford University Press; 2011.
- Stanovich KE, West RF, Toplak ME. The complexity of developmental predictions from dual process models. *Developmental Review*. 2011; 31:103-118.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jecp.2013.04.004>.

Fig. 1.

Mean ratings in Experiments 1 and 2 separately for 9- and 13-year-olds. Endorsement of testimony was measured on a 4-point scale in which 4 corresponded to *definitely choose the alternative endorsed by testimony* and 1 corresponded to *definitely choose the alternative supported by statistical evidence*. Error bars depict standard errors. The dashed line corresponds to the scale midpoint. Means above the midpoint depict endorsement of testimony, and those below the midpoint depict endorsement of statistical evidence.

Table 1

Number of individuals at each age who relied on different types of evidence.

<i>Experiment 1: Casual versus expert testimony</i>					
Age (years)	Casual testimony		Expert testimony		
	Testimonial evidence	Statistical evidence	Inconsistent	Testimonial evidence	Statistical evidence
8–10	18	7	11	27	1
12–14	2	19	7	16	1
<i>Experiment 2: Weak versus strong statistical evidence</i>					
	Weak statistical evidence		Strong statistical evidence		
	Testimonial evidence	Statistical evidence	Inconsistent	Testimonial evidence	Statistical evidence
8–10	16	13	19	6	25
12–14	11	18	6	2	30