

## Manual for the Startle & Surprise Inventories and Visual Analogue Scales

Chen, J.; Landman, H.M.; Stroosma, O.; van Paassen, M.M.; Mulder, Max

**DOI**

[10.4233/uuid:4aa39791-4d21-4427-b86a-628f52d17fbe](https://doi.org/10.4233/uuid:4aa39791-4d21-4427-b86a-628f52d17fbe)

**Publication date**

2025

**Document Version**

Final published version

**Citation (APA)**

Chen, J., Landman, H. M., Stroosma, O., van Paassen, M. M., & Mulder, M. (2025). *Manual for the Startle & Surprise Inventories and Visual Analogue Scales*. Delft University of Technology.  
<https://doi.org/10.4233/uuid:4aa39791-4d21-4427-b86a-628f52d17fbe>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# **Manual for the Startle & Surprise Inventories and Visual Analogue Scales**

Jiayu Chen  
Annemarie Landman  
Olaf Stroosma  
M. M. (René) van Paassen  
Max Mulder

Section Control & Simulation  
Faculty of Aerospace Engineering

June 19, 2025



## **Contents**

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Instruments overview</b>	<b>2</b>
<b>3</b>	<b>Administration guidelines</b>	<b>2</b>
<b>4</b>	<b>Psychometric properties</b>	<b>3</b>
4.1	Reliability . . . . .	3
4.2	Validity . . . . .	3
4.2.1	The Startle and Surprise Inventories . . . . .	3
4.2.2	The Visual Analogue Scales for Startle and Surprise . . . . .	3
<b>5</b>	<b>Contact and permissions</b>	<b>4</b>
	<b>References</b>	<b>5</b>
<b>A</b>	<b>The Startle Inventory (Startle-I)</b>	<b>6</b>
<b>B</b>	<b>The Surprise Inventory (Surprise-I)</b>	<b>7</b>
<b>C</b>	<b>The Visual Analogue Scales for Startle (Startle-VAS)</b>	<b>8</b>
<b>D</b>	<b>The Visual Analogue Scales for Surprise (Surprise-VAS)</b>	<b>9</b>

# 1 Introduction

Accurate and non-obtrusive measurement of startle and surprise is essential for advancing our understanding of human responses under stress, identifying causal factors, and assessing their distinct effects on operator performance. Insights gained from such research can inform the development of evidence-based safety protocols and targeted training interventions.

To support this aim, the Startle and Surprise Inventories (Startle-I; Surprise-I), and the Visual Analogue Scales for Startle and Surprise (Startle-VAS; Surprise-VAS) were developed and psychometrically validated as self-report measuring instruments for assessing startle and surprise in response to specific events, situations, or stimuli [1, 2]. Although initially developed within the context of human factors research in aviation, these instruments are designed to be broadly applicable across domains where unexpected events may impact human performance, including healthcare, maritime navigation, and military operations.

These instruments are intended for use in experimental, operational, and training environments, particularly those where acute stress and unexpected events are likely to occur, such as in-flight system failures. This manual provides guidance for human factors researchers and applied psychologists, on the standardized and scientifically rigorous use of the instruments. It is structured to include an overview of the instruments, administration guidelines, and a summary of their psychometric properties to support accurate application and interpretation in research and operational contexts.

## 2 Instruments overview

The Startle-I consists of six items (Appendix A), and the Surprise-I comprises five items (Appendix B). Simple and accessible language was used on the items to allow the inventories to be used in research with non-native English speakers. Response options for all items are presented on a 5-point Likert scale, which captures varying levels of agreement with each statement (1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree). For the Startle-I and Surprise-I, the score of each inventory is defined to be the average of all items' ratings, ranging from 1 to 5.

Additionally, item 2 ("I predicted it beforehand.") and item 4 ("It was consistent with my expectation.") in the Surprise-I are reverse-coded. For these items, a response of 5 should be recoded as 1, 4 as 2, 3 remains unchanged, 2 as 4, and 1 as 5.

The single-item Startle-VAS (Appendix C) and Surprise-VAS (Appendix D) each consist of a 100 mm horizontal line with tick marks at 10 mm intervals. The left endpoint is labelled "0" and the right endpoint is labelled "10". The left endpoints are additionally labelled with "not startled at all" and "not surprised at all", respectively. The right endpoints are labelled with "extremely startled" and "extremely surprised", respectively. For the Startle-VAS and Surprise-VAS, users are required to place a cross/mark on the line as the answer to the question and the resulting score is the distance of the centre of the cross/mark to the left endpoint in centimetres.

## 3 Administration guidelines

Consistent administration is essential to ensure the validity and comparability of data across studies. In the introductory text for both the multi-item inventories and visual analogue scales (VASs), clearly specify which particular stimulus or event the pronoun 'it' refers to in each item. To ensure consistency across participants and conditions, standardized instructions should be

provided before application. Below is a suggested script:

*“You are about to complete a questionnaire about your immediate reaction. Please respond honestly based on how you actually felt at the time of the specific event, not how you think you were supposed to feel or how you would normally respond. There are no right or wrong answers.”*

The inventories and VASs should be completed as soon as possible after the event or stimulus of interest. Considering that startle is a rapid and transient response [5, 7] and surprise could involve cognitive appraisal [6, 10, 12], a delay in the measurement may affect the accuracy of self-reports, due to factors such as memory decay or reinterpretation of the stimulus.

## 4 Psychometric properties

### 4.1 Reliability

The internal consistencies of the Startle-I and Surprise-I have been evaluated across multiple samples and varied contexts. In a sample of 729 observations, 81 participants were exposed to nine video stimuli. Cronbach’s  $\alpha$  [3] indicated acceptable to excellent internal consistency, ranging from  $\alpha = 0.714$  to  $\alpha = 0.929$  for Startle-I, and  $\alpha = 0.843$  to  $\alpha = 0.955$  for Surprise-I [1].

In a separate study comprising 208 observations, 26 airline pilots experienced eight varied startling and surprising scenarios in an ecologically-valid aviation setting. McDonald’s  $\omega$  [8] indicated similarly high internal consistency, with values ranging from  $\omega = 0.88$  to  $\omega = 0.96$  for the Startle-I, and  $\omega = 0.77$  to  $\omega = 0.96$  for the Surprise-I [2].

### 4.2 Validity

#### 4.2.1 The Startle and Surprise Inventories

To evaluate the psychometric properties of the Startle-I and Surprise-I, multiple validity assessments were conducted. Content validity was established through expert evaluation by seven specialists in the fields of Cognitive Science and Psychology. These experts assessed the relevance of an initial set of items developed for measuring startle and surprise, derived from fundamental and applied literature [1]. An item was retained if at least 50% of the experts rated that item to be relevant for its construct [4].

Construct validity was initially examined using 729 observations from 81 participants, each of whom rated the retained items nine times following exposure to nine video clips [1]. Multilevel exploratory factor analysis with oblique, direct oblimin rotation was employed [11, 13].

Further evidence for construct validity in an ecologically-valid setting was obtained from 208 observations involving 26 airline pilots, each exposed to eight simulated in-flight scenarios designed to elicit varied levels of startling and surprising responses [2]. Multilevel confirmatory factor analysis [9] was conducted to confirm the factor structures of the Startle-I and Surprise-I. These assessments collectively confirmed that the inventories reliably measure self-report startle and surprise responses in operational contexts.

#### 4.2.2 The Visual Analogue Scales for Startle and Surprise

Concurrent validity of the Startle-VAS and Surprise-VAS was assessed with Spearman correlations between the Startle-VAS and Startle-I, as well as the Surprise-VAS and Surprise-I, using 729 observations from 81 participants [1]. The ratings of Startle-VAS showed strong correlations

with the Startle-I scores,  $\rho = 0.778$  to  $\rho = 0.877$ . The ratings of Surprise-VAS highly correlated with the Surprise-I scores,  $\rho = 0.681$  to  $\rho = 0.903$ . All correlations were statistically significant, supporting the concurrent validity of the Visual Analogue Scales for Startle and Surprise.

## 5 Contact and permissions

This instruction manual is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#). When using or adapting these instruments, please cite:

*doi: [10.4233/uuid:4aa39791-4d21-4427-b86a-628f52d17fbe](https://doi.org/10.4233/uuid:4aa39791-4d21-4427-b86a-628f52d17fbe)*

For questions regarding the administration, scoring, or interpretation of the Startle and Surprise Inventories and Visual Analogue Scales for Startle and Surprise, please contact the research team: [startle-surprise-inventories@tudelft.nl](mailto:startle-surprise-inventories@tudelft.nl).

## References

- [1] Chen, J., Landman, A., Derumigny, A., Stroosma, O., van Paassen, M. M., and Mulder, M. (2025a). Development and Validation of the Startle and Surprise Inventories and Visual Analogue Scales [Manuscript submitted for publication].
- [2] Chen, J., Landman, A., Derumigny, A., Stroosma, O., van Paassen, M. M., and Mulder, M. (2025b). Preliminary Multilevel Confirmatory Factor Analysis of the Startle & Surprise Inventories in an Operational Aviation Context [Manuscript submitted for publication].
- [3] Cronbach, L. J. (1951). Coefficient Alpha and the Internal Structure of Tests. *psychometrika*, 16(3):297–334.
- [4] DeVellis, R. F. (2012). *Scale Development: Theory and Applications*. SAGE Publications, third edition.
- [5] Koch, M. (1999). The Neurobiology of Startle. *Progress in Neurobiology*, 59(2):107–128.
- [6] Landman, A., Groen, E. L., Van Paassen, M. M., Bronkhorst, A. W., and Mulder, M. (2017). Dealing With Unexpected Events on the Flight Deck: A Conceptual Model of Startle and Surprise. *Human Factors*, 59(8):1161–1172.
- [7] Martin, W. L., Murray, P. S., Bates, P. R., and Lee, P. S. Y. (2015). Fear-Potentiated Startle: A Review from an Aviation Perspective. *The International Journal of Aviation Psychology*, 25(2):97–107.
- [8] McDonald, R. P. (2013). *Test Theory: A Unified Treatment*. Psychology Press.
- [9] Mehta, P. D. and Neale, M. C. (2005). People Are Variables Too: Multilevel Structural Equations Modeling. *Psychological Methods*, 10(3):259.
- [10] Noordewier, M. K., Topolinski, S., and Van Dijk, E. (2016). The Temporal Dynamics of Surprise. *Social and Personality Psychology Compass*, 10(3):136–149.
- [11] Reise, S. P., Ventura, J., Nuechterlein, K. H., and Kim, K. H. (2005). An Illustration of Multilevel Factor Analysis. *Journal of Personality Assessment*, 84(2):126–136.
- [12] Reisenzein, R., Horstmann, G., and Schützwohl, A. (2019). The Cognitive-Evolutionary Model of Surprise: A Review of the Evidence. *Topics in Cognitive Science*, 11(1):50–74.
- [13] Tabachnick, B. G., Fidell, L. S., and Ullman, J. B. (2013). *Using Multivariate Statistics*. Pearson, seventh edition.

## A The Startle Inventory (Startle-I)

The following statements refer to [the stimulus]<sup>1</sup>. Please read each statement and circle the number that best represents your agreement with the statement.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1. It startled me.	1	2	3	4	5
2. It immediately made me feel scared or angry.	1	2	3	4	5
3. It made me physically flinch.	1	2	3	4	5
4. It caused my heart to suddenly beat harder or faster.	1	2	3	4	5
5. It shocked me.	1	2	3	4	5
6. It immediately caused stress or frustration to me.	1	2	3	4	5

---

<sup>1</sup>Specific stimulus description should be inserted here.



## B The Surprise Inventory (Surprise-I)

The following statements refer to [the stimulus]<sup>2</sup>. Please read each statement and circle the number that best represents your agreement with the statement.

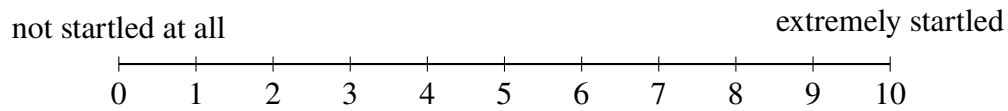
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1. It surprised me.	1	2	3	4	5
2. I predicted it beforehand.	1	2	3	4	5
3. I did not see it coming.	1	2	3	4	5
4. It was consistent with my expectation.	1	2	3	4	5
5. It was unexpected.	1	2	3	4	5

---

<sup>2</sup>Specific stimulus description should be inserted here.

## C The Visual Analogue Scales for Startle (Startle-VAS)

Please indicate, by placing a cross/mark on the line below<sup>3</sup>, how **startled** you were by [the stimulus]<sup>4</sup>.



---

<sup>3</sup>When printed on A4 paper at 100% scale, the Visual Analogue Scale measures precisely 100 mm, as intended.

<sup>4</sup>Specific stimulus description should be inserted here.

## D The Visual Analogue Scales for Surprise (Surprise-VAS)

Please indicate, by placing a cross/mark on the line below<sup>5</sup>, how **surprised** you were by [the stimulus]<sup>6</sup>.



---

<sup>5</sup>When printed on A4 paper at 100% scale, the Visual Analogue Scale measures precisely 100 mm, as intended.

<sup>6</sup>Specific stimulus description should be inserted here.