Hays, W. (1973). *Statistics for the social sciences* (2nd ed.). New York: Holt.

Jones, L.V. (1955). Statistics and research design. *Annual Review of Psychology, 6,* 405–430.

Loftus, G.R. (1985). Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11,* 396–405.

Loftus, G.R. (1991). On the tyranny of hypothesis testing in the social sciences. *Contemporary Psychology, 36,* 102–105.

Loftus, G.R. (1993a). Editorial comment. *Memory & Cognition, 21,* 1–3.

Loftus, G.R. (1993b). Visual data representation and hypothesis testing in the microcomputer age. *Behavior Research Methods, Instruments, & Computers, 25,* 250–256.

Loftus, G.R. (1995). Data analysis as insight. *Behavior Research Methods, Instruments, & Computers, 27,* 57–59.

Loftus, G.R., & Bamber, D. (1990). Weak models, strong models, unidimensional models, and psychological time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 916–926.

Loftus, G.R., & Masson, M.E.J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review, 1,* 476–490.

Lykken, D.T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70,* 151–159.

Maltz, M.D. (1994). Deviating from the mean: The declining significance of significance. *Journal of Research in Crime and Delinquency, 31,* 434–463.

Meehl, P.E. (1967). Theory testing in psychology and physics: A methodological paradox. *Philosophy of Science, 34,* 103–115.

Meehl, P.E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald and the slow process of soft psychology. *Journal of Consulting and Clinical Psychology, 46,* 806–834.

Meehl, P.E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports, 66*(Suppl. 1), 195–244.

Nunnally, J. (1960). The place of statistics in psychology. *Educational and Psychological Measurement, 20,* 641–650.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183–192.

Rosnow, R.L., & Rosenthal, R. (1989). Statistical procedures and the justification of knowledge in psychological science. *American Psychologist, 44,* 1276–1284.

Rozeboom, W.W. (1960). The fallacy of the null-hypothesis significance test. *Psychological Bulletin, 57,* 416–428.

Schmidt, F. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods, 1,* 115–129.

Slamecka, N.J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 384–397.

Tyler, R.W. (1935). What is statistical significance? *Educational Research Bulletin, 10,* 115–118, 142.

The Water-Level Task: An Intriguing Puzzle

Ross Vasta and Lynn S. Liben[1]

Department of Psychology, SUNY Brockport, Brockport, New York (R.V.), and Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania (L.S.L.)

Take a moment to look at Figure 1. It presents one of several variations of an intriguing problem known as the water-level task (WLT). The correct response to the problem is to draw a horizontal line across the bottle, reflecting the general principle that the surface of a liquid is invariantly horizontal regardless of the orientation of its container. Variations of the task have included presenting the tilted bottle alone, using real containers rather than drawings, and asking subjects whether a waterline in a tilted container looks "correct" (rather than having them draw a line).[2]

The WLT might appear to be a simple problem. In reality, researchers have found that a surprisingly large proportion of adolescents and adults draw slanting lines in the tilted bottles (often with considerable confidence!), and are unable to articulate or identify the physical principle underlying the task. Determining which subjects are most likely to make errors, and why they do so, has been a 30-year scientific puzzle that continues to challenge investigators.

In this article, we begin by tracing the WLT to its source and original purpose—Piaget's work on children's spatial development. We then examine how the task provided an inadvertent battleground for the theoretical debate surrounding gender differences that emerged during the 1970s. Finally, we consider current attempts to explain the fascinating data that continue to be generated by the WLT, and we suggest some directions future research might take.

ORIGINS OF THE PROBLEM

The WLT was developed by Piaget and Inhelder (1948/1956) as part of their investigation of children's emerging spatial concepts. Piaget and Inhelder proposed that

Recommended Reading

Kalichman, S.C. (1988). Individual differences in water-level performance: A component skills analysis. *Developmental Review, 8,* 273–295.

Liben, L.S. (1991). The Piagetian water-level task: Looking beneath the surface. In R. Vasta (Ed.), *Annals of child development: Vol. 8* (pp. 81–144). London: Kingsley.

Pascual-Leone, J., & Morra, S. (1991). Horizontality of water level: A neo-Piagetian developmental review. In H.W. Reese (Ed.), *Advances in child development and behavior: Vol. 23* (pp. 231–276). San Diego: Academic Press.

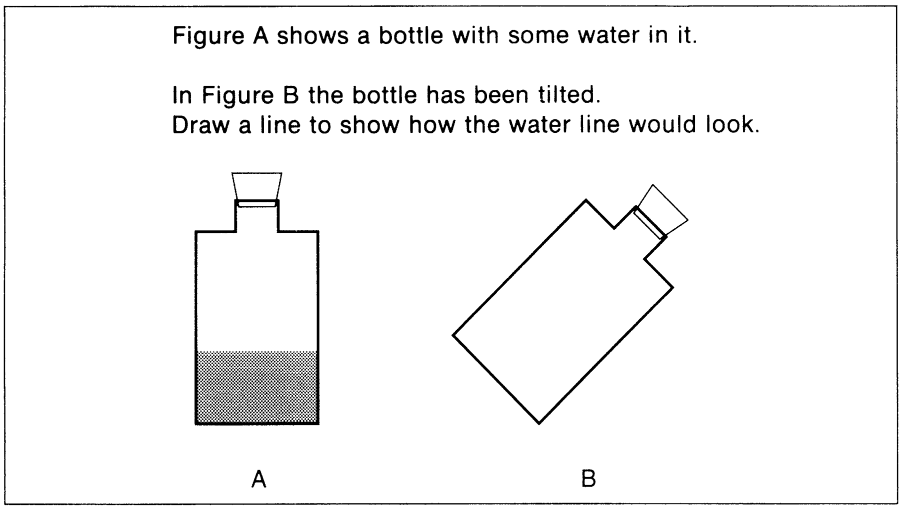

Fig. 1. The water-level task.

children gradually come to construct a euclidean (three-dimensional) conceptual system of horizontal and vertical axes with which to represent space. This reference system functions as an abstract geometric container and is independent of the objects found within it. However, because the physical environment itself contains horizontals (e.g., the horizon, tabletops) and verticals (e.g., flagpoles, intersections of walls), a true test of "whether the child has any real understanding of these notions . . . [requires determining how the child] discovers real physical laws . . . such as the constancy of the surface of a liquid whatever the angle of the container" (Piaget & Inhelder, 1948/1956, p. 380). Piaget and Inhelder thus viewed the ability to perform accurately on the WLT as a key indicator that the child has developed a mature euclidean reference system.

Based on their study of children at various ages, Piaget and Inhelder charted the developmental sequence illustrated in Figure 2. They reported that the drawings of very young children (under age 4) do not even portray the planar surface of the water (Stage IA). Slightly older children typically represent the waterline as fixed relative to the sides of the container by drawing the waterline parallel to the container's base (Stage IIA). Children next begin to indicate that the water's position changes relative to the sides of the container, first mistakenly showing the water as tilted in all but upright containers (Stage IIB), and later erring only when one axis of the container is not horizontal (Stage IIIA). Finally, by about age 9, children consistently produce horizontal lines (Stage IIIB). From these original reports, it seemed a given that adolescents and adults would have little trouble with the WLT.

In 1964, however, Rebelsky reported that some of her graduate and undergraduate students at Boston University had considerable difficulty with the task. And since that time, many other investigators have confirmed that many college-educated adults do not respond correctly on the task. In addition, Rebelsky reported that females were less accurate than males, a finding that has been replicated by virtually all subsequent researchers.

MODERN STUDY OF THE TASK

Since Piaget's early developmental work on the WLT, most studies have involved older subjects, whose failure to perform accurately poses the major puzzle. The research also is no longer exclusively Piagetian, but now invokes biological, information processing, and other theoretical models and mechanisms to explain the findings.

Gender Differences

The 1970s witnessed an explosion of psychological research concerned with gender differences and their origins. Correspondingly, in the era immediately following Rebelsky's (1964) empirical revelations, much of the work on the WLT focused on the observed gender differences. Meta-analyses of spatial abilities research have found these differences to be significant and meaningful, as indicated by statistical effect sizes[3] on the WLT and related tasks that range from .44 for all subjects to around .60 for adults (Linn & Petersen, 1985; Voyer, Voyer, & Bryden, 1995). It is difficult to pinpoint the proportions of males and females who have difficulty with the task, as these vary with methodological factors such as the version of the task that is used. Nevertheless, a recent report of three WLT experiments with undergraduate subjects is probably representative: Results for males showed that about 50% performed very well and about 20% performed poorly, whereas results for females showed that about 25% performed very well and about 35% performed poorly (Sholl & Liben, 1995).

Note, then, that although the differences are robust, not all males perform well and not all females perform poorly. Thus, any explanation that ties WLT performance exclusively to the subject's biological sex is untenable. In fact, Thomas and Lohaus (1993) have

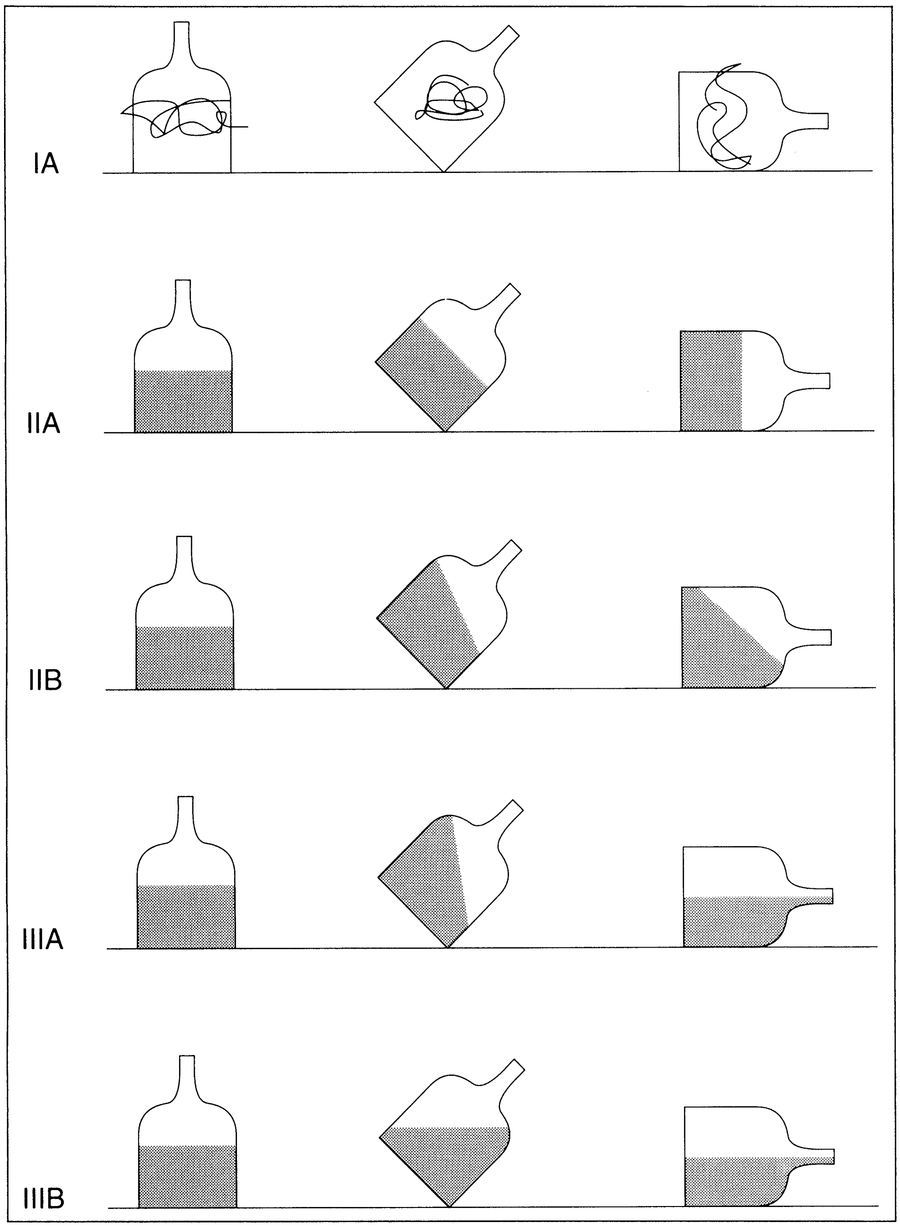

Fig. 2. Stages of water-level task development proposed by Piaget and Inhelder (1948/1956). From Downs and Liben (1991).

gone so far as to assert that there are no differences between the males and females who do well on the task or between the males and females who do poorly; the only gender difference is in the proportions of people who perform well and who perform poorly. This conclusion grows out of the researchers' application of a mathematical model to the performance of subjects ranging from young children to adults. Although provocative, the conclusion is preliminary and needs additional support because, in at least some studies, performance differences between males and females have been detected even within low- and high-scoring groups (Sholl & Liben, 1995) and within groups defined by whether or not they know the principle (Vasta, Lightfoot, & Cox, 1993).

Some explanations of the gender differences have involved biological mechanisms. For example, researchers have proposed that (a) a recessive gene on the X chromosome both facilitates acquisition of the horizontality principle and is more frequently expressed in men than in women (Thomas & Jamison, 1981), (b) different levels of exposure to sex-related hormones, such as androgen and estrogen, during the prenatal period cause the brains of males and females, including those areas that involve spatial ability, to develop differently (Collaer & Hines, 1995), and (c) gender differences in the effectiveness of the vestibular system of the brain may influence the perception of gravitational upright, and thus the perpendicular horizontal (Liben & Stern, 1996; Sholl, 1989). Although each of these explanations has supporting evidence, none can account for all of the findings.

Socialization models, too, have been proposed to explain the gender differences on the task. Most such models are based on the assumption that in our culture, boys are more strongly encouraged than girls to participate in activities that promote the development of spatial skills (e.g., block play, carpentry, math and science courses). Some correlational evidence does relate good WLT performance to self-reports of early participation in spatial activities, but methodological and cause-effect issues cloud firm conclusions (Baenninger & Newcombe, 1995). Similarly, there is evidence that performance on the WLT is related to subjects' vocational choices, although gender differences persist within occupations and the cause-effect relations therefore remain unclear (Robert & Harel, 1996; Vasta, Rosenberg, Knott, & Gaze, in press).

Another method of investigating possible socialization influences is through the use of a training paradigm. To the extent that gender differences result from differential experience with certain spatial activities, a training regimen designed to provide such experiences could improve the performance of females on the task and thereby reduce or even eliminate the gender differences. Of course, any salutary effects of a training program could result from females' learning to use a strategy that, although effective, is different from the one used by males or from the "natural" path of learning. Even in this case, however, the results would indicate that, at the very least, the gender differences in task performance are not inevitable, as might be predicted by a strong biological model. Some training studies, in fact, have reported success at improving the WLT performance of women, even to the point of bringing it to the level of males' performance (Baenninger & Newcombe, 1995; Liben & Golbeck, 1984; Vasta, Knott, & Gaze, 1996).

Interactional models probably hold the greatest likelihood of explaining the gender differences. Such models include variations on the "bent-twig" idea that males begin with stronger biologically based spatial interests and abilities. These characteristics presumably lead them to seek out more spatial activities and experiences, thereby promoting their spatial skills even further (Sherman, 1978). Recent data reported by Casey (1996), for example, extend this model to differences in spatial abilities among women. Casey's research was based on a theory of brain lateralization which posits that right-handed females with left-handed or ambidextrous relatives are more likely to be genetically endowed with strong spatial abilities than are women who display other handedness patterns (Annett, 1994). Casey hypothesized that only women who possessed this particular biological spatial advantage would benefit from experiences that have the potential to promote spatial skills. And indeed, her research demonstrated that among women who had taken many math and science courses, those with the stipulated pattern of familial handedness performed best on a spatial task requiring them to rotate a three-dimensional figure in their minds. A specific combination of biological potential and prior experiences thus appears necessary to produce in women a high level of the spatial ability assessed by this task. It is conceivable, then, that a similar model accounts for the finding that many women perform poorly on the WLT, whereas a minority perform very well.

Individual Differences: Perceptual Factors

Much of the recent study of the WLT has subsumed the gender differences under the more general search for variables that separate good from poor performers. One such set of variables operates at the sensory-perceptual level.

The WLT is an example of a spatial task involving spatial perception ability. Such tasks require subjects to locate the horizontal or vertical axis in the face of competing perceptual cues (Linn & Petersen, 1985; Voyer et al., 1995). Spatial perception may be related to the concept of field dependence/independence (FD/I), developed by Witkin (Witkin & Goodenough, 1981). In his analysis of cognitive style, Witkin contended that individuals who are field independent are able to focus on a task or problem independent of its context, whereas those who are field dependent have more difficulty doing so. If spatial perception is indeed linked to FD/I, individuals who are field dependent would be expected to err in determining the horizontal axis (i.e., waterline) when it is located within a tilted frame. A number of studies, in fact, have reported significant positive correlations between subjects' scores on the WLT and on Witkin's principal task for assessing FD/I, the rod-and-frame test, in which subjects attempt to place a movable rod at vertical within a tilted frame (Liben, 1978; Sholl, 1989).

Spatial perception ability plays a role in WLT performance because the oblique sides of a tilted container provide cues that compete with subjects' ability to locate the horizontal axis. Indeed, the WLT may be thought of as similar to a visual illusion in which the tilted frame induces subjects to perceive internal horizontal or vertical lines as displaced in the direction opposite the frame's tilt (Fig. 3; Coren & Hoy, 1986).

Support for the role of this perceptual process has come from studies showing that accuracy on the WLT improves when the container has rounded sides that interfere less with drawing a waterline that is horizontal (Vasta et al., 1993), and from studies showing that subjects have difficulty even when water is removed from the problem entirely—for example, when subjects are asked simply to draw horizontal lines within tilted rectangles (Liben & Golbeck, 1986; Vasta et al., 1993), or when they are asked to draw a horizontal bar pivoted on a tilted rod surrounded by a tilted frame (Liben, 1991).

But perceptual processes alone cannot explain the differences between successful and unsuccessful performers on the WLT. Even among subjects who understand the behavior of water and who perform well on the task, illusory tilt promotes the perception of a horizontal line appearing nonhorizontal. Thus, some other variable or variables must separate these groups (Sholl & Liben, 1995).

Published by Cambridge University Press

Individual Differences: Cognitive Factors

It may be clear, by this point, why most theoretical models of the WLT have viewed it as a task involving a number of different skills and competencies. These competencies include the categories of motor skills, sensory-perception abilities, graphic skills, and cognitive processes. In recent years, some especially exciting research on the WLT has focused on the last of these areas, as researchers have begun to identify the cognitive mechanisms that are active during the task and that are instrumental in determining a subject's success or failure.

The variable that most clearly predicts performance on the WLT is subjects' knowledge of the physical principle that the surface of a liquid remains invariantly horizontal. As would be expected, those who can articulate or identify the principle are considerably more accurate on the task than those who cannot (Liben & Golbeck, 1984; Vasta et al., 1993). Nevertheless, this measure does not predict performance perfectly; some subjects who know the principle perform poorly, and others who do not know it perform well.

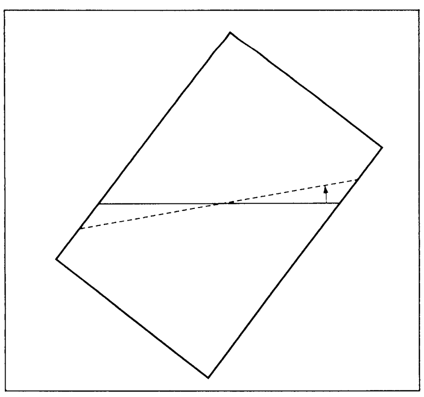

Fig. 3. The visual frame illusion. When viewing a tilted frame, most subjects perceive the horizontal (solid line) as displaced in the direction opposite the frame tilt (dashed line). From Sholl and Liben (1995).

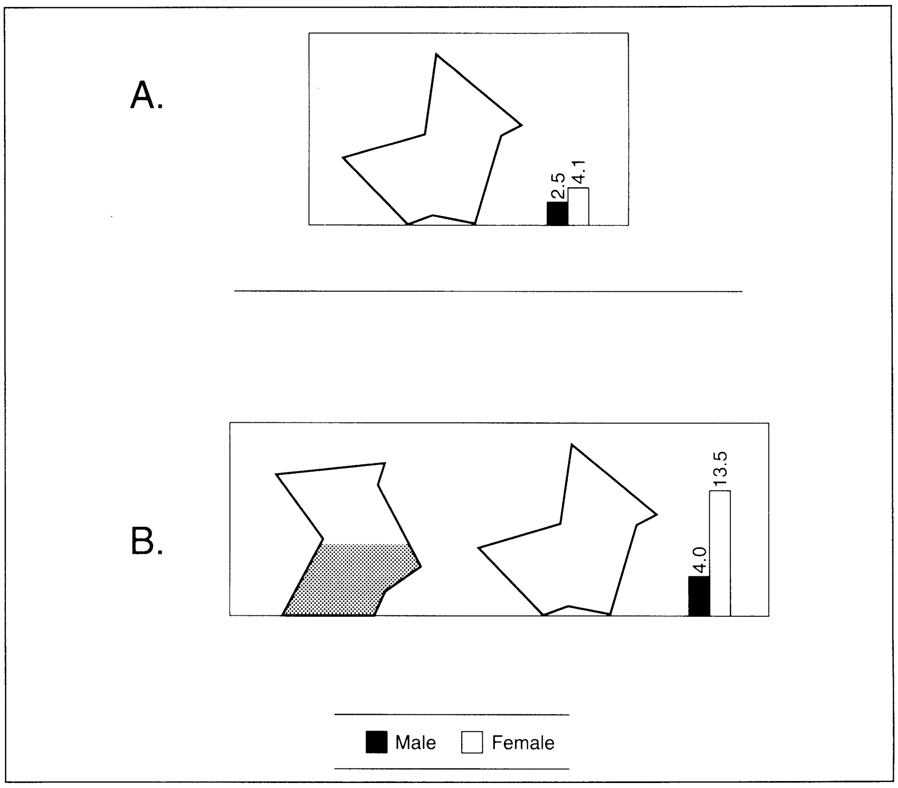

Fig. 4. Tasks and results from a study of water-level task performance using containers of various shapes, and presented in a one-bottle (a) or two-bottle (b) format. The bars represent the average degrees of deviation from horizontal for the waterlines drawn by male and female subjects.

Some subjects, after drawing their waterlines in the tilted containers, report they imagined the water was in motion, rather than at rest. These subjects tend to produce larger errors on the task. And, perhaps not surprisingly, more females than males have been found to fall into this category (Robert & Morin, 1993).

Research that currently holds great promise for understanding the WLT involves the cognitive strategies subjects bring to bear on the task. Consider, for example, Figure 4a. When subjects are told the object is a glass container and that they should draw a line to show how the container would look if half full of water, virtually every subject draws a near-horizontal line. Yet, when instead shown Figure 4b and asked to draw the waterline in precisely the same empty container, many subjects draw markedly nonhorizontal lines (Vasta, 1994). Because the containers are identical in the two tasks, the errors by the latter subjects cannot result from, for example, perceptual characteristics (oblique lines) inherent in the shape of the container. Rather, before drawing their waterlines in the empty container, these subjects must perform some sort of cognitive operation that apparently is induced by the presence of the partially filled upright container.

Research has shown further that the particular relation between the two containers also affects subjects' drawings. For example, one study compared the lines subjects drew in the empty containers in Figures 5a and 5b, which differ in the degree to which the empty containers appear to have been rotated from the partially filled containers. Even though the empty containers are identical, subjects' lines deviated more from horizontal in Figure 5a, in which the empty container appears to

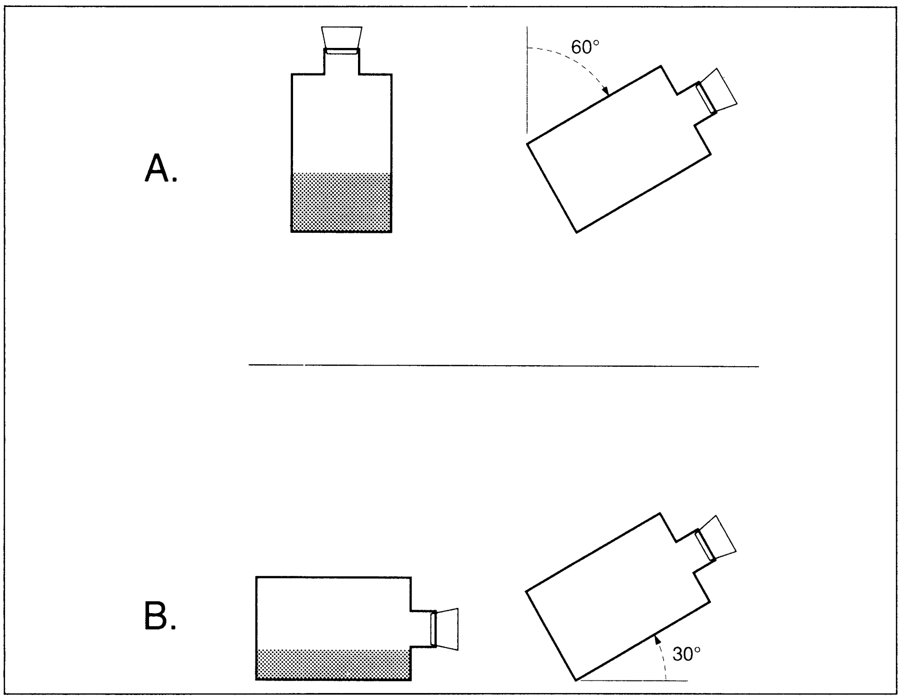

Fig. 5. Two tasks used in a study investigating the potential role of mental rotation in the water-level task. The reference lines have been added for illustration and were not present in the original task. From Vasta, Belongia, and Ribble (1994).

have been rotated 60°, than in Figure 5b, in which the empty container appears to have been rotated 30°.

One cognitive mechanism proposed to account for these findings is mental rotation. According to this hypothesis, subjects who encounter a tilted container mentally rotate it from its original position before drawing their waterlines, and the amount of required mental rotation is positively related to the size of subjects' errors (Vasta, Belongia, & Ribble, 1994).

The potential role of cognitive strategies was also demonstrated in research comparing subjects who scored poorly on the WLT with subjects who scored well. The task involved determining whether lines drawn in tilted rectangles were horizontal or slightly tilted. In one condition, subjects simply made their responses, whereas in a second condition, they were instructed to first superimpose an imaginary grid on the page to assist in locating the horizontal and vertical axes. The results showed that subjects who were low scorers performed worse in the first condition than when using the grid strategy. Subjects who were high scorers performed the same in both conditions (Sholl & Liben, 1995). These data suggest that adults who perform well on the WLT spontaneously activate cognitive strategies to assist them on the task, whereas those who perform poorly do not—even though they apparently could if so instructed.

FUTURE DIRECTIONS

Evidence from a number of different areas is now beginning to reveal why the WLT proves difficult even for adolescents and adults, and why some individuals are especially prone to having difficulty. But the puzzle is far from solved, and a number of important questions and issues remain. Fortunately, several directions of inquiry appear to offer considerable promise for addressing these questions.

One fundamental issue involves the developmental origins of the skills needed for success on the task. One approach to determining which biological factors and early experiences contribute to success on the WLT later in life would be to identify the factors that appear to affect performance in adults, and then study these factors prospectively in children using longitudinal research designs. Thus, data from the same individuals would be available from childhood, adolescence, and adulthood, and the various pathways to success and failure could be examined more directly. This approach not only would shed light on why many adults do poorly on the task, but also might go a long way toward explaining the observed gender differences. Although a longitudinal approach requires considerable patience, it promises to provide important information about the developmental course that separates successful from unsuccessful subjects.

Another potential focus of future work is studying the WLT in its larger scientific context. The growing literature on adults' misunderstanding of everyday physical principles suggests that ignorance of the physical behavior of water is hardly an isolated problem. Adults have been shown to hold naive beliefs regarding many aspects of physics and often err, for example, when asked to draw the correct trajectory of a car driving off a cliff or a ball emerging from a spiral tube (McCloskey, 1983). From this perspective, it seems important to continue to explore the correlates of good and poor WLT performance, correlates that might transcend the laboratory to include learning tasks posed by educators in the classroom and supervisors in the workplace.

The finding that so many adults perform poorly on a task designed originally for children represents a particular challenge to developmentalists, especially those who endorse a universalist perspective regarding the processes and end states of cognitive development. It is important to recognize, however, that adults and children may have difficulty with the task for very different reasons. For example, children's errors may indeed reflect the absence of a euclidean spatial system, as argued by Piaget and Inhelder half a century ago, but it is doubtful that adults' errors signify the same deficit. Rather, it is more likely that most normal adults have developed a mature spatial system (probably, as predicted, by about age 9), but that a considerable proportion fail to apply it when necessary. However, to whatever extent some adults have truly failed to establish a euclidean reference system, it will be necessary to identify the alternative spatial concepts they apply in the WLT and why these, in particular, have developed.

The WLT remains an enigma, despite being one of the most widely investigated spatial problems. But as psychologists continue to identify mechanisms that contribute to task mastery, they will move ever closer to solving this intriguing puzzle.

Acknowledgments—We thank Norm Frisch for his assistance with the artwork.

Notes

1. Address correspondence to Ross Vasta, Department of Psychology, SUNY Brockport, Brockport, NY 14420; e-mail: rvasta@acsprl.acs.brockport.edu.

2. For more variations on the task, see the Recommended Reading.

3. Effect size is a statistic used to calculate the influence of the independent variable on the dependent variable. When we report that scores from two levels of an independent variable are significantly different, we mean that the difference between their mean scores is not likely to result from chance. But in order to say something about the size of this effect, we must consider more than just the difference between the mean scores; we must also consider the variability within each group. For example, one way to calculate effect size is to divide the difference between the two means by the standard deviation of the groups being compared (Cohen, 1977). Effect size thus reveals more than whether the group differences are statistically significant; it can tell us whether the effect is large or small.

References

Annett, M. (1994). Handedness as a continuous variable with dextral shift: Sex, generation and family handedness in subgroups of left and right handers. *Behavioral Genetics, 24,* 51–63.

Baenninger, M., & Newcombe, N. (1995). Environmental input to the development of sex-related differences in spatial and mathematical ability. *Learning and Individual Differences, 7,* 363–379.

Casey, M.B. (1996). Understanding individual differences in spatial ability within females: A nature/nurture interactionist framework. *Developmental Review, 16,* 240–261.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Collaer, M.L., & Hines, M. (1995). Human behavioral sex differences: A role for gonadal hormones during early development? *Psychological Bulletin, 118,* 55–107.

Coren, S., & Hoy, V.S. (1986). An orientation illusion analog to the rod and frame: Relational effects in the magnitude of distortion. *Perception & Psychophysics, 39,* 159–163.

Downs, R.M., & Liben, L.S. (1991). The development of expertise in geography: A cognitive-developmental approach to geographic education. *Annals of the Association of American Geographers, 81,* 304–327.

Liben, L.S. (1978). Performance on Piagetian spatial tasks as a function of sex, field dependence, and training. *Merrill-Palmer Quarterly, 24,* 97–110.

Liben, L.S. (1991). Adults' performance on horizontality tasks: Conflicting frames of reference. *Developmental Psychology, 27,* 285–294.

Liben, L.S., & Golbeck, S.L. (1984). Performance on Piagetian horizontality and verticality tasks: Sex-related differences in knowledge of relevant physical phenomena. *Developmental Psychology, 20,* 595–606.

Liben, L.S., & Golbeck, S.L. (1986). Adults' demonstration of underlying Euclidean concepts in relation to task context. *Developmental Psychology, 22,* 487–490.

Liben, L.S., & Stern, R.M. (1996, November). *Water-level task performance and susceptibility to visually-induced motion sickness.* Paper presented at the meeting of the Psychonomic Society, Chicago.

Linn, M.C., & Petersen, A.C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56,* 1479–1498.

McCloskey, M. (1983). Intuitive physics. *Scientific American, 248,* 122–130.

Piaget, J., & Inhelder, B. (1956). *The child's conception of space* (F.J. Langdon & J.L. Lunzer, Trans.). London: Routledge & Kegan Paul. (Original work published 1948)

Rebelsky, F. (1964). Adult perception of the horizontal. *Perceptual and Motor Skills, 19,* 371–374.

Robert, M., & Harel, F. (1996). The gender difference in orienting liquid surfaces and plumb lines: Its robustness, its correlates, and the associated knowledge of simple physics. *Canadian Journal of Experimental Psychology, 50,* 280–314.

Robert, M., & Morin, P. (1993). Gender differences in horizontality and verticality representation in relation to initial position of the stimuli. *Canadian Journal of Experimental Psychology, 47,* 507–522.

Sherman, J. (1978). *Sex-related cognitive differences.* Springfield, IL: Charles C. Thomas.

Sholl, M.J. (1989). The relation between horizontality and rod-and-frame and vestibular navigational performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 110–125.

Sholl, M.J., & Liben, L.S. (1995). Illusory tilt and Euclidean schemes as factors in performance on the water-level task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1624–1638.

Thomas, H., & Jamison, W. (1981). A test of the X-linked genetic hypothesis for sex differences on Piaget's water-level task. *Developmental Review, 1,* 274–283.

Thomas, H., & Lohaus, A. (1993). Modeling growth and individual differences in spatial tasks. *Monographs of the Society for Research in Child Development, 58*(9, Serial No. 237).

Vasta, R. (1994, June). *Do adults perceive tilted bottles on the water-level task as rotated?* Paper presented at the annual meeting of the American Psychological Society, Washington, DC.

Vasta, R., Belongia, C., & Ribble, C. (1994). Investigating the orientation effect on the water-level task: Who? When? and Why? *Developmental Psychology, 30,* 893–904.

Vasta, R., Knott, J.A., & Gaze, C.E. (1996). Can spatial training erase the gender differences on the water-level task? *Psychology of Women Quarterly, 20,* 549–567.

Vasta, R., Lightfoot, C., & Cox, B.D. (1993). Understanding gender differences on the water-level problem: The role of spatial perception. *Merrill-Palmer Quarterly, 39,* 391–414.

Vasta, R., Rosenberg, D., Knott, J.A., & Gaze, C.E. (in press). Experience and the water-level task revisited: Does expertise exact a price? *Psychological Science.*

Voyer, D., Voyer, S., & Bryden, M.P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117,* 250–270.

Witkin, H.A., & Goodenough, D.R. (1981). *Cognitive styles: Essence and origins.* New York: International Universities Press.