



Kathy McNamara, Ph.D., Cleveland State University

# The Ethics of Scientific Thinking: Assessment, Intervention, and Decision-Making

Copyright © 2017 by Kathy McNamara;  
please do not reproduce any portion of this  
presentation without permission.

# Learner Objectives

**Understand the nature of “scientific thinking” as the basis for ethically appropriate practice.**

**Be able to identify specific cognitive errors and biases that influence school psychologists’ and teams’ decisions.**

**Be able to cite strategies that minimize the impact of bias on the decision-making process**

# Foundations of School Psychological Service Delivery

## Legal, Ethical, and Professional Practice

- Knowledge of the history and foundations of school psychology; multiple service models and methods; ethical, legal, and professional standards; and other factors related to professional identity and effective practice as school psychologists



# Ethics & Scientific Thinking

- “Ethical practice” is usually understood to mean knowledge and application of the “rules” of practice, so we study the rules and use them to make decisions
- The “rules” **do** outline for us four principles that govern practice and decisions
  - Respecting the Dignity and Rights of All Persons (identify, address, and minimize bias)
  - Professional Competence and Responsibility (engage only in evidence-based practice)
  - Honesty and Integrity in Professional Relationships (collaborate as team member)
  - Responsibility to Schools, Families, Communities, the Profession, and Society (self-monitor and be a continuous learner)
- We are going to examine the limitations of using what we see and think to make decisions, and develop an understanding of how “scientific thinking” enables us to pay attention to and minimize these limitations, whether we act individually or as members of groups or teams.

## Scientific Thinking: The Basis for Ethical Practice

## Sample Decisions Made by School Psychologists and School-Based Teams

Decision	Individual School Psych	Team or Group
Would Differential Reinforcement be a good choice of intervention for this child's disruptive behavior?	X	X
Is this number of progress monitoring data points enough to determine if the intervention is effective?	X	
Is the "evidence base" for our Tier 2 intervention strategies adequate?	X	
Based on the results of our Fall screening, which students should receive Tier 2 intervention?	X	X
Is the new reading program better than the old program?	X	X
What are the possible interventions for a child's problem, and which are likely to be accepted by teachers?		X
If the parents request immediate "testing," should we comply with their request?		X
Does this child meet the eligibility criteria for SLD?	X	X

## NASP Ethical Standard: Responsible Assessment Practices

- Data-driven decisions should be based on results of assessment measures that have been *validated* for the purpose for which they are being used.
- For example, use of IQ tests for the purpose of classifying children for special education purposes ... a valid practice?

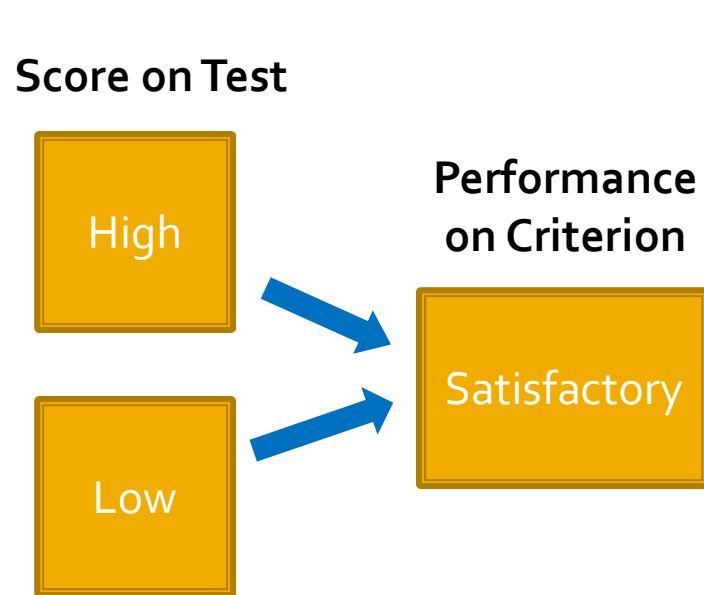
# Decision Validity

Is a decision defensible on the basis of *the technical adequacy of the methods used* to make it (assessment measures, decision rules), the *manner in which methods were applied*, and the *outcome* to which the decision will lead?

## Validity of the measurement method

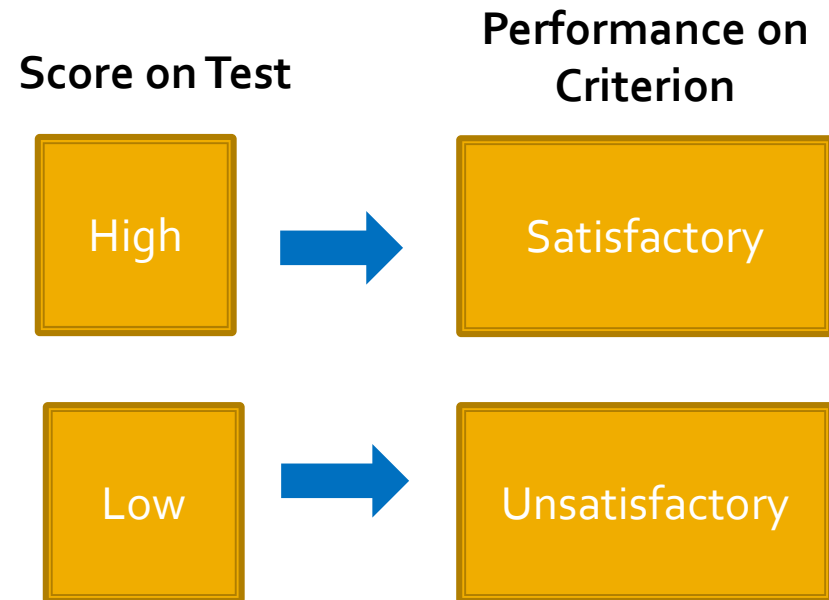
- Although far from perfect, IQ tests yield scores that are excellent predictors of academic achievement and job performance across just about every major occupation.
- However, there is a difference between the *average* performance of African-American and White students on IQ tests; and evidence of differential performance at the *item* level. Does this mean the tests are biased, and therefore invalid for African-American students?
  - Larry P. vs. Riles (1972) decision found that a test is *unbiased* only if it yields the **same pattern of scores** when administered to different groups of people; since IQ tests yield a different pattern of scores between groups, they were judged to be biased (Bersoff, 1981)

# Validity of measurement method (Is it biased?)



## BIASED TEST

Group average scores are different, but performance on criterion (e.g., success in school) is the same.



## UNBIASED TEST

Group average scores are different, AND performance on the criterion is different.

# Validity of application; validity of outcome

## Validity of the manner in which the method is applied

- Does the method yield accurate results (predictions) when used with particular groups or individuals (e.g., English Language Learners)?

## Validity of the outcome to which use of the method leads

- Does the use of standardized, norm-referenced IQ tests contribute to disproportionate representation of African-American children in special education programs?
- If so, why? Is it possible that group differences in IQ scores are due not to characteristics of children, but to environmental influences such as unequal educational opportunities, inadequate instruction, etc.?

AND ...

- Does placement of children in special education programs adequately address their behavioral and mental health needs?

# Measurement Error

## Ensure adequacy of data before making decisions

Concern about reliability across CBM “equivalent” reading passages (Christ & Ardoyn, 2009); reliability of R-CBM “trend” (slope of trendline) with too few data points (Ardoyn et. al, 2013)

- 8 to 10 data points, collected over at least 6 weeks, for academic measures
- 3 data points for baseline, followed by about 5 progress monitoring data points, for behavioral measures
- As decisions become “higher stakes,” increase the number of data points, and employ corroborative data sources

# Errors of Knowledge (easily fixed!)

Fact-based solutions (not opinion-based; not negotiable)

Requires ongoing attention (evolving knowledge base; professional development)

Requires familiarity with standards to establish an “evidence base” (peer-reviewed, methodologically sound, replicable studies)

Requires familiarity with “evidence-based practice” resources

Amenable to “training” or educational interventions

# Examining the Evidence Base for Aptitude-Based PSW Approach to Identifying Specific Learning Disabilities

- **Third Method: Alternative, research-based procedures (PSW) *Aptitude-Based* (uses measured “cognitive processes”)**

Low academic achievement is related to discrepancies in cognitive abilities: “Cognitive weakness in a sea of strengths”

- **Concordance-Discordance Method (C/DM; Hale, et. al)**
  - Analyzes differences between scores on measures of “cognitive processes” (C-DM; Hale)
- **Discrepancy/Consistency Method (D/CM; Naglieri, et. al)**
  - Analyzes differences between scores on achievement measures and measures of factors of Planning, Attention, Simultaneous, and Successive (PASS) factors, and between PASS factors
- **Cross Battery Assessment (XBA; Flanagan, et. al)**
  - Based on Cattell-Horn-Cattell (CHC) theory of intelligence, using 7 cognitive clusters from Woodcock-Johnson III cognitive measures; analyzes standard (norm-referenced) scores on tests; SLD if a cognitive cluster deficiency exists within otherwise normal profile.

Stuebing, K., Fletcher, J., Branum-Martin, L., & Francis, D. (2012). Evaluation of the technical adequacy of three methods for identifying specific learning disabilities based on cognitive discrepancies. *School Psychology Review*, 41, 1, 3-22.

McGill, R., Styck, K., Palomares, R., & Hass, M. (2016). Critical issues in specific learning disability identification: What we need to know about the PSW model. *Learning Disability Quarterly*, 39, 3, 159-170.

# The Evidence

- Models tend to work well in identifying “Not SLD” (true negatives), but poorly in identifying “SLD” (true positives), and they tend to over-identify as SLD cases that are not SLD (false positives).

Model	Sensitivity or Positive Predictive Value (Probability that case is true SLD, given that test results indicate SLD; True Positives)	Specificity or Negative Predictive Value (Probability that case is not SLD, given that test results indicate not SLD; True Negatives)
C/DM	24%	99%
D/CM	17%	99%
XBA	48-53% (depending on cognitive cluster used)	96-97% (depending on cognitive cluster used)

- Substantial concern about the “treatment validity” of PSW results; that is, the obtained information does not inform intervention planning or treatment selection in any meaningful way (i.e., using PSW for this purpose does not result in intervention outcomes that are any better than what could have been obtained without PSW information).

## **Rank order the following factors and practices in terms of their impact on student achievement:**

- A. Instruction individualized for students.
- B. Extent/quality of parental home supervision.
- C. Teachers' use of formative evaluation.
- D. School financial support.
- E. Retention.
- F. Classroom strategies to decrease disruptive behavior.
- G. Circumstances of students' home environments.

# Rank order based on research (Hattie, 2009)

1. Classroom strategies to decrease disruptive behavior ( $d = .93$ )
2. Teachers' use of formative evaluation ( $d = .90$ )
3. Circumstances of students' home environments ( $d = .57$ )
4. School financial support ( $d = .23$  across all categories of spending, but teacher salary and classroom supplies showed strong positive effect)
5. Instruction individualized for students ( $d = .23$ )
6. Extent/quality of parental home supervision ( $d = .17$ )
7. Retention ( $d = -.16$ )

## So, many issues can be addressed and resolved through improved *knowledge*, but ...

- “Research can generate crucial information on ... incidence, effectiveness, and consequences ...
- “The evidence-based practice agenda is not just about adopting and implementing research-supported practices. It is about our way of *thinking scientifically* to reduce bias and errors in our practice” (Kratochwill, 2012, p. 38, emphasis added).

# Errors of Perception and Cognition

“The scientific method is a toolbox of skills  
designed to prevent scientists from  
fooling *themselves*”

(Lilienfeld, et. al, 2010, p. 9)

# Heuristics (“shortcuts and fixes”)

- Humans are “wired” to search for patterns in information, and do so reflexively
- The use of “heuristics” (mental shortcuts) conferred a survival advantage in quickly perceiving patterns, but their persistence and influence in learning, perception, and memory contributes to bias and error ...
- Which, in turn, contributes to less-than-optimal decisions, particularly in situations requiring consideration of complex and sometimes conflicting information (sound familiar?)

# Heuristics

(perceptual and cognitive “shortcuts and fixes”)

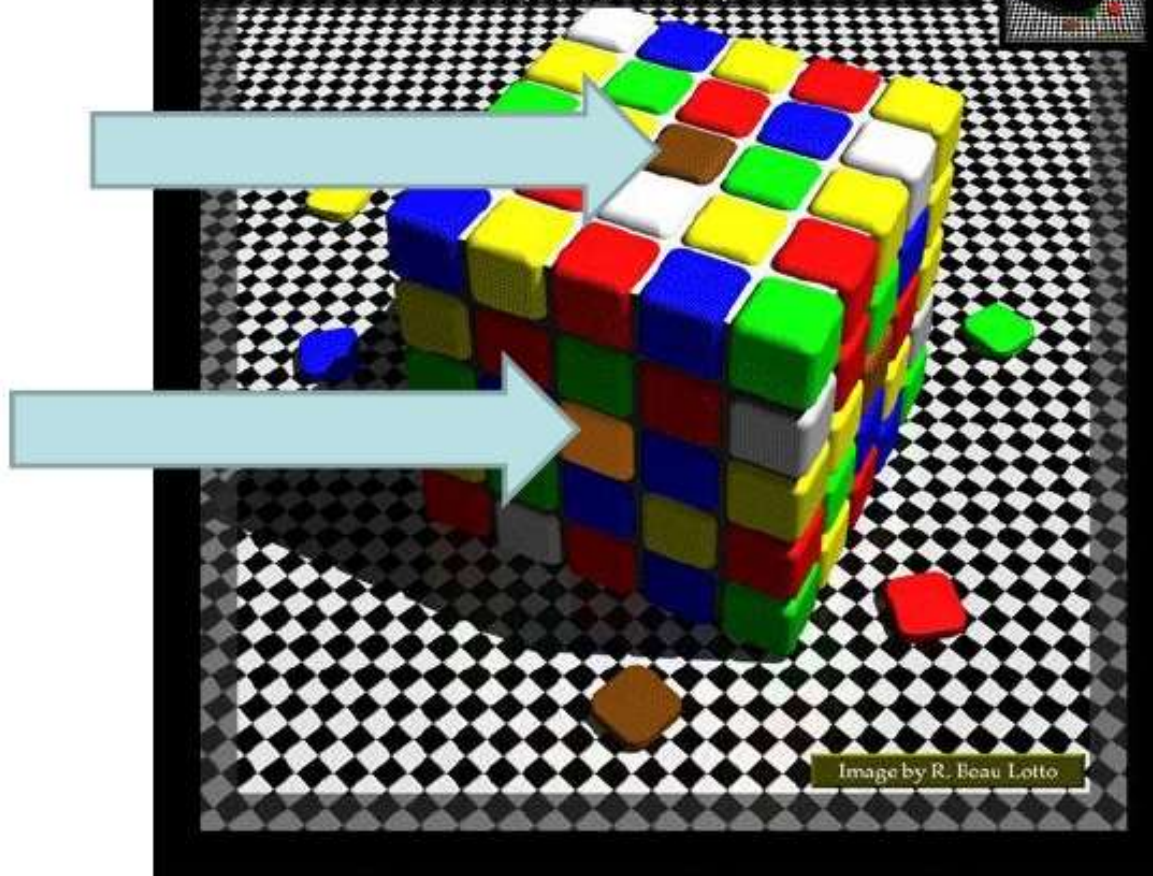
- If you were traveling from Wichita KS to Fort Lauderdale FL, in what direction would you go?
  - Heuristic: You know Wichita is in the middle of the U.S., and Florida is near the bottom of the country, near the ocean, so you (correctly) guess “southeast.”
- If you were traveling from Reno NV to San Diego CA, in what direction would you go?
  - Heuristic (“shortcut”): You know NV is north of CA and not on the ocean, but CA is on the ocean, so you (incorrectly) guess “southwest.” It is actually “southeast.”
- Here, your “heuristic” is only partly right, because although most of CA is indeed west of NV, the bottom portion is actually east of NV.

Lilienfeld, S., & Lynn, S.J. (2015). Errors/Biases in Clinical Decision Making. In R. Cautin & S. Lilienfeld, The Encyclopedia of Clinical Psychology (1-9). NY: Wiley.

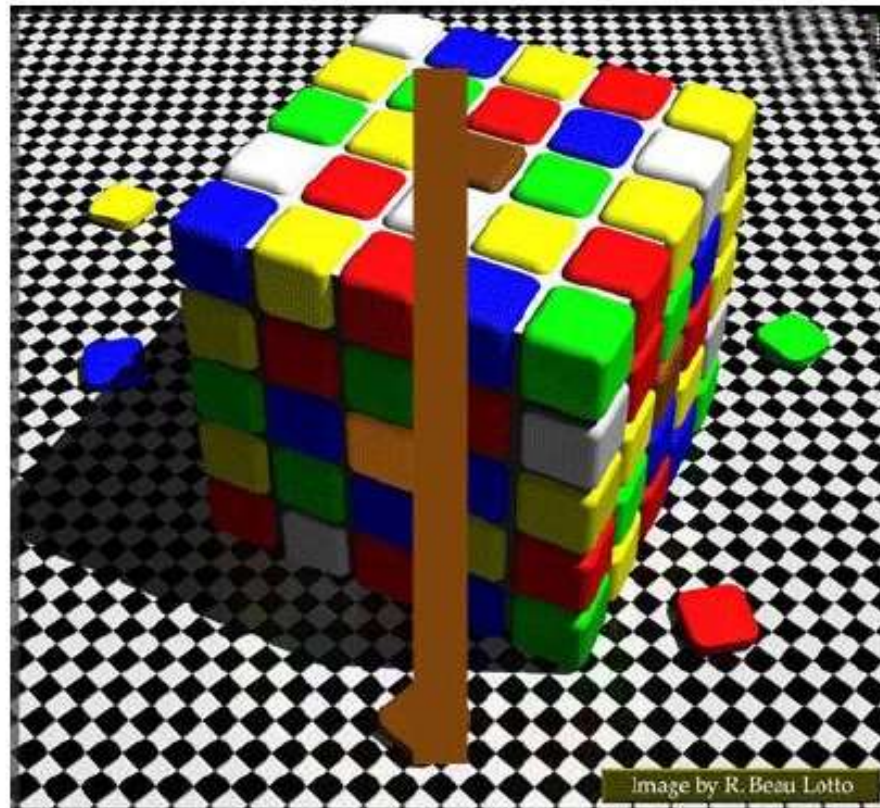


The central squares on the upper and lower surfaces of this cube appear very different in colour: Brown on the top and bright orange on the bottom. Move your mouse over the 'mask' to reveal their 'true' physical similarity.

'MASK'







The perceptual system responds automatically to the darkening effect created by the shadow surrounding the lower square by making it appear lighter than it is.

# Representativeness Heuristic

- Concluding that two events (or qualities) belong together because of some superficial resemblance or quality, or because they are in the same “mental category”
  - Examples:
    - Children from low-income or transient home circumstances ... are assumed to have poor hygiene, based on the association we make between poverty and a dirty, unkempt appearance; or because poverty and homelessness bring to mind the kind of people encountered while working in a soup kitchen
    - At a fund-raising event, you are introduced to a loud, assertive, opinionated, obviously intelligent and self-confident woman, and you assume that she is either an attorney or an activist (when she is, in fact, an elementary school teacher).

# Biases

- Tendency to draw conclusions, or to make judgments/decisions based on some factor of which we are unaware

# Confirmation Bias

- Tendency to test an explanation by looking for more instances of when it's true, than by seeking instances of when it's false.
  - Contrary to the scientific method, in which we do our best to prove that we are wrong!
- Seek (and preferentially notice) information that confirms initial impressions;
- Interpret ambiguous evidence consistent with initial impressions;
- Tend to forget evidence that contradicted our initial impression
- Over-analyze data or “cherry pick” information that is consistent with our hypothesis.
- Notice ... Do you want your explanation to be true?

“Morton’s Demon was a demon who sat at the gate of my sensory input apparatus and if and when he saw supportive evidence coming in, he opened the gate. But if he saw contradictory data coming in, he closed the gate. In this way, the demon allowed me to believe that I was right and to avoid any nasty contradictory data” (Morton, 2002).

- Is the population of Venezuela greater or less than 65 million?
- How many people do you think live in Venezuela?
- ANSWER: About 28 million



- But MOST people guess a number that is much higher than 28 million ... a number closer to 65 million.
- Because, while you don't know the population of Venezuela, you heard the number "65 million," and that influenced your answer!

# Anchoring Bias


- Tendency for initial information to serve as an “anchoring (reference) point” from which subsequent judgments or decisions are made
  - Price tags in stores or auto showrooms that show “Manufacturers Suggested Retail Price”
- Search for an answer begins with information that is immediately available; adjustments (if any) are made to that initial information.



The following ten corporations were ranked by Fortune magazine to be the largest 500 United States based firms according to sales revenues for 2003 :

- Group A: Reebok International, Hilton Hotels, Starbucks, Radioshack, Hershey Foods
- Group B : CoconoPhillips, American International Group, McKesson, AmerisourceBergen, The Altria Group

Which group of organizations (A or B) had the *larger* total sales revenues in 2003? How do you know?

- 
- **True or False: There has been a recent, dramatic increase in the percentage of children with autism.**
  
  - **Do more people in the U.S. die each year from cause “A” or cause “B”?**
    - (A) Murder    (B) Diabetes
    - (A) Murder    (B) Suicide
    - (A) Car accidents    (B) Abdominal Cancer

- Regarding rates of autism ... False: Although the US Dept. of Education reported a 657% increase in rates of autism (1993 to 2003), it is due primarily to changes in diagnostic criteria (DSM) and increased reporting (IDEA).
  - In addition to increased reporting, there has been increased attention to autism (films, television), and depictions are typically vivid and anecdotal in nature.
  - When *diagnostic criteria were held constant* over 7 years, no increase in autism prevalence (Chakrabarti & Fombonne, 2005; Gernsbacher, Dawson & Goldsmith, 2005). Further, increase in autism rate has been matched by a decrease in the number of diagnoses of learning disability and “mental retardation” combined (Bishop, et. al, 2008).
- Regarding causes of death ... Most people answer “B” for the first pair, and “A” for the other two pairs ... but the correct answer in all three pairs is “B.” Why? Murder more likely than suicide to be publicized/discussed; car accidents more shocking and memorable than cancer, the common cause of many deaths.

# Availability Bias

Judging the likelihood of an event (or accuracy of a statement) on the basis of how easily or readily it comes to mind (including believing a statement to be true if it's been repeated often enough)

## ***Examples of beliefs influenced by the availability heuristic:***

- Homeless people are more likely than non-homeless people to be mentally ill (media portrayals of or personal encounters with homeless people who were behaving oddly are more likely to come to mind)
- Gun violence, especially in schools, occurs more frequently than in the past: Media reporting of school shootings leads most people to believe that gun violence is increasing, although it has decreased in the past 20 years, and is still a very rare phenomenon in schools (Cornell, U VA Youth Violence Project, 2015)

A mock trial defendant was accused of “driving under the influence” of alcohol, resulting in his running a stop sign and colliding with a garbage truck. His blood alcohol level at the time of the collision was (is) unknown.

**Two separate “juries” were formed, and each was given a different description of the defendant’s behavior at a party prior to the accident.**

*Jury A. On his way out the door, the defendant staggered against a serving table, knocking a bowl to the floor.*

*Jury B. On his way out the door, the defendant staggered against a serving table, knocking a bowl of guacamole dip to the floor and splattering guacamole on the spotless white carpet.*

**How did each jury rule on the question of the defendant’s guilt?**

- The effect of the “Availability Bias” becomes more pronounced when the depiction of an event is more vivid, detailed, or emotionally charged.

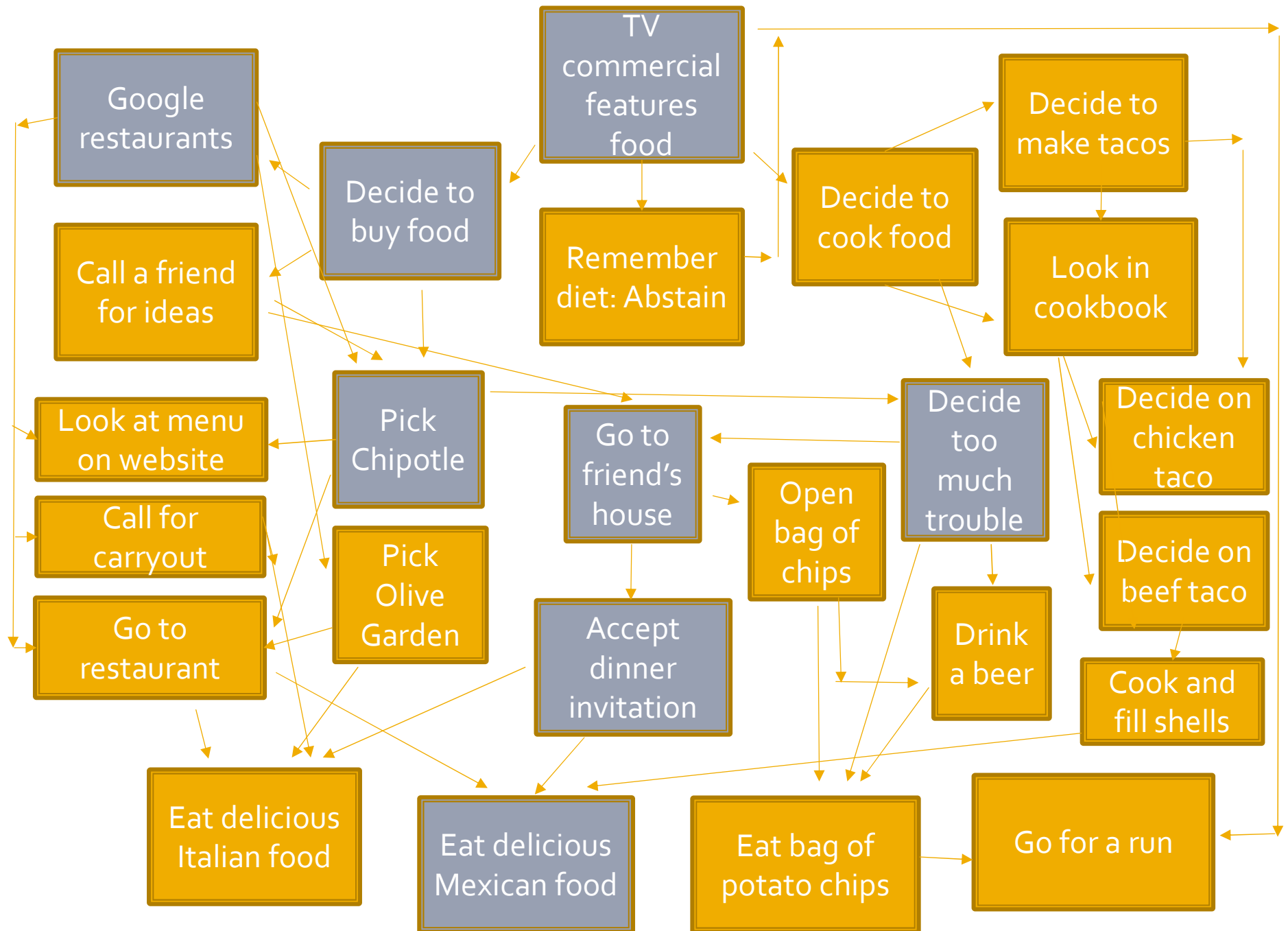
# Making Predictions?

- A friend is eating delicious Mexican food. You can trace the history of your friend's decision to eat this food all the way back to the time a few hours ago when she saw a tantalizing food commercial on TV.
- HOWEVER, knowing that your friend is currently watching a tantalizing food commercial on TV doesn't allow you to predict with certainty that, in a few hours, she will be eating delicious Mexican food.



# Hindsight Bias

- “I knew it all along” ... perceiving events as more predictable *after* they’ve occurred than before they occurred
  - Although it is almost always possible to look at something that is happening now, and trace back through a history of events that may have led or contributed to the occurrence of the phenomenon, it is never possible to reverse this procedure (i.e., make prediction about the occurrence of a future event based on knowledge of current events or circumstances)
    - Why? Because there are so many possibilities at each step along the way, each governed by probability and circumstances, that accurate prediction is impossible.



# Thought Error: Illusory Correlation

(Lilienfeld, et. al, 2010)

Example: Widely held belief that children learn better if the method of instruction matches their “learning style.”

	High Score	Low Score
Visual learner: Visual presentation	A (“hit”)	B (“miss”)
Visual learner: Auditory presentation	C (“miss”)	D (“hit”)

- A focus on “hits” (Table Cells A and D: memorable co-occurrences), while overlooking “misses” (Table Cells B and C: absence of memorable co-occurrences)
- “Modality preference” (among students) does not interact with teaching method; i.e., there is no difference in student outcomes when the teacher uses a method (e.g., visual presentation) designed to match the student’s preference/style (e.g., visual learner) (Pashler, McDaniel, Rohrer & Bjork, 2009).
  - Also exemplifies “representativeness bias,” because of resemblance between seeing (the student) and visual presentation of material.

# Thought Error: Inferring causation from correlation (Lilienfeld, et. al, 2010)

- **A and B are correlated, but A doesn't necessarily *cause* B.**
  - The possibilities (all of which must be considered) are:
    - A causes B (maybe ...)
    - B causes A (no ... because the cause must precede the effect)
    - **C (often unknown or unmeasured) causes both A and B (maybe ...)**
- **Example:**
  - Physical abuse in childhood (A) is correlated with aggression in adulthood (B)
  - But the “cycle of violence” explanation (A causes B), although widely believed, ignores the plausible possibility of a genetic factor that “causes” both A and B (Krueger, et. al, 2001)
- Further, the “post hoc, ergo propter hoc” error frequently occurs (A comes before B; therefore A caused B)
  - A child receives a series of vaccinations at around age 2. Not long after, the pediatrician diagnoses the child with autism. Because the diagnosis followed the vaccination, the vaccination must have caused the autism.

# Thought Error: Neglect of ...

## ■ Neglect of Missing Data

- Tendency to overlook or discount missing data/information
- Example: A school psychologist overestimates the probability that an adolescent who approaches her to talk about her depression will attempt suicide, because of her lack of contact with typical adolescents who occasionally experience depression and do NOT attempt suicide.
- Example: A school psychologist is studying the effectiveness of a classwide intervention using a pretest/posttest design. Although 62 students were enrolled in the intervention and took the pretest, 16 of them dropped out before the intervention ended.

## ■ Base Rate Neglect

- Failure to take into consideration the naturally-occurring rate of a phenomenon – i.e., although it may strike us as noteworthy and possibly rare, it actually is a relatively common occurrence.
- Example: A teacher concludes that a difficult-to-teach student must have a disability because the student's reading screening score fell below the national 25<sup>th</sup> percentile, whereas the class average score falls at the national 30<sup>th</sup> percentile – not much higher!
- Think: When hearing the sound of hoofbeats from a distance, think "horses," not "zebras." (The base rate occurrence of "horses" is much greater than that of zebras.)

# Summary of common shortcuts and fixes ...

## ■ Naïve Realism

- Tendency to believe that the world is exactly as we see it, uncontaminated by preferences, preconceptions, or interpretations (e.g., “checkerboard” demonstration)

## ■ Bias Blind Spot

- Belief that others are biased, but we are not (although can detect biases in others, we can’t detect them in ourselves)

## ■ Confirmation Bias

- Tendency to seek and believe only information consistent with our position, and to ignore or minimize other information

## ■ Availability Bias

- Making a judgment that is influenced by whatever comes to mind most easily (because of publicity, vividness, one’s own experiences, etc.)

## ■ False Consensus Effect

- Tendency to overestimate the extent to which others share our views (“ad populum” fallacy – if a belief is prevalent, it must be true)

## ■ Insider Perspective

- Regarding a given perspective as a fact, even though it is only a perception or opinion on which others (“outgroup”) might differ.
- Teacher lounge discussion: Parents who don’t attend school conferences place less importance on children’s school performance than parents who do attend.
- Neglects to consider how people outside one’s insulated ingroup might perceive the situation at hand (e.g., parents always hear bad news at conferences and go home discouraged, so they choose not to attend; would people not raised and schooled in the U.S. perceive Russia as evil)?

## Prescriptions for Thinking Scientifically (Lilienfeld, Lynn, Ruscio, & Beyerstein, 2010, pp. 251-252)

- Although first impressions may be helpful in “sizing up” people, they’re typically inadequate when it comes to evaluating scientific claims or making decisions;
- Many shared beliefs are nothing more than “urban legends,” so we shouldn’t assume they’re accurate;
- Good stories aren’t always accurate stories; media coverage, repetition, and anecdotes can lead us to over-estimate the frequency of sensational events, and under-estimate the frequency of less sensational events;
- Biased samples result in equally biased conclusions. If we’re exposed primarily to one group of people in our line of work, our perceptions of the prevalence of certain traits in people at large will be skewed;
- Certain biases, such as illusory correlation, confirmation bias, and the representativeness and availability heuristics, lead us to draw erroneous conclusions. Heuristics are helpful shortcuts, but if we rely on them blindly and uncritically, we’ll often make mistakes;
- Correlation isn’t causation, so knowing that two things are statistically associated doesn’t tell us what’s causing what. Also, just because one thing comes before another, the first doesn’t necessarily cause the second;
- Carefully conducted scientific research (although not foolproof) is our best safeguard against error.

# Errors of Social Influence

Many of the decisions in which school psychologists are involved (especially those of a high-stakes nature) are actually made by *teams*.

So, group “process” variables influence decisions, to a greater degree than the “rational” considerations that are assumed to be the strongest influence.

# Do Teams Make Better Decisions?

- Broader knowledge and skills base *should* allow group to make better decisions than each member could have made on his/her own.
- BUT ... group members tend to use discussions to communicate/advocate their own position, rather than to exchange information (of which other members may be unaware) to make a good decision.
  - Information supporting one's position is emphasized and repeated
  - With repetition, the position gains support (group moves toward consensus)
  - False consensus effect: Tendency to overestimate the extent to which others share our views
  - Once there is a predominant position, further information exchange becomes meaningless to members

## Information sharing in groups or teams

Three pieces of **favorable information about Candidate B** (B1, B2, and B3) were seen by **all of the group members**, but all four pieces of **favorable information about Candidate A** (A1, A2, A3, and A4) **were not given to everyone**. Because the group members did not share the information about Candidate A, Candidate B was *erroneously* seen as a better choice (Stasser & Titus, 1985).

Group Member	Information Favoring Candidate A	Information Favoring Candidate B
X	A1, A2	B1, B2, B3
Y	A1, A3	B1, B2, B3
Z	A1, A4	B1, B2, B3

Best solution (A) is hidden unless all available information is shared during discussion.

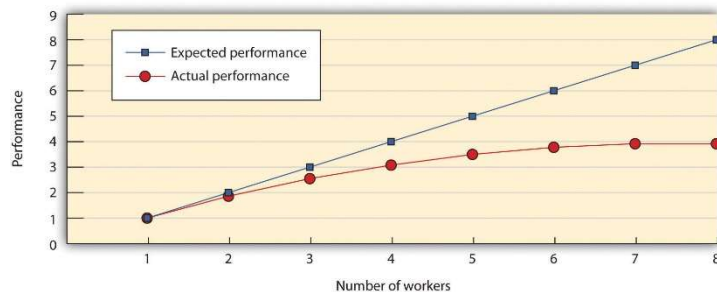
The shared information will be repeated during discussion, so it will be seen as more valid and will have a greater influence on decisions, because it is more “cognitively accessible,”  
AND

Higher status members more likely to share new information and dominate discussion, even if their information is not more important or valid (Wittenbaum, 1998; Hinsz, 1990).

# The sum of its parts ...

## Shouldn't "more people" = "better solutions"?

In a group, the cumulative effort/contribution of all members should equal or exceed the total value that could have been achieved if each member had worked on his or her own.



### The Ringelmann Effect ...

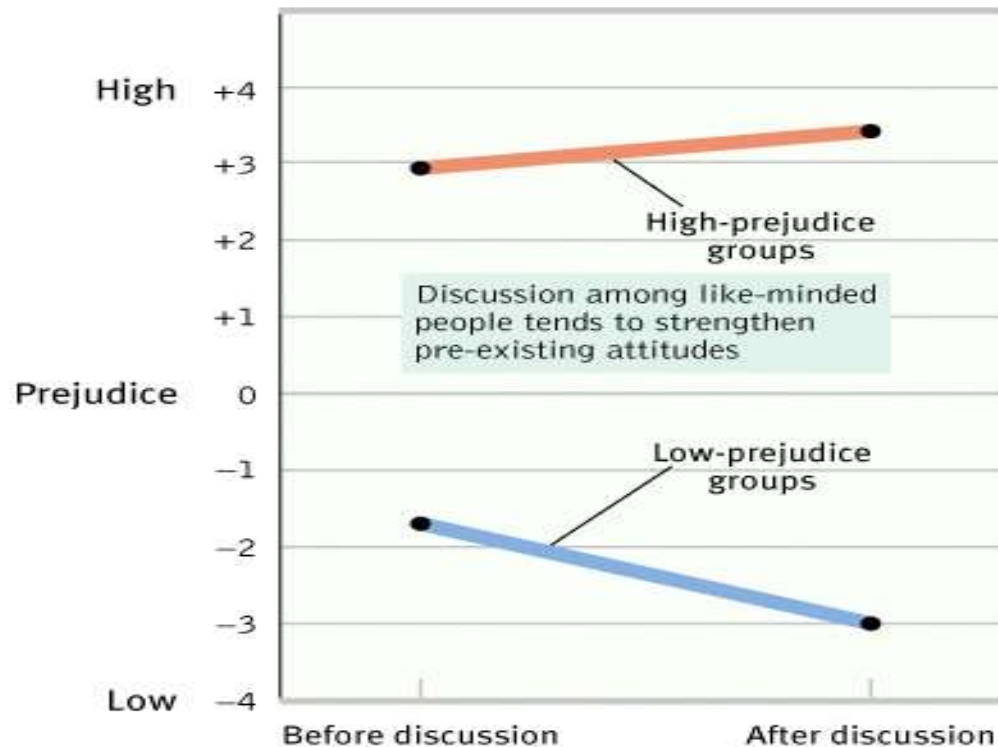
Although more men pulled harder on a rope than fewer men did, there was a substantial process loss in comparison with what would have been expected on the basis of their individual performances.

# True or False?

- **“Brainstorming” new ideas in groups works better than asking people to generate ideas on their own.**
- **False.** Most studies show that the number and quality of generated in brainstorming sessions is less/poorer than is generated by individuals (Kohn, 1990).
- “Social loafing” or “free-riding”... reduced engagement/effort among members of a group; exacerbated when:
  - “Higher status/credibility” group members are present
  - Enforcement mechanisms are absent
  - Information held by group members is redundant
  - Task-relevant skills overlap from member to member
- “Production blocking” ... listening to others or waiting to speak (resulting in forgetting your idea), or concentrating on own idea (missing others’ input)



# Group Polarization

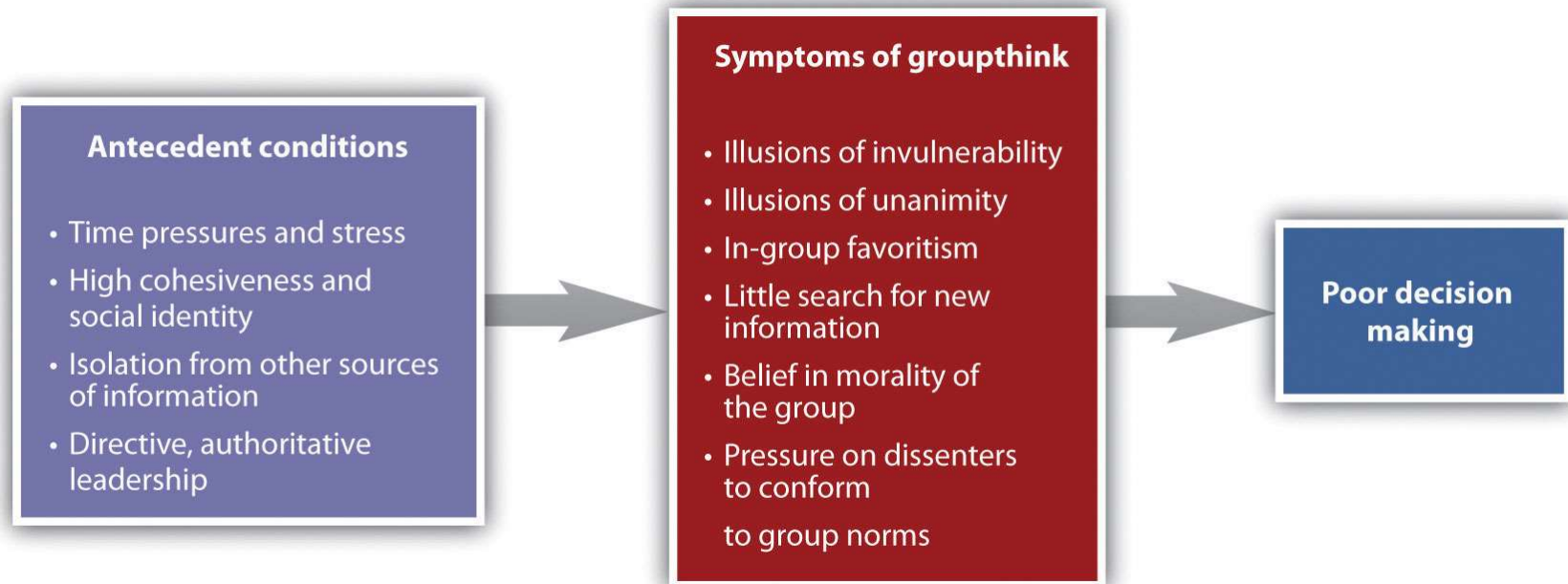


Discussions in groups tend to increase the strength of initial position/opinions (i.e., the group's position is more extreme than the position of any one member of the group).

# Groupthink



**Preoccupation with group unanimity that impairs critical thinking**



# What Helps in Team Decision-Making?

- Modify the size of the team depending on the purpose of the meeting (e.g., broad input to understand student's difficulty, vs. communication with parents to elicit their cooperation); also consider possible impact of seating, meeting length, etc. on success of meeting
- Leader refrains from giving opinions, especially initially, and concentrates instead on encouraging input
- "Nominal technique" (write ideas individually before the meeting)
- "Round robin" (Sequenced turn-taking among speakers)
- Actively promote minority dissent, rather than striving for premature consensus
- Appoint a "devil's advocate" to raise questions about group decisions
- "Consider the opposite/an alternative explanation"
- Delayed decision-making
- Periodically discuss how the team is functioning and make plans for improvement

# “Clinician, Heal Thyself”

## Clinician Bias

- Too often, clinicians ignore research findings in making decisions about a particular case. Why?
  - (1) This situation/person is unique, so the rules of probability or the findings of research don't apply to him/her;
  - (2) Research studies using “p” values to determine significance are credible, but statistics are dry and esoteric, so irrelevant when considering *human behavior*
  - (3) While a clinician knows that screening using research-based methods yields accurate results (and clinicians' judgment is more error-prone), she believes that she does a better job than the screening test in identifying problems due to her expertise and many years of experience (i.e., “bias blind spot”).

# Clinician Bias

- But ...
  - Zebras vs. Horses (again): Experts routinely over-identify “counter-examples,” with too great a focus on “unique” aspects and too little focus on commonalities, resulting in poor judgment accuracy (Grove et. al, 2000)
  - Clinicians’ routine exposure to a sample of people experiencing more severe or persistent problems leads them to erroneously view most people as less resilient than they are, and most problems as requiring more intensive intervention than is actually needed (Cohen & Cohen, 1984)
  - Clinicians are subject to the same biases (confirmation, representativeness, availability, etc.) as their colleagues and the population at large (Lilienfeld & Lynn, 2015).

# Fortunately, we can count on our clinical skills and judgment (or *can we?*) ...

- When supplied with the same case study information, and comparing “clinical method” (judgment and intuition applied to case data) to “mechanical method” (algorithm or “decision rule”), the latter is at least as (and sometimes more) accurate in making clinical predictions (psychiatric diagnoses, psychotherapy outcomes, suicidality, college and job performance, etc.) (Dawes, et. al, 1989)
- Malcolm Gladwell’s assertions in his book “Blink” notwithstanding, studies demonstrate that intuition and “hunches” lead to poor quality of decisions in professional practice, although intuition can be a useful signal that something is amiss, and that a solution, once derived, is ethically acceptable (Cottone & Claus, 2000)
- Most clinicians think their judgment improves with experience (although it doesn’t); advocate using both rule-based and clinical methods together (which works as long as both methods agree); or insist that the matter at hand is sufficiently unique as to represent an exception to the rule (which it usually isn’t)... (Dawes, 1994; Grove, et. al, 2000; Smith & Dumont, 1997)

# Ten Prescriptions for School Psychologists

- Seek out disconfirming evidence (to prove your hunch/hypothesis wrong);
- Don't become overly attached to your hypotheses ("know all theories, love some, wed none");
- Consider rival hypotheses (accept hypothesis only if it beats at least one other rival hypothesis);
- Don't cherry-pick (examine *all* evidence/data);
- Put your intuition to the test (hunches may be a good starting point, but they don't work well for decision-making);
- Be skeptical of clinical judgment and long-standing clinical wisdom ("eminence-based practice");
- Be aware of the existence of blind spots (run ideas past others to detect weaknesses or biases);
- Encourage dissent (reinforce others who offer alternative views);
- Quantify, quantify, quantify (assess "impressions" numerically; measure outcomes);
- Maintain a self-critical attitude (willingness to acknowledge that one might be mistaken), and be willing to change beliefs.

# References & Further Reading

- Burns, M., Jacob, S., & Wagner, A. (2008). Ethical and legal issues associated with using response-to-intervention to assess learning disabilities. *Journal of School Psychology, 46*, 263-279.
- Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. New York: Routledge.
- Jacob, S., Decker, D. & Hartshorne, T. (2010). *Ethics and law for school psychologists*. New York: Wiley.
- Kratochwill, T. (2012). Comments on "Distinguishing science from pseudoscience in school psychology: Science and scientific thinking as safeguards against human error": Evidence-based interventions for grandiose bragging. *Journal of School Psychology, 50*, pp. 37-42.
- Lilienfeld, S., Ammirati, R., & David, M. (2012). Distinguishing science from pseudoscience in school psychology: Science and scientific thinking as safeguards against human error. *Journal of School Psychology, 50*, pp. 7 – 36.
- Lilienfeld, S., & Lynn, S. (2015). Errors/Biases in clinical decision making. In R. Cautin & S. Lilienfeld (Eds.), *The Encyclopedia of Clinical Psychology*, pp. 1-9. New York: Wiley.
- Lilienfeld, S., Lynn, S.J., Ruscio, J., & Beyerstein, B. (2010). *50 great myths of popular psychology*. West Sussex, UK: Wiley-Blackwell.
- Mlodinow, L. (2008). *The drunkard's walk: How randomness rules our lives*. New York: Vintage.
- Nisbett, R. & Ross, L. (1980). *Human inference: Strategies and shortcomings*. Englewood Cliffs, NJ: Prentice.
- Watkins, M. (2009). Errors in diagnostic decision-making and clinical judgment. In T. Gutkin & C. Reynolds (Eds.), *The handbook of school psychology* (pp. 210-229). New York: Wiley.