

# THE STAR MODEL

A research-based review

Research Paper by Challenging Frontiers Ltd.

**October 2021**



**Challenging Frontiers**

Independent insight, advice and research



---

“

*This paper not only takes a deep dive into STAR, but the way it does this provides a blueprint to follow for any of us within the HR or L&D community striving to take a more evidence-based approach to our practice.*

Rob Baker, Chief Positive Deviant and Founder of Tailored Thinking and author of Personalization at Work.

---

“

*This paper helps to outline that despite being highly practical and widely used in various ‘people’ activities (e.g., recruitment and learning and development) there is very little scientific research around the STAR model.*

*This sort of dialogue is critical in order to not only further HR/Learning functions practices, but also help test their validity. This will help not only drive commercial and strategic organisational benefits but also avoid instances of bias and support an agenda of inclusivity.*

Patrick Mullarkey,  
L&D Lead at Snyk

## Foreword

Early in my job hunting career, I recall receiving highly prescriptive advice for interviewing – advice I still receive to this day. My mentors, friends, and colleagues have always encouraged me to use the STAR model to answer behavioural questions in job interviews. It is widely considered the sure-fire formula for interview success; missing any step of the model means losing points for an interview question, and potentially even the interview itself. While the model has long been a part of my interview preparation toolkit, both as a job candidate and hiring manager, it has never quite “grown” on me. While I appreciate the STAR Model’s methodical approach to evaluating job interview responses, I have never been wholly convinced that it is, without a doubt, the gold standard for evaluating job interviews or providing feedback. Something has always been missing for me: evidence. To enable evidence-based decision-making and practice related to hiring and coaching people, it is our responsibility as L&D practitioners to ensure there is evidence for why we should use the STAR Model, evidence that goes beyond just popular opinion.

In this paper, Sukhvinder provides a critical analysis of the STAR Model to uncover this important evidence. Using a rigorous and scientific approach, the paper presents the many gaps in evidence surrounding the STAR Model. Furthermore, he makes key recommendations for testing, refining, and measuring the STAR Model to ensure evidence-based L&D practice. We now know that there is little evidence behind the STAR Model - it’s time for us to heed Sukhvinder’s recommendations and call to action to build the evidence.

### **Gurprit K. Randhawa**

MSc, EMBA, PhD, CHE

Director  
eHealth & Virtual Care Innovation  
First Nations Authority  
Canada



## The STAR Model: why does research matter in L&D?

Learning and Development (L&D) uses many models, theories and practises in the design, development and delivery of learning solutions. But when people use these models and theories, they often neglect to ask a couple of very important questions. Where has the evidence base come from? And what research has been used to develop these models and theories?

Some models are designed using sound psychological principles. Others however, are based on a person's or consultant's own experience. There are even models that are derived from ideas.

The benefit of using research driven models and theories is that it enables us to directly attribute an intended outcome with individual models or theories, and to therefore prove the model has practical utility. (Rigdon et al., 2014).

This report explores the STAR model, examining the evidence base for the model and what the implications are for its use in L&D.

## What is the STAR Model?

The STAR model is a highly popular technique (vital, some people say) when preparing for behavioural and situational interviews. It provides a clear structure for showcasing relevant experience (Indeed, 2020) and for providing feedback in a critical yet constructive way (Pasaribu, 2019). Do a Google search and you will find many, many articles that give tips to ace interviews or give feedback to team members that reference the model.

The inventors of the STAR model, Development Dimensions International (DDI), created it as a basic and effective way to communicate during an interview or feedback. According to DDI, the acronym is as follows:

- ST:** Situation/ Task – Describe the particular situation or task to provide the initial context
- A:** Action – Provide tangible actions you or another person took in response to address the situation in hand
- R:** Result – Describe the outcomes and why they were either effective or ineffective, depending on the context



# STAR



The STAR model has a high level of applicability and is incredibly practical. It is a good example of how a simple model can improve communication. I have used the tool in many professional settings, for the training of line managers and for supporting improved recruitment practice.

The model offers a clear way for people to structure a feedback conversation, or to present a clear and coherent answer in a job interview setting.

But, this is part of the problem of the model. Which of the two purposes is the model actually meant to be used for? If it has dual purposes, where's the evidence to prove that it can yield the desired outcome in both circumstances?

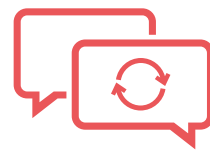
When used in recruitment settings, how has the model performed against other models that are equally simplistic? And how has it performed in feedback settings? How does it help managers differently to other feedback models?

# What are the different variations of STAR?



## STAR model for recruitment

During the recruitment process, the STAR model can be used to provide concise answers to interview questions. For instance, Doyle (2021) showed that the STAR method is very useful for answering competency based and behavioural questions. Doyle claimed that past performance is a good predictor of future performance, therefore recruiters need examples of a candidate's past work that are applicable, tangible and definitive.



## STAR model for feedback

When used for providing feedback, STAR can enable performance reviews to be conducted in a more casual or relaxed way (Clarke, 2016). According to Clarke, it can be integrated into existing performance review systems.

The main advantage of using STAR for providing feedback, according to Clarke, is that it reduces the time it takes to complete appraisal forms. And without STAR, there is the risk that feedback is too vague and lacks description. This can result in employees not knowing how they performed a task, even when they did the task or showed behaviour correctly. This would indicate that STAR is a simple and effective way to provide feedback to employees.



## The STAR/AR model

DDI also created the STAR/AR method, an extension of the original STAR method, for situations where reflection is needed. This variation of the STAR model fits well within the learning and development field.

The STAR/AR model acronym can be described as follows:

- ST:** Situation/Task - Explain the situation or task
- A:** Action - Give details about what you or another person did to handle the situation
- R:** Result - Describe the consequences
- A:** Alternative action - Discuss what could have been differently
- R:** Alternative Result - Share how the different action could have produced a better result



## STAR model for Coaching

Cited in the book, 'Coaching for High Performance: How to Develop Exceptional Results through Coaching', Cook (2009) highlighted the STAR model as a future-focused coaching model. The acronym is described as follows:

- Situation
- Target
- Action
- Results or Review

As shown, the 'T' differs. Cook (2009) said target explains how a coach might encourage their client to think about the future, temporally 'targeting' their situation. For instance, "What specifically would you like to be different?" Furthermore, 'R', as well as results, is also used to review the situation, where the participant is asked how they would feel as a result of the completed task. According to Cook, this variation of the STAR method is particularly useful for management feedback discussions where an employee may not be reaching a target.





# What is the existing research?

So, we've established that there is a lot of media and materials surrounding the STAR model in recruitment and feedback. And there is much support and practical advice for using STAR. However, there is no scientific research to support and build the STAR model and to confirm its validity and reliability.

Cook (2009) does argue for the support of STAR in coaching in her book. But, there were not any studies or research cited within the text. DDI, the inventors of the STAR model, doesn't have any published science backing either.

Conversely, the University of New Mexico Human Resource Departments has an infographic describing the STAR model within the feedback context. In agreement with the article written by Clarke (2016) with Trakstar, the infographic highlights the importance of using the acronym to phrase and set up feedback to employees and colleagues.

We cannot ignore this lack of research or studies to provide a scientific lens on the model.

“

*We cannot ignore this lack of research or studies to provide a scientific lens on the model.*



# The case for evidence-based research methods of tools in L&D

The purpose of this paper is to highlight the importance of evidence-based practice in L&D. Without such informed practice, L&D can easily be swayed by increasingly persuasive marketing and sales techniques. We have to be much more rigorous in our practice if we are to understand the inherent strengths or weaknesses in any proposed solutions or interventions.

Research helps demonstrate two important things. Firstly, what is the evidence-based approach to learning and development? And, how does this research help inform the design of learning and development solutions?

According to Kraiger and Ford (2020), within the L&D field, “the science of workplace instruction is the application of evidence-based principles that have been found to help individuals learn knowledge, skills and attitudes that impact job performance and organisational effectiveness”. However, Thompson (2002) argues that researchers often fail at reporting the reliability of such tools and seemingly don’t provide the links that demonstrate validity in their research. And if there isn’t an evidence base for a solution or intervention, it’s very likely that there isn’t any validity to the tool itself.

## When building a tool

In psychology, one of the main aims of scientific research is to solve practical problems. In this scenario, the STAR model is a tool to help elevate feedback sessions, job interview answers or coaching sessions. Building a tool, therefore, would call for sturdy theoretical and methodological grounding (Tashakkori & Teddlie, 2010).

A paper by Alavi and colleagues (2018) claims that in order for research to be meaningful, it needs to be grounded in theory. Alavi and colleagues also said that this should be an ongoing process during the design stages of any tool or model.

The research we conducted prior to writing this paper did not yield any evidence of theoretical grounding or support for the STAR model.

## When testing a tool

A paper by Mohajan (2017) suggests that reliability and validity are fundamental features in evaluation. The assessment of tools helps to reveal the accuracy – or validity – and the consistency – or reliability – of those tools (Mooi et al., 2018). The lack of such assessment of models and tools in L&D a major concern (Rigdon et al., 2014). The concern is it’s possible to make claims of efficacy without any evidence to support those claim.

According to Mohajan (2017), reliability relates to the level of stability in findings, whereas validity relates to the truthfulness. In simple terms, it’s a case of: Is the tool we’re using consistently producing the same results (reliability), and is it measuring what it should be measuring (validity)? Testing a tool in terms of validity and reliability ensures transparency, decreases bias in research (Singh, 2014) and leads to a level of quality in the interpretation of results (Cook & Beckham, 2006).

Another argument that supports the testing of measurements is that not assessing reliability or validity leaves room for error and claims of application or transferability that were not tested (Forza, 2002).





# Validity

In terms of validity, there are different forms that can suggest a tool's 'truthfulness'.

## Face Validity

One such form of validity is face validity, also known as content validity. It is defined as the apparent consistency within the definition of each construct and the items (Boyer, 2009). Put simply, does it look like it measures what it suggests it measures?

In the case of STAR, the prevalence and popularity of the model indicates that there could be a level of face validity, where the 'S' for situation directly guides a person to describe their situation, and so forth. In research, face validity is usually tested by asking researchers and participants (either from the general or from specific populations), if the description matches the item. If inconsistencies were found, definitions would be modified to reflect that. The STAR model was founded by professionals and researchers in the field – DDI – who also coined the definitions for each letter in the acronym. The fact that the model has been so extensively used by professionals, HR personnel and managers would attest to the face validity.

One thing to note however, is the variations in the STAR model, as discussed earlier on in this report – T for task or target or R for result or review. This variability in definitions may question the true face validity of the STAR model.

## Construct Validity

In research, construct validity tests if items are measuring what they are supposed to measure (Boyer, 2009). This is usually confirmed through statistical testing, such as confirmatory factor analysis. Our search did not find any evidence of this in the building of the STAR model. This finding does not confirm that the STAR model does or does not have construct validity. The lack of construct validity suggests that it is difficult to determine if the acronyms, Situation, Task, Action and Result, are the right words, if they have the right corresponding definitions, and if each word actually achieves the result they are supposed to achieve.

This isn't to say the STAR model doesn't have utility or practical application, but to highlight that there is good faith in the construct of the tool, without an evidence base to support that faith.

“  
The fact that the model has been so extensively used by professionals, HR personnel and managers would attest to the face validity.





As discussed, reliability of tool helps us understand if we can rely on a tool to reproduce the same results time and again - across different demographic groups, across different geographies and populations, across different professions, and across different organisations, settings or contexts.

Reliability is enforced by standardising the administration of a tool.

From an anecdotal and lived experience point of view, there seems to be strong reliability for the STAR model. That is to say that when people use it, time and again, for different purposes, it seems to produce the desired results. A closer look, though, suggests that the reliability of the tool lacks a strong evidence base.

### Test-retest Reliability

The importance of test-retest reliability is that it allows a researcher to utilise the same tool, in the same conditions, at different times, and the results are the same (Hobbs, 2016). Test-retest reliability therefore, assesses the consistency of the tool in question, and if the tool is representative and stable over time.

Again, statistical analysis is important to calculate a tool's test-retest reliability. And again, that seems to be missing in the case of the STAR model.

### Inter-rater Reliability

Inter-rater reliability refers to the degree to which different 'raters' or people conducting consistent observations (Frey, 2018) produce similar or different results. For example, if 20 different line managers use the STAR model for structuring feedback to their team members, do they all score the efficacy of the STAR model in similar ways? Or is there variation? If scores are aligned, it suggests strong inter-rater reliability. If scores are variable, it suggests weak inter-rater reliability.

In L&D, tools that encourage open-ended answers often reduce inter-rater reliability by a significant amount (Campbell et al., 2013). This is because variability in the answers - content, length or interpretation, for example - decreases consistency in the results. One such method to assess inter-rater reliability in psychology is triangulation, the process of obtaining a debrief of one researcher's interpretations from another (Belotto, 2018).

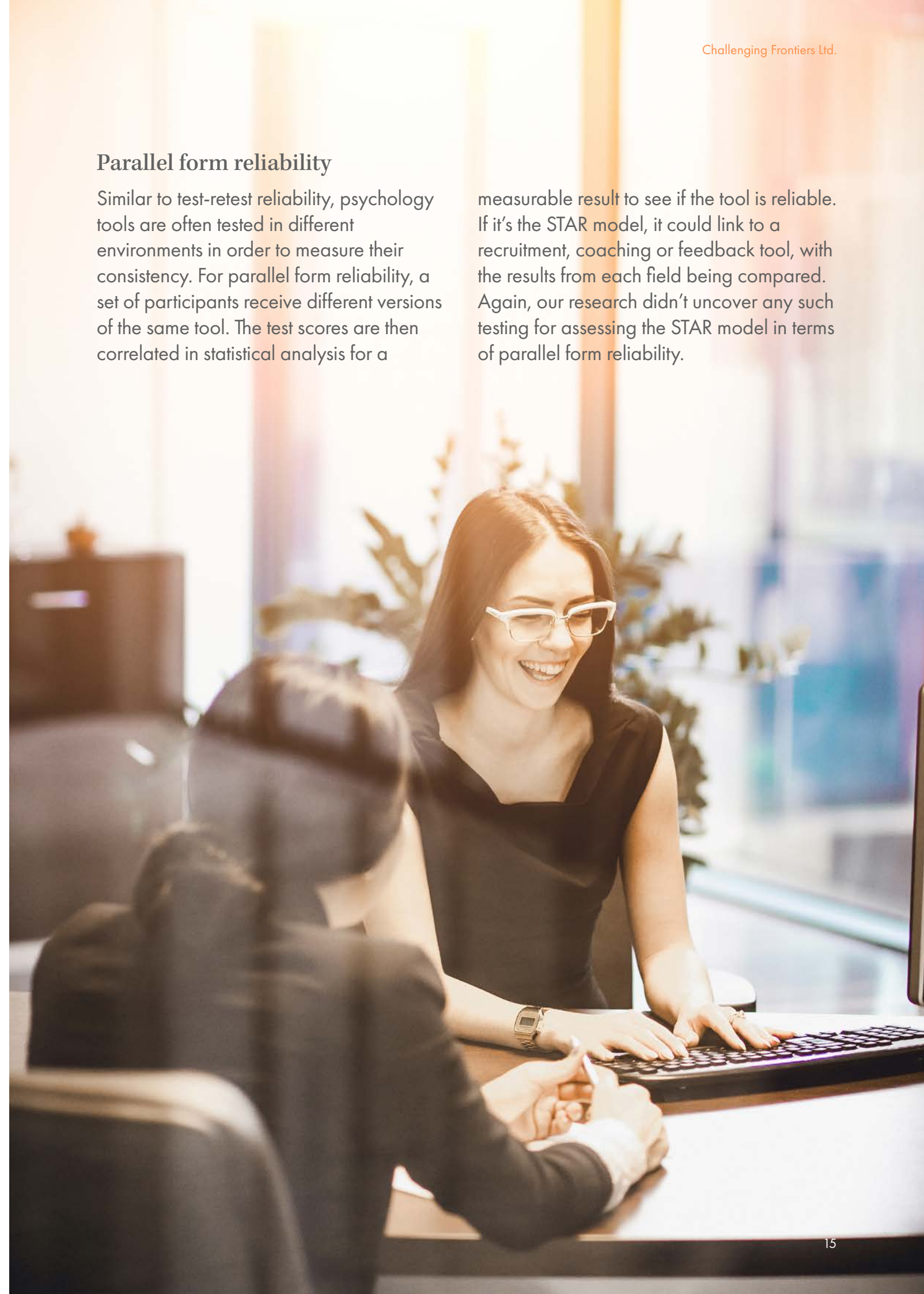
Standardised administration avoids personal interpretation or personal utilisation of a model. It ensures everyone hears the same messaging, reducing bias and prejudice. It is difficult to determine if there is inherent reliability in the STAR model because of the variability in answers, length of answers and a lack of standardised administration of the tool.

### Parallel form reliability

Similar to test-retest reliability, psychology tools are often tested in different environments in order to measure their consistency. For parallel form reliability, a set of participants receive different versions of the same tool. The test scores are then correlated in statistical analysis for a

measurable result to see if the tool is reliable. If it's the STAR model, it could link to a recruitment, coaching or feedback tool, with the results from each field being compared. Again, our research didn't uncover any such testing for assessing the STAR model in terms of parallel form reliability.

# Reliability





# Conclusion

In the context of the STAR model, we know there is a certain degree of usefulness for answering job interview questions, providing feedback and coaching. People's lived experience is a valid form of evidence and the thousands of people who use STAR model would likely argue it has served them well.

We can draw a few specific conclusions from this report and the research we undertook:

**You may achieve the desired results, but you can't ever be sure it's because of the STAR model.**

It is clear that the STAR model is a popular and widely used tool, but there is a lack of evidence and research for the building and testing of the model.

**1. There is doubt about the consistency and reliability of the STAR model, including in relation to administration.**

A scientific approach to building and testing tools in psychology requires reliability and validity testing. Only face validity seems to exist in relation to the STAR model. There is not any rigorous guidance or structure with regards to administering the tool also has no rigorous guidance or structure. This has decreased the inter-rater reliability.

**2. There is face validity, but it is only anecdotal.**

The discussion of reliability and validity leads to the conclusion that the STAR model may have a certain degree of face validity, because so many professionals and practitioners have used and recommended the tool.

**To allow for stronger utilisation of the STAR model, we recommend the following points:**

**1. Set out the context.**

Ask yourself: what would I, or the organisation, be using the STAR model for? As this report demonstrates, the STAR model has its place in recruitment, coaching and feedback. Contextual grounding is important to set the tone. For example: "Who are the audience?" or "What are the main performance needs of our line managers?"

**2. Clear definitions**

From there, you and your organisation will need to set out clear definitions for the STAR acronym. As discussed, there can be variations in the letters and definitions, such as Results or Review for "R". Setting out clear definitions will ensure a standardised and consistent approach - ie, a reliable approach. Using a standardised tool will therefore, also lead to valid answers.

**3. Avoid risk of variability.**

Try to avoid variability in the acronym. For example, in a feedback situation, use a standardised work document with helpful guidance and the STAR framework to ensure consistent use.

**4. Set a measure of success**

What outcomes do you expect to achieve from using the STAR model? If it's improved recruitment outcomes, how will you know STAR was a success factor in achieving that?

If it's for performance management purposes, how will you know STAR was a success factor? If it's for coaching, how will you know STAR helped produce a great set of actions?



*I hope you have found this report and all our findings useful.*

## With thanks

Vicki Xian for researching and co-writing the paper.

Rob Richardson for design and Insights Media for the production and publishing of the paper.

## Reference List

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42(3), 294-320. doi: 10.1177/0049124113500475

Belotto, M. J. (2018). Data analysis methods for qualitative research: Managing the challenges of coding, interrater reliability, and thematic analysis. *Qualitative Report*, 23(11).

Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vols. 1-4). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781506326139

Chang, H. J. B. (2020). *Social-Emotional Development and School Readiness: Assessing the Predictive Validity of a Kindergarten Screening Module* (Doctoral dissertation, Widener University).

Hobbs, M., (2016). What is test-retest reliability and why is it important? Retrieved from: <https://www.cambridgecognition.com/blog/entry/what-is-test-retest-reliability-and-why-is-it-important>

Boyer, S. (2009). Self-directed learning: Measures and models for salesperson training and development.

Indeed, (2020). How To Use the STAR Interview Response Technique. Retrieved from: <https://www.indeed.com/career-advice/interviewing/how-to-use-the-star-interview-response-technique>.

Pasaribu., MS., (2019). Giving Feedback with STAR. Retrieved from: <https://medium.com/@marthinpsrb/giving-feedback-with-star-148959e59bec>

DDI (no date). What's the STAR Method? Retrieved from: <https://www.ddiworld.com/solutions/behavioral-interviewing/star-method>.

Cook, S. (2009). Coaching for high performance: How to develop exceptional results through coaching. IT Governance Ltd

Doyle, A., (2021)., How to Use the STAR Interview Response Method. Retrieved from: <https://www.thebalancecareers.com/what-is-the-star-interview-response-technique-2061629>

UNM (2010). The Star Feedback Model. Retrieved from: <https://hr.unm.edu/docs/eod/the-star-feedback-model.pdf>

Rotolo, C. T., Church, A. H., Adler, S., Smither, J. W., Colquitt, A. L., Shull, A. C., ... & Foster, G. (2018). Putting an end to bad talent management: A call to action for the field of industrial and organizational psychology. *Industrial and Organizational Psychology*, 11(2), 176-219.

Kraiger, K., & Ford, J. K. (2020). The Science of Workplace Instruction: Learning and Development Applied to Work. *Annual Review of Organizational Psychology and Organizational Behavior*, 8.

Parker, S. K. (2017). Work design growth model: How work characteristics promote learning and development.

Idrisov, G., Mau, V., & Bozhechkova, A. (2017). Searching for a new growth model. *Voprosy ekonomiki*, (12), 5-23.

Mooi, E.A., Sarstedt, M. and Mooi-Reci, I. (2018), *Market Research. The Process, Data, and Methods using STATA*, Springer, Berlin.

Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economic Series*, 17(4), 59-82.

Cook, D. A., & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine*, 119, 166.e7-166.e16.

Forza, C. (2002). Survey Research in Operations Management: A Process-based Perspective. *International Journal of Operations and Production Management*, 22 (2), 152-194.

Rigdon, E.E., Becker, J.-M., Rai, A., Ringle, C.M., Diamantopoulos, A., Karahanna, E., Straub, D. and Dijkstra, T.K. (2014), "Conflating antecedents and formative indicators: a comment on Aguirre-Urreta and Marakas", *Information Systems Research*, Vol. 25 No. 4, pp. 780-784.

Thompson, B. (2003). Understanding Reliability and Coefficient Alpha, Really. In B. Thompson (Ed.). *Score Reliability: Contemporary Thinking on Reliability Issues*. Thousand Oaks, CA: SAGE.

Alavi, M., Archibald, M., McMaster, R., Lopez, V., & Cleary, M. (2018). Aligning theory and methodology in mixed methods research: Before design theoretical placement. *International Journal of Social Research Methodology*, 21(5), 527-540.

Tashakkori, A., & Teddlie, C. (2010). *SAGE handbook of mixed methods in social & behavioral research*: SAGE







**Challenging Frontiers**

Independent insight, advice and research

Challenging Frontiers Ltd.

196A High Street

Epping

Essex

CM16 4AQ

Sukh (+44) 07961-700305

[hello@challengingfrontiers.com](mailto:hello@challengingfrontiers.com)

