

Short Commentary

When Doctors Try to Count the Uncountable

Larsen T^{1*} and Holgersen RB²¹Partner at Beier & Larsen, Conductor, Ph.D., Denmark²Partner at Beier & Larsen, MD, Associate Professor, Master of Health Professions Education (MHPE)

*Corresponding author: Ture Larsen, Partner at Beier & Larsen, Conductor, Ph.D., Oldfuxvej 20, 2400 NV, Denmark

Received: November 27, 2020; Accepted: December 21, 2020; Published: December 28, 2020

Three decades of research of physician's leadership in acute scenarios are based on the assumption that everything can be made up in numbers. Maintaining this way of looking at the world indicates that it might be a good idea to consider if it makes sense to do as all other researchers do.

In 'Tintin and the Black Gold', there is a breathtaking scene where Thomson and Thompson drive in a jeep in the desert. They get lost and run in large circles lasting one-hour per. turn and one additional set of wheel tracks is visible in the sand. When seven wheel tracks can finally be seen, enthusiastically they confirm to each other: 'We're obviously getting near a big town' (Figure 1).

A recurring challenge in research is to create a construction of a reality where it is possible to measure what it has to do with the real world. It makes very good sense if a new medical treatment is invented and a factor is required that can measure the likelihood of it actually working. The factor medical researchers are working on to ensure that a drug has the desired effect is called 'p-value'. It is common practice for scientific results to be 'statistically significant', which is considered when the p-value is 0.05 or less.

Statistical significance and the p-value are excellent tools if used correctly. But when the p-value is used to maintain or prove something irrelevant, it is possible to become blind to the obvious. It is also possible to find research where there is doubt about causality because the researcher in an attempt to count and measure ends up having to twist the actual research question. Ultimately, the researcher may end up having to conclude on results that differ from reality and the problem the researcher originally intended to illuminate.

No Development for Three Decades

In a study in the journal Heliyon [1], we reviewed the literature of the last three decades to find out how younger physicians are trained in leadership. We researched medical, psychological and educational articles on the subject and found 40 articles written in the period 1986-2016. The results suggested that the medical world for 30 years has called for an effective way to train the clinical team leader when he has to lead acute teams in a life-threatening situation for a patient. In the case of cardiac arrest, the alarm will sound Code Red and call the professional staff. This team has a few minutes to save the patient and the situation is led by a team leader who does not have 'hands on'

himself, but focuses solely on leading. The literature agrees that his or her role is crucial to the quality of the team's work and ultimately as to whether the patient survives [2].

In the articles examined, a total of 30 of the 40 articles requested a useful training in team leadership for emergencies. The point that there is a lack of leadership training of the younger doctors was emphasized in both the oldest and the most recent study. The point was surprisingly homogenous and therefore discouraging.

Iserson stated in 1986: "...physicians are never taught clinical leadership. This is a serious deficiency in medical training. Most people are not born leaders: they must be taught leadership techniques" [3]. McCue and colleagues reported in 1986: "Leadership has been neglected as a part of the education and training of physicians" [4]. Nicksa and colleagues reported in 2015: "Traditionally, surgical education has not formally taught leadership skills ..." [5]. Leenstra and colleagues reported in 2016: "...practical guidance needed for the deliberate practice of leadership skills" [6]. Robinson and colleagues reported in 2016: "We recommend that training in and assessment of leadership skills in emergency scenarios should be an integral component of postgraduate medical training" [7].

It is not Good Enough to say that Leadership is Acquired through Experience

We found no leadership training in the 40 articles. However, we found 29 measurement tools to document that the professional clinical training, which was repeatedly changed from the originally requested training in leadership made a difference compared to the team leader. Obviously, the scientists aim to accompany Thomson and Thompson in their desert journey.

But the consensus that clinical leadership training is needed remained. The most effective way to learn leadership at the hospital, has for many years been and continues still to be, 'learning by doing'. Apparently there is a clear belief that it comes with experience. It

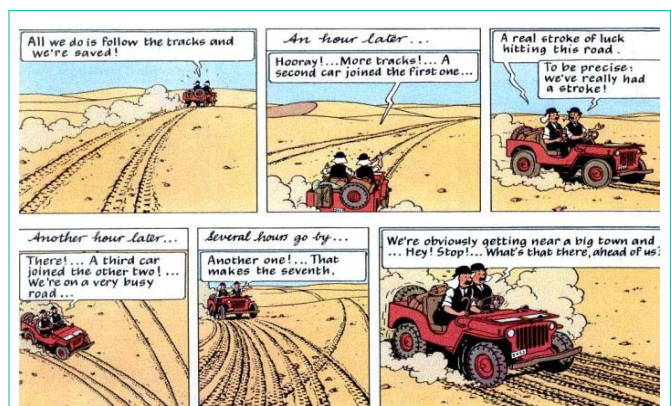


Figure 1: The cartoon elegantly demonstrates how it is possible to confirm a hypothesis and analysis without considering whether it is possible that something completely different is taking place.

just does not apply to all doctors. Like Thomson and Thompson, the research ran around itself in an attempt to document a difference using numbers without relating to the lack of training.

Impressive Numbers but They Do Not Say Anything about Leadership Training

The 40 articles we examined showed that the researchers, for example, chose to work with timekeeping on the team's work in order to be able to measure the quality of leadership. Other researchers used behavioral markers with scoring systems to measure the effectiveness of leadership. Behavioral markers are descriptions of different pre-defined appropriate types of behavior. These are collected on a piece of paper called a taxonomy where different professionals must impartially assess the quality of each person's individual behavioral marker, e.g. on a scale of one to seven. The results can be put into tables and translated into statistical significance and p-values. There have been many discussions about whether the scale should be three, five, seven or nine points. But this made no difference compared to what was asked for.

Quantitative measures (i.e. numbers) have been chosen because it is important to be able to document that the research made a difference. If this is to be proven, there must be statistical significance in the medical world, i.e. the results must be able to be translated into numbers, and if you are lucky, you will find the coveted statistical significance, where the p-value is <0.05 .

That kind of evidence makes good sense in the rest of the scientific research in the medical world. If we can prove that a new drug or procedure works with great significance, everyone will be convinced.

When we talk about what goes on between people (social interaction), it does not always make sense to count, measure and weigh.

Albert Einstein was perhaps the greatest scientist, and he pinned a note to his office wall that read: 'Not everything that can be counted counts and not everything that counts can be counted.'

A Human Error

Perhaps the research ignored the difficult core question 'How do we train in leadership?' And they replaced it with the more manageable question 'How do we measure in leadership?' This took place without anyone noticing the difference in the essentially different questions. The latter is in itself also a relatively complex maneuver, which has now been proven over time, as this challenge has kept researchers busy for 30 years, resulting in many PhD dissertations and research projects.

This is a common human error according to psychologist and Nobel Prize winner Daniel Kahneman [8], who has dealt with the psychological phenomena around judgement and decision-making and the irrational human being and human errors based on heuristics and biases.

He calls the phenomenon 'substitution', that is, when we reformulate a difficult-to-handle question and answer it with a seemingly homonymous question instead without noticing the difference.

For humans are not rational but are characterized by the reality

they are in. And in the medical world, the reality of research has become dependent on being able to document and prove significance.

Thinking, Fast and Slow

Another explanation for the research's jumping off track, can also be found with Kahneman. He distinguishes between the fast, intuitive thinking (we can name it the 'everyday brain') and the slow, rational and thoughtful thinking (we could call it the 'nerd brain').

In everyday life, we are all predominantly in the 'everyday brain' mode. In the 'everyday brain', we act on emotions and impulses, and the system is irrational but fast. The 'nerd brain' works slowly, is troublesome and requires deliberation and effort.

We experience the interaction between people in this case the acute team leader and the team members – as an 'everyday brain' function. People who need to react quickly to each other and to all the unpredictability that can arise in such a situation.

This is in stark contrast to the world of medical researchers, where intuition and sensations are difficult to recognize as legitimate or manageable quantities. But scientists find themselves deep inside the 'nerd's brain', where they strive to construct a measurable reality.

Medical Scholars Risk Running in Circles in a Parallel World

This world of research behind closed doors can thus be described as a parallel universe in relation to the real world where things have to work.

In the 40 articles we worked on, we found 29 different measuring instruments on leadership. Every research group seems to have found a new and better way to measure almost competing in making or adapting tools in this parallel world constructed deep in their nerdy brain.

Some suggest that it may be a good idea to challenge standards and expectations as well as consider whether it is a good idea to do like all the others.

Yule and colleagues stated in 2015: 'The rate of developing assessment tools in this area has outpaced development of workable interventions to improve non-technical skills in surgical training and beyond' [9].

Thus, scholars must be careful about how the p-value can be used to support claims that appear unfounded when they consider the total amount of knowledge or in plain text: keep alive some nonsense.

Nevertheless, this seems to have happened with the research in leadership for the acute team leader. We count and measure and count again to find a p-value, while the actual training in leadership remains a mirage in the desert.

References

1. Larsen T, Beier-Holgersen R, Meelby J, Dieckmann P, Ostergaard D. A search for training of practising leadership in emergency medicine: A systematic review. *Heliyon*. 2018; 4: e00968.
2. Larsen T, Beier-Holgersen R, Ostergaard D, Dieckmann P. Training residents to lead emergency teams: A qualitative review of barriers, challenges and learning goals. *Heliyon*. 2018; 4: e01037.
3. Iserson KV. Critical leadership. *The Journal of Emergency Medicine*. 1986;

- 4: 335-340.
4. McCue JD, Magrinat G, Hansen CJ, Bailey RS. Residents' leadership styles and effectiveness as perceived by nurses. *Journal of Medical Education*. 1986; 61: 53-58.
 5. Nicksa GA, Anderson C, Fidler R, Stewart L. Innovative approach using interprofessional simulation to educate surgical residents in technical and nontechnical skills in high-risk clinical scenarios. *JAMA Surgery*. 2015; 150: 201-207.
 6. Leenstra NF, Jung OC, Johnson A, Wendt KW, Tulleken JE. Taxonomy of trauma leadership skills: A framework for leadership training and assessment. *Academic Medicine: Journal of the Association of American Medical Colleges*. 2016; 91: 272-281.
 7. Robinson PS, Shall E, Rakhit R. Cardiac arrest leadership: In need of resuscitation? *Postgraduate Medical Journal*. 2016; 92: 715-720.
 8. Kahneman D. *Thinking, Fast and Slow*. 2011; 499.
 9. Yule S, Parker SH, Wilkinson J, McKinley A, MacDonald J, Neill A, et al. Coaching non-technical skills improves surgical residents' performance in a simulated operating room. *Journal of Surgical Education*. 2015; 72: 1124-1130.