


# Possible etiologies for low correlations between diagnostic performance and diagnostic confidence

Frank Papa, DO, PhD

University of North Texas Health Science Center



# Diagnostic error is a major health concern

- ▶ DDX error may be the third leading cause of death
  - ▶ In 2015, the National Academy of Medicine (NAM) subsequently ...
    - ▶ Suggested that American medical training programs and licensing boards are “not adequately preparing individuals to become skilled diagnosticians”
    - ▶ Recommended that new approaches to diagnostic instruction and assessment, at all levels of medical training (undergraduate, residency and continuing medical education), “employ approaches that are aligned with evidence from the learning sciences”.
- 



# Assumptions that could drive improvements in DDX performance

- ▶ First, accepting that differential diagnose (DDX) represents a categorization task
- ▶ Second, appreciating that DDX largely represents decision making under uncertainty because the majority of human illnesses/diseases represent 'ill-defined' categories.
  - ▶ That is, categories having neither 'individually necessary features', nor 'combinations of sufficient features' by which a correct diagnosis can be reliably rendered.
- ▶ Third, the ill-defined nature of most human illnesses thereby makes DDX an inherently difficult and error prone task.
  - ▶ A task arguably calling for, if in need of, the development of the decision maker's capacity to 'self-monitor' their decisions ...
  - ▶ A construct called Confidence?



# Assumptions that could drive improvements in DDX performance

- ▶ Fourth, in accordance with Dual Processing Theories (DPT, a framework for studying judgements involving ill-defined categories), accepting that categorization proceeds via two distinct cognitive mechanisms (System 1 and System 2).
  - ▶ System 1 enables the decision maker to estimate the degree to which a given instance approximates or is similar to either:
    - ▶ 1) One or more instances (exemplars) previously encountered and correctly categorized, and/or
    - ▶ 2) The prototypical portrayal of one or more of the disease categories under consideration.
  - ▶ System 2 enables the decision maker to estimate 'analytically' the degree to which a given instance is likely/probably a member of one or more of the disease categories under consideration.




# Assumptions that could drive improvements in DDX performance

- ▶ Fifth, despite differences in the knowledge and information processing routines used to arrive at a categorization decision, the author suggests that both System 1 and System 2 share in common information processing/reasoning sub-systems attempting to determine ...
  - ▶ Where along a gradient (i.e., 'typicality/likelihood/probability') an unknown instance belongs.
  - ▶ Partial evidence for this position is based upon research revealing that in terms of DDX, the more a patient's signs/symptoms approximate those characteristic of a given disease's prototype, the more likely it will be correctly diagnosed; the less typical the case, the less likely.
  - ▶ Thus, diagnostic accuracy is in part, a function where along a typicality/likelihood/probability gradient, an unknown case resides



# Assumptions that could drive improvements in DDX performance



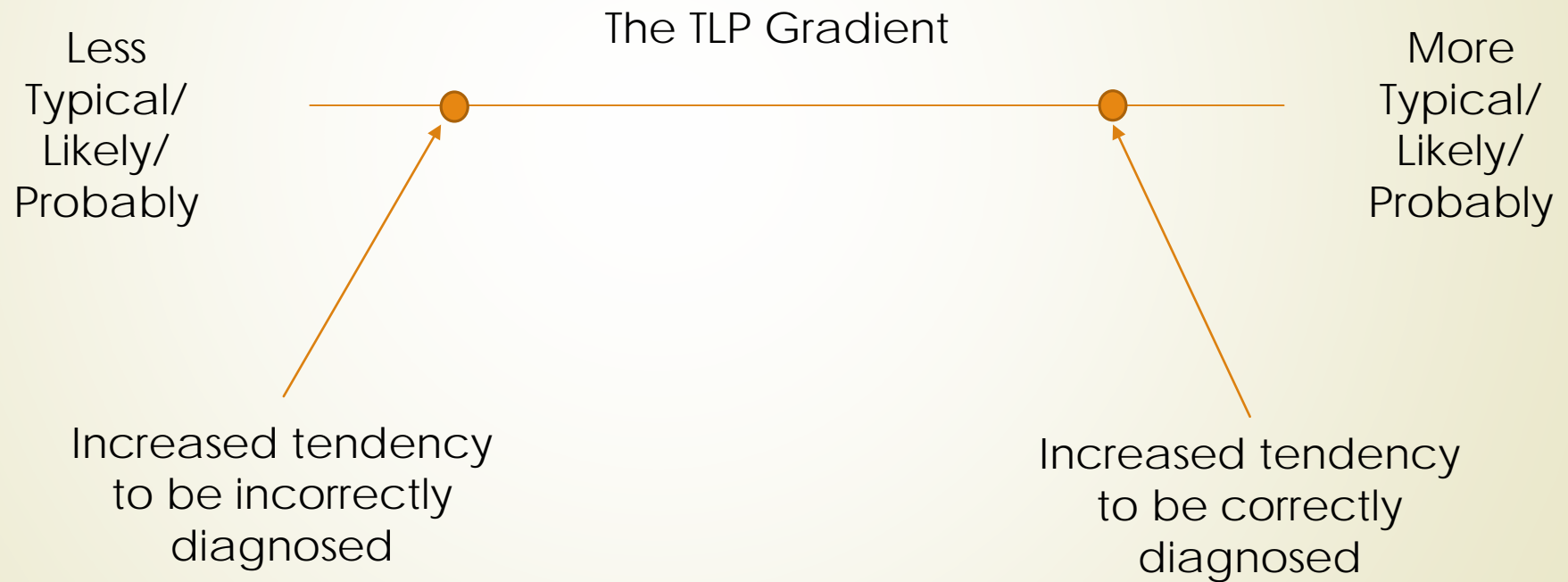
Less  
Typical/  
Likely/  
Probably

The TLP Gradient

---

More  
Typical/  
Likely/  
Probably

# Assumptions that could drive improvements in DDX performance





# Assumptions that could drive improvements in DDX performance

- ▶ Sixth, diagnostic accuracy (categorization) involving ill-defined categories, is in part, a function of the features comprising an unknown instance. That is, in part ...
  - ▶ A function of the point along a typicality/likelihood/probability gradient where the instance resides.
- ▶ Diagnostic accuracy (a dependent variable), is in part, a function of an independent variable representing where along this typicality/likelihood/probability gradient, an unknown instance resides.



# Assumptions that could drive improvements in DDX performance

- ▶ Categorization Accuracy (CA);
- ▶ The decision maker's Knowledge Base (KB);
- ▶ The decision maker's Information Processing Mechanism (IPM);
- ▶ The point along a typicality/likelihood/probability gradient (TLPG) where the instance resides ...

$$\text{▶ } CA \mapsto f(KB) + (IPM) + (TLPG) + E$$



# Assumptions that could drive improvements in DDX performance

- ▶ Another framework used to determine possible causes of sub-optimal diagnostic performance involves the notion of confidence ....
  - ▶ More specifically, diagnostic error is related to 'over confidence'.
  - ▶ Corroborating evidence comes from generally low correlations between diagnostic performance and confidence.
    - ▶ That is, individuals achieving low levels of diagnostic performance reporting high levels of diagnostic confidence.
- ▶ Error is in part, the result of 'poor calibration' between an individual's performance and confidence in their performance, has been reported in a variety of settings.



# Assumptions that could drive improvements in DDX performance

- ▶ So, what is confidence?
- ▶ Confidence may be a manifestation of one's self-monitoring capabilities or executive control skills.
  - ▶ Are individuals demonstrating a poorly calibrated relationship between their performance and confidence, in part, the product of programs providing little to no training in self-monitoring/executive control skills?
- ▶ There may be a means of training tomorrow's practitioners to be better calibrated in terms of their diagnostic performance and confidence.



# Assumptions that could drive improvements in DDX performance

- ▶ More specifically, tomorrow's health care providers could be explicitly trained in ...
  - ▶ DDX as a categorization task
  - ▶ The ill-defined nature of human illnesses
  - ▶ The error prone nature of DDX



# Assumptions that could drive improvements in DDX performance

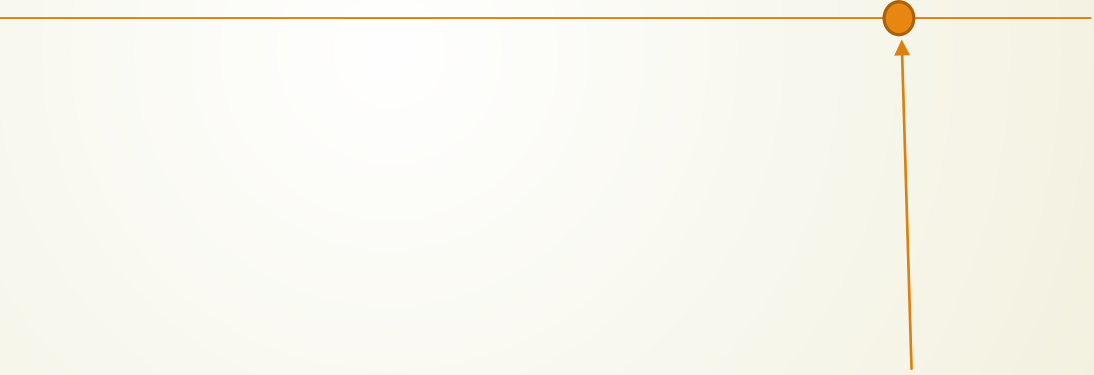
- ▶ Partial Solution ...
  - ▶ Training in how to construct a case typicality/likelihood/probability estimate for any given case
  - ▶ Students could use their own case typicality/likelihood/probability estimate to reflect upon, and determine how confident they might/should be in terms of their diagnosis for the case at hand, and ...
  - ▶ Then render their diagnosis

# Assumptions that could drive improvements in DDX performance

Less  
Typical/  
Likely/  
Probably

The TLP Gradient


More  
Typical/  
Likely/  
Probably



Student's estimate of the  
typicality/likelihood/probability that  
the case is an instance of Disease 'X'



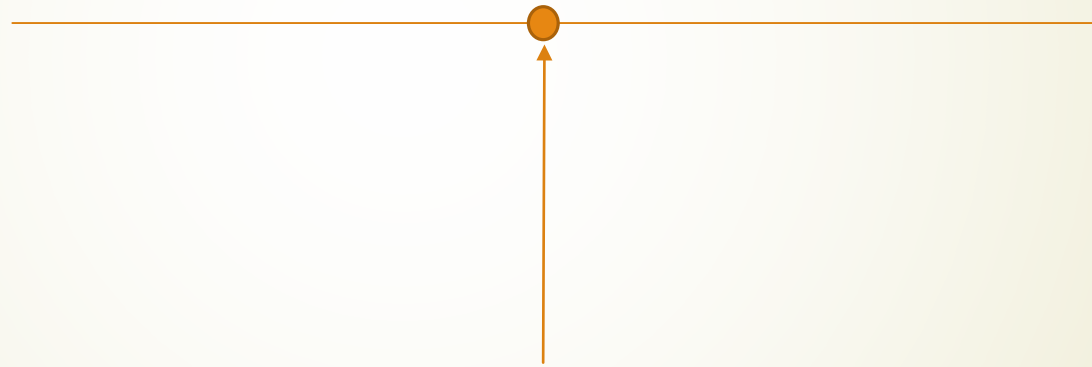
# Assumptions that could drive improvements in DDX performance

- ▶ Students could then be given a technology-based, objective estimate of that case's typicality/likelihood/probability and then asked to reflect upon and determine how confident they might/should be in terms of their diagnosis for the case at hand.
- 

# Assumptions that could drive improvements in DDX performance

Less  
Typical/  
Likely/  
Probably

The TLP Gradient



More  
Typical/  
Likely/  
Probably

Technology-based estimate of the typicality/likelihood/probability that the case is an instance of Disease 'X'

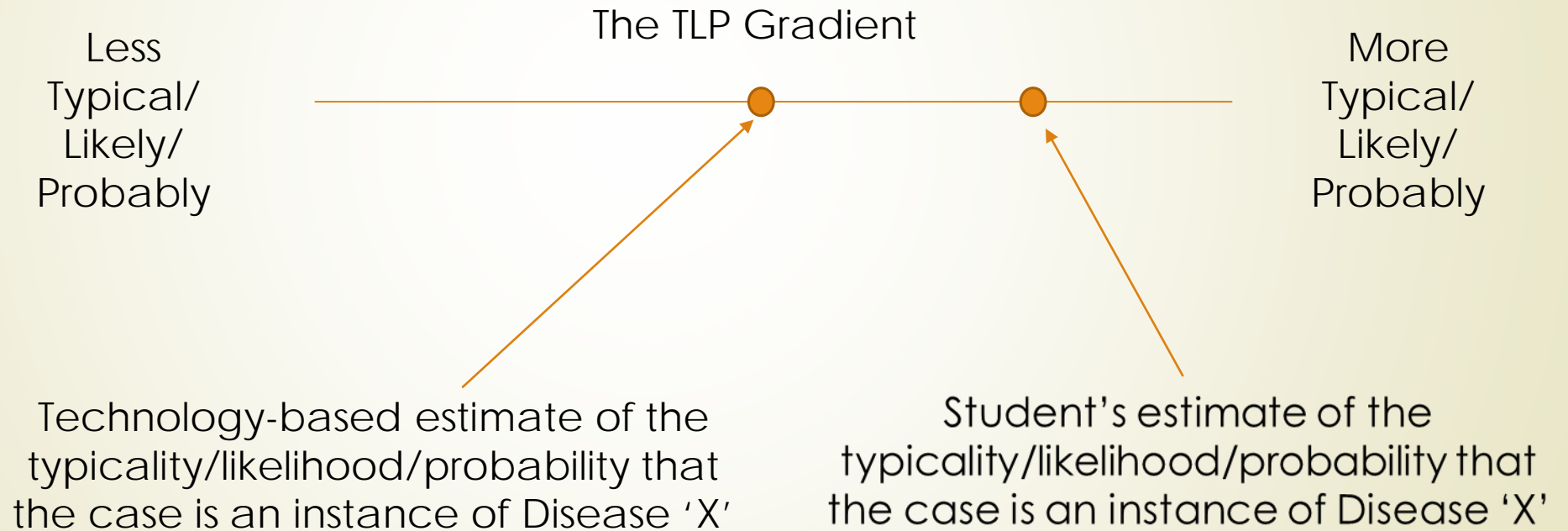




# Assumptions that could drive improvements in DDX performance

- ▶ Performance feedback in the form of 'your diagnosis is correct/incorrect', in conjunction with
  - ▶ A reminder of their own estimate the case's typicality/likelihood/probability ',
  - ▶ And the technology-based, objective estimate of that case's typicality/ likelihood/probability ...

# Assumptions that could drive improvements in DDX performance





# Assumptions that could drive improvements in DDX performance

- ▶ A comparison of these two estimates could then be used to support the students' in better calibrating their confidence (i.e., self-monitoring/executive control skills) with their performance ...
- ▶ And hopefully, improve their self-calibration capabilities (evolving self-monitoring/executive control skills), and thereby improve their diagnostic performance.





# Assumptions that could drive improvements in DDX performance





# Assumptions that could drive improvements in DDX performance

