

AVERAGE ERROR

Authored by
mohammad looti

November 11, 2025

RECOMMENDED CITATION

mohammad looti (2025). *AVERAGE ERROR*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=68852>

AVERAGE ERROR

Primary Disciplinary Field(s): Statistics, Experimental Psychology, Metrology, Data Science

1. Core Definition

The **Average Error** is a fundamental statistical measure utilized across multiple quantitative disciplines, particularly experimental psychology and metrology, to quantify the typical magnitude of inaccuracy within a set of observations or measurements. It refers specifically to the degree by which a series of recorded data points deviate from a predefined criterion. This criterion can be either an **absolute standard**, such as a known physical constant, standard weight, or accepted length, or a **relative criterion**, often represented by the mean of the observations themselves within a specific experimental factor or condition.

In essence, Average Error provides a concise summary statistic detailing the overall level of precision or deviation present in the measurement process. Unlike measures that focus on the direction of error (bias), Average Error typically concentrates on the magnitude of the disparity between the observed value and the true or expected value. A low average error suggests high precision and reliability in the measurement system or observer, indicating that, on average, the observations cluster closely around the target or criterion value. Conversely, a high average error signals significant variability or inaccuracy across the collected data points, suggesting issues with the measurement instrument or methodology.

While often used synonymously with related concepts like Mean Absolute Error (MAE) or Root Mean Square Error (RMSE) in modern statistical modeling, the historical definition of Average Error, especially in early psychological and physical measurement contexts, implies calculation based on the absolute differences. This methodological choice is designed to avoid the cancellation of positive errors (overestimations) and negative errors (underestimations) that would occur if one simply calculated the mean of the raw differences, which yields only the systematic error or bias. This measure is therefore critical for researchers assessing the fundamental quality and validity of their experimental data, especially when calibrating instruments or evaluating the consistency of human performance in tasks requiring objective accuracy.

2. Etymology and Historical Development

The formal concept of quantifying observational error is deeply rooted in the history of statistics and the physical sciences, tracing its origins back to the needs of 18th-century astronomers and surveyors performing geodetic measurements. Early practitioners, faced with inherent imprecision in tools and observation, sought systematic ways to derive the most probable true value from a series of discrepant readings. This foundational concern led to the pioneering work of

mathematicians such as Carl Friedrich Gauss and Adrien-Marie Legendre, whose development of the Method of Least Squares provided a standardized mathematical framework for minimizing error, thus laying the groundwork for formal error analysis.

The specific term **Average Error** gained distinct prominence in the burgeoning field of experimental psychology during the late 19th and early 20th centuries. With the establishment of the first psychological laboratories by figures like Wilhelm Wundt, there arose an urgent need for standardized metrics to quantify human sensory abilities, reaction times, and the consistency of subjective judgment. In psychophysics, for example, researchers required a metric that could easily and intuitively summarize how far, on average, a participant's judgment deviated from the objective physical stimulus. Average Error served this purpose by offering a practical, easily calculated metric of typical deviation, thereby becoming a standard reporting tool.

It is important to note that the historical calculation of Average Error was not always uniform, which has contributed to terminological ambiguity over time. In classical texts, it was often strictly defined as the arithmetic mean of the absolute values of the errors, corresponding precisely to the modern Mean Absolute Error (MAE). However, sometimes the term was used more generally to refer to any measure of central tendency applied to the error distribution. Regardless of slight methodological variations, the core objective of the calculation--to provide a single, representative number that encapsulates the overall spread or inaccuracy inherent in the observational data--has remained constant throughout its history. While subsequent advances in statistical modeling have led to the prevalence of more theoretically complex and robust measures, the concept of Average Error reflects a crucial, foundational analytical approach to data quality.

3. Calculation Methodology and Criteria Differentiation

The operational definition of **Average Error** centers on its straightforward calculation, derived by dividing the total aggregated error score by the total number of cases or observations (N). As typically defined in statistical practice relevant to measurement accuracy, the calculation necessitates three key stages. First, the error for each individual case (e_i) is determined by calculating the difference between the observed value (x_i) and the established criterion value (C). Second, the absolute value of each of these errors is summed to yield the Total Error Score ($\sum |x_i - C|$). Third, this total score is divided by the amount of cases (N) to yield the Average Error.

The use of the absolute value function in the summation ($\sum |e_i|$) is a critical methodological feature distinguishing Average Error from measures of bias. If signed errors were summed, any overestimation (positive error) would cancel out an equivalent underestimation (negative error), potentially yielding a mean of zero even if all individual observations were significantly inaccurate. By utilizing absolute values, the calculation effectively captures the magnitude of inaccuracy

regardless of its direction, providing a true measure of typical measurement deviation. This ensures that the resultant statistic accurately reflects the precision, or lack thereof, within the entire set of observations.

Furthermore, the nature of the criterion ($\bar{x}$) chosen dictates the type of error being quantified. When an **Absolute Criterion** is used--such as a certified standard of weight or length--the resulting Average Error reflects the absolute accuracy of the measurement system relative to an objective truth. This is essential in metrology for instrument calibration. Conversely, when a **Relative Criterion** is employed--typically the mean of the collected observations ($\bar{x}$) within a specific experimental condition--the Average Error reflects the internal consistency or precision of the data points relative to one another. This latter application is common in fields where the true value is unknowable, and the researcher is interested in the variability or spread of the responses within a controlled group.

4. Relationship to Modern Error Metrics

While historically significant, the term **Average Error** has largely been supplanted by more precise and mathematically grounded metrics in contemporary statistics and machine learning, primarily the **Mean Absolute Error (MAE)** and the **Root Mean Square Error (RMSE)**. Understanding the distinctions between these metrics is crucial for modern data analysis. Average Error, when calculated using absolute differences, is mathematically identical to MAE. MAE measures the average length of the error vector, offering a metric whose value is in the same units as the dependent variable, providing highly intuitive interpretation.

However, MAE and, by extension, Average Error, contrast sharply with RMSE. RMSE is calculated by squaring the errors, averaging the squared values, and then taking the square root. The crucial difference introduced by squaring the errors is the disproportionate weighting given to large deviations. RMSE inherently penalizes large errors much more severely than MAE does. This characteristic makes RMSE the preferred loss function in many fields, particularly engineering and physical sciences, where mitigating extreme errors is paramount. If a few very large errors exist in the dataset, the RMSE score will be dramatically higher than the MAE score, signaling the severity of these outliers.

Furthermore, both MAE (Average Error) and RMSE differ fundamentally from the **Standard Deviation**. Standard Deviation measures the spread of data points around the mean of the sample itself and is optimized for inferential statistical tasks due to its close ties to the properties of the normal distribution. Average Error measures the typical distance from a chosen criterion (which may or may not be the sample mean), and while it offers robustness against extreme outliers compared to RMSE, its reliance on the non-differentiable absolute function limits its utility in advanced optimization techniques used in complex modeling, such as those relying on continuous

gradient descent algorithms.

5. Applications in Applied Sciences

The use of **Average Error** is pervasive in measurement-intensive disciplines. In metrology, it forms the basis for certifying equipment accuracy. For example, a manufacturer of laboratory scales will calibrate the device against a series of certified reference weights. The Average Error calculated from these trials quantifies the typical deviation of the scale's readings from the known standards, providing a clear measure of the instrument's operational accuracy and its necessary tolerance levels. This metric is essential for establishing uncertainty budgets in high-precision manufacturing and research.

In experimental psychology, Average Error remains a core metric for evaluating human performance consistency. Consider a study on perceptual-motor tasks, such as tracking a moving target or adjusting a dial to match a perceived magnitude. The Average Error calculated from the participant's attempts provides an unambiguous quantitative score for their accuracy. Changes in Average Error over training sessions are critical indicators of learning and skill acquisition, as a decrease in this value demonstrates greater precision and reduced variability in performance. It allows researchers to compare the effectiveness of different training protocols or the impact of external factors (e.g., fatigue, distraction) on observational precision.

Moreover, in geographical and surveying applications, Average Error can be used to assess the precision of measurements taken over large distances. When multiple surveyors independently measure the same segment, the Average Error calculated relative to the combined mean of their results provides immediate insight into the homogeneity of their measurement techniques and the reliability of their instruments. In these applied contexts, the metric's simplicity and direct interpretation in the original units of measure are substantial advantages for practical reporting and decision-making.

6. Limitations and Ambiguity

A significant limitation of **Average Error** stems from its reduced mathematical tractability when compared to squared-error metrics. Because the absolute value function is not differentiable at zero, the application of Average Error (or MAE) can complicate its integration into sophisticated statistical and machine learning models that depend on continuous gradient calculations for parameter optimization. This mathematical constraint often necessitates the use of RMSE or variance-based measures when building complex predictive frameworks.

A second major limitation involves the inherent ambiguity of the historical term itself. Before the standardization of statistical reporting, "Average Error" was sometimes used loosely, leading to confusion regarding whether the reported figure represented the mean of the absolute differences

(precision/MAE) or the mean of the signed differences (systematic bias). This lack of explicit definition can severely impede the replicability and comparability of older research findings. Modern statistical reporting standards mitigate this issue by requiring explicit designation of the metric used, ensuring clarity for the scientific community.

Finally, compared to RMSE, Average Error minimizes the sum of absolute errors rather than the sum of squared errors. While this makes it more robust to outliers (as it doesn't heavily penalize them), it also means that if the research goal is specifically to eliminate or reduce the largest potential deviations--a necessity in fields like safety engineering or quality assurance--Average Error may be insufficient. In such high-stakes scenarios, the stronger penalty imposed by squaring large errors (as utilized by RMSE) is generally preferred to ensure that model improvements prioritize the reduction of catastrophic errors.

7. Further Reading

[Mean absolute error \(Wikipedia\)](#)

[Experimental Psychology \(Wikipedia\)](#)

[Metrology \(Wikipedia\)](#)

[Precision and accuracy \(Wikipedia\)](#)

[Fisher, R. A. \(1922\). On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 222, 309-368.](#)