Ensuring the validity of scientific research

Sherman Xie. Zhejiang University of Technology (República Popular China)

Recibido 26/06/2026 • Aceptado 17/10/2026

ORCID: <https://orcid.org/0000-0003-2544-8277>

Abstract

If «scientific» implies precision, reliability, validity, and universality, much contemporary research across disciplines fails to meet these standards. Circular measurement poses a significant threat to the natural sciences, social sciences, and humanities by undermining research validity and reproducibility. Some published studies are even retracted, not necessarily because of academic misconduct, but potentially due to insufficient measurement literacy.

This paper argues that circular measurement, while seemingly arising from conceptual ambiguity, fundamentally occurs when measurement validity depends on the very object, concept, or theory being measured. This dependency generates circular reasoning rooted in the paradox of self-reference, traditionally avoided in classical mathematics. Drawing on mathematicians' and philosophers' responses to the third foundational crisis in mathematics, the paper proposes clarifying the intensions and extensions of concepts as a possible strategy for overcoming self-referential paradoxes and mitigating the methodological problems associated with circular measurement.

Keywords: validity, circular measurement, coordination, logic, self-reference, paradox.

Resumen

Garantizar la validez de la investigación científica

Si lo «científico» implica precisión, fiabilidad, validez y universalidad, buena parte de la investigación contemporánea en diversas disciplinas no satisface estos criterios. La medición circular representa una amenaza significativa para las ciencias naturales, las ciencias sociales y las humanidades, al comprometer la validez y reproducibilidad de las investigaciones. Algunos estudios publicados llegan incluso a ser retractados, no necesariamente por malas prácticas académicas, sino posiblemente por una insuficiente formación en materia de medición.

Este artículo sostiene que la medición circular, aunque aparentemente deriva de la ambigüedad conceptual, surge fundamentalmente cuando la validez de una medición depende del propio objeto, concepto o teoría que pretende medir. Esta dependencia genera un razonamiento circular vinculado a la paradoja de la autorreferencia, tradicionalmente evitada en las matemáticas clásicas. A partir de las respuestas de matemáticos y filósofos a la tercera crisis de los fundamentos de las matemáticas, el artículo propone clarificar la intensión y extensión de los conceptos como posible estrategia para superar las paradojas autorreferenciales y mitigar los problemas metodológicos asociados a la medición circular.

Palabras clave: validez, medición circular, coordinación, lógica, autorreferencia, paradoja.

ISSN-e: 1885-5679

ISSN-e: 1885-5679

Ensuring the validity of scientific research

Sherman Xie. Zhejiang University of Technology (República Popular China)
Recibido 26/06/2026 • Aceptado 17/10/2026
ORCID: <https://orcid.org/0000-0003-2544-8277>

§ 1. Introduction[1]

The progress of human intelligence is primarily reflected in the academic research activities of three major cognitive systems: natural sciences, social sciences, and humanities. Especially over the past four hundred years, since science separated from religion and became an independent discipline, academic research —represented by the natural sciences— has significantly accelerated the advancement of human material civilization. As a result, religion has gradually declined, while science has flourished.

However, both «science» and «religion» are merely concepts that have emerged in the course of human civilization, shaped by various conditions. The debate between science and religion is essentially a contest over their respective understandings of causal logic, and both sides have inherent limitations in their grasp of it.

In contemporary academic contexts, the term «science» seems to imply precision, truthfulness, reliability, validity, universality, etc.; this is merely a superstitious consensus held by ordinary people. In fact, most so-called «scientific research» across various disciplines today is, in fact, not truly scientific.

For instance, the problem of circular measurement poses a common threat to natural sciences, social sciences, and humanities, as it undermines the validity and accuracy of research. On the surface, circular measurement stems from the ambiguity of the concepts involved. However, at its core, the issue arises because the validity of

[1] Statements/declarations. *Ethical approval*: this research does not involve human participants and(/or) animals and does not analyze or generate any datasets, because it proceeds with a theoretical approach. The IRB of the author's institution declared the study exempt from ethical review. *Funding*: no funds, grants, or other support were received. *Author's contribution*: this manuscript is the author's original work. *Competing interests*: there are no (/potential) conflicts of interest. *Availability of data and materials*: this research does not analyze or generate any datasets. *Use of GenAI in writing*: this article used ChatGPT to correct spelling and grammar errors.

ISSN-e: 1885-5679

measurement depends on the object it attempts to measure —such as certain concepts or theories— rather than being independent of it. This inevitably leads to circular reasoning in the results.

In other words, the problem of circular measurement originates from the self-referential paradox forbidden by classical mathematics. The validity of classical mathematics is built upon the avoidance of self-reference. Drawing on the approaches of mathematical philosophers to address the current Third Mathematical Crisis, namely, by clarifying the intension and extension of concepts to circumvent self-referential paradoxes. This method may also help mitigate the threat of circular measurement.

Validity is considered a key methodological concept for evaluating the quality of measurement and assessment tools across multiple scientific disciplines in natural and social sciences (Barlas and Carpenter 1990; Garaway 1997; Hood 2009; Markus and Borsboom 2013; Alexandrova and Haybron 2016). In general, this term refers to the extent to which an assessment achieves its intended objectives. For example, a «valid» intelligence test provides information about intellectual ability, while a «valid» tuberculin test indicates the presence of tuberculosis infection. Although the technical term «validity» was first explicitly introduced in the context of standardizing intelligence tests in the 1920s, it has gradually evolved into a methodological concept applied beyond psychology (Monteiro *et al*., 2022) and education (James-Brabham *et al*., 2023), including performance testing (Traini *et al*., 2024), medical diagnostics (Yao *et al*. 2020), market research (Punniyamoorthy and Prasanna Mohan Raj, 2007), and more. Despite these developments leading to numerous validity concepts and classifications, intense debates remain regarding the appropriate meanings and applications of validity under different assessment objectives. (Markus, 1998; Strauss and Smith, 2009; Borsboom and Markus, 2013; Zhao 2023)

Behind the debates over the validity of measurement tools often lies a fundamental epistemological issue: the threat of circular measurement. This threat arises when claims are made that two measurement procedures are measuring the same characteristic, or that a particular procedure is indeed measuring the characteristic of interest, without sufficient or appropriate justification. This issue has been extensively discussed in recent measurement epistemology, referred to as the naming

ISSN-e: 1885-5679

measurement problem or the coordination problem. (Luchetti, 2020) The traditional concept of coordination in epistemology emerged in the philosophical debates about the relationship between precise science and reality, sparked by fundamental developments in physics at the turn of the 20th century. However, philosophers and methodologists have recently begun to recognize the importance of this concept in fields outside the physical sciences, such as evidence-based medicine, physical anthropology, perceptual physiology, neuroscience, and psychometrics, within the context of emerging quantitative concepts (Mari, 2003, 2005; Goertz and Mahoney, 2012; Ghilardi and Keller, 2012; Micheli and Mari, 2014).

In summary, measurement and evaluation serve as both the starting point and endpoint for modern scientific research. Ensuring the validity of measurements without independent validation methods presents a common challenge for natural sciences, social sciences, and even the humanities. (Chang, 2004; Tal, 2013) This issue is significant in epistemological and metrological discussions because it relates to how we understand and verify the reliability of scientific measurements. If the measurement process is threatened by circularity, the conclusions drawn may lack external validity, meaning they might not be applicable in contexts beyond the measurement tool itself. This paper aims to explore the origins and nature of the circular measurement problem and propose potential solutions.

§ 2. The origin and nature of the circular measurement problem

The threat of circular measurement directly stems from the ambiguity of concepts, and the essence is the paradox of self-reference. On these two points, a brief description is as follows:

2. 1. The circular measurement problem stems from the ambiguity of concepts

The definition of concepts is the premise of all scientific research and thus is fundamental. Generally speaking, the traditional method used to define concepts is definition by attributes, which involves defining a concept based on its core connotative attributes. (Wang, 2017) This method works well in most natural science

ISSN-e: 1885-5679

fields, but in the fields of social sciences and humanities, concepts defined in this way can leave people feeling confused. As a result, a significant proportion of concept definitions in social sciences and humanities literature are not very successful. The reason is that many concepts in the natural sciences are, in fact, about establishing a standard attribute and using this standard attribute to change or filter nature. Therefore, it is possible to use a very simple attribute to precisely define all its boundaries. For example, when we define oxygen, there is almost no pure oxygen in nature; it is mostly a mixture of gases. However, we can define oxygen as a standard —two identical atoms with an atomic number of 8 combined to form a molecule, which under certain conditions forms a gas, and select substances that meet this definition in nature. Thus, the attribute definitions of concepts in the natural sciences can be very simple and can be entirely conducted within the discipline.

However, in the fields of social sciences and humanities, including some fields that cannot be arbitrarily changed and filtered by humans (such as geology, meteorology, etc.), the attribute definition method of concepts will encounter great difficulties. (Wang, 2017) Common reasons include the following points:

1) Different people, for various reasons, are likely to give different names or expressions to the same cognitive object. This is not just a matter of different names for a concept, but it may also make the extension and connotation of the concept different, with overlaps and differences. In the natural sciences, it is relatively easy to unify concept names, symbols, etc., through the efforts of the professional community. In the fields of social sciences and humanities, this work is relatively more difficult and is deeply influenced by different ethnicities, cultures, schools, parties, enterprises, or interest groups, and even different political or ideological influences. There are also those who, to reduce the plagiarism rate of their papers, to show their characteristics, or simply to be different from others, artificially give different names to the same objects, making it seem like their own innovation, their unique viewpoints, and opinions. To avoid being completely the same as others, they artificially create some subtle differences that can be magnified.

2) Concept definitions often need to take into account the names people use for them in everyday life, which makes their attributes easily become very complex, making it difficult to define them accurately.

ISSN-e: 1885-5679

3) Concept definitions are likely to be developing and changing rather than fixed, which makes simple concept definitions easily deviate more and more as times change.

4) Concept definitions often involve a very wide range, and different people from different perspectives and disciplines will have different cognitions. Therefore, traditional research focused on concepts within a discipline is difficult to adapt to.

5) Different disciplines have their particularities, and in addition to the common reasons mentioned above, there must be other reasons.

2. 2. The circular measurement problem is essentially a self-referential paradox

Modern mathematics is the core and foundation of natural science, and set theory is the core and foundation of modern mathematics. However, the current third mathematical crisis, caused by the flaws in set theory, has not been truly resolved. The first mathematical crisis involved irrational numbers and was essentially a conflict between countable and uncountable. The second mathematical crisis involved calculus and was essentially a conflict between finite and infinite. The current third mathematical crisis, involving set theory, is essentially a conflict between the whole and the part. Modern set theories, such as the ZF system (Felgner, 2006) and the NBG system (Bodea, 2013), avoid this contradiction by adding principles, which to some extent compensate for the shortcomings of Cantor's set theory (Van Dalen, 1972; Vallin, 2013), but are like computer operating systems that have been patched and are merely usable. In fact, they have not been able to resolve the conflict between the whole and the part.

Naive set theory is one of the foundations of modern set theory, but it is non-formalized and prone to logical contradictions. To address the many paradoxes that arise in naive set theory, Bertrand Russell accepted Henri Poincaré's idea of impredicativity and proposed the vicious circle principle (VCP). (Whitehead and Russell, 2019) This principle states that in any set, there cannot be an individual that depends on the definition of the whole set, nor can there be a whole that contains such an individual. By prohibiting such circular definitions, i.e., avoiding self-reference, the generation of paradoxes can be avoided.

ISSN-e: 1885-5679

Self-reference involves a description of something about itself and encompasses three dimensions or levels: self-negation, vicious circles, and self-referentiality. Here is a brief description of each:

The first dimension or level is the self-negation. The ambiguity of self-reference is primarily manifested in self-referential paradoxes, where logic can derive a judgment that is contrary to itself, that is, the self-negation of logic. Generally, what we need is a consistent logic that can describe each other. A system composed of consistent logic should not exhibit contradictions. However, Gödel's incompleteness theorems show that in a consistent logical system, there will always be self-referential propositions that cannot be proven. Therefore, self-reference not only negates itself but also negates the system in which it resides (Gödel, 1992).

The second dimension or level is the Vicious Circle Principle. Russell's understanding of self-reference is more profound and nuanced. Through repeated observation of self-referential paradoxes, Russell discovered a relationship between the individual and the whole within them. For example, in the Barber Paradox, a barber who shaves those and only those who do not shave themselves should or shouldn't shave himself. The barber's scope of work defines a category of people who do not shave themselves, and it is precisely because the barber is defined within this category that the paradox arises. Similarly, there is the Liar Paradox, such as the statement «Everything I say is a lie», which involves all statements, including the statement itself (Dummett, 2006).

Self-reference describes a whole that contains an individual who needs the whole to be defined, or an individual whose definition involves the whole. Russell found that this phenomenon is universal in self-referential paradoxes. He believed that by prohibiting such wholes or individuals, self-reference could be avoided, achieving consistency in proof and computation processes. Russell named the relationship where individuals and wholes define each other in self-reference the Vicious Circle Principle. (Priest, 1984)

Simplifying this relationship further, we find that in the Barber Paradox, the existence of the barber alone negates the definition of all people who do not shave themselves. In the Liar Paradox, the statement negates all statements that are lies. (Parsons, 1974) Thus, self-referential paradoxes include a phenomenon where a whole

ISSN-e: 1885-5679

is negated by an individual it contains. As mentioned earlier, in Gödel's incompleteness theorems, self-reference not only negates itself but also the completeness of the system it is in. In Russell's Vicious Circle Principle, self-reference not only negates itself as an individual but also the definition of the whole it is part of. In fact, this is the essence of Gödel's incompleteness theorems, and this level of understanding elevates self-reference to a more concise and abstract level.

The third dimension or level is the Self. Why must self-reference always be a negation of itself? Why does an affirmation of itself not lead to an opposite judgment? For example, «What I am saying is true», why isn't this statement a paradox? To understand this, self-referential paradoxes must be distilled to a more concise and abstract level. Taking the Liar Paradox as an example, if the statement is a lie, it can be summarized as: defining oneself in a form that negates oneself, so an affirmation of oneself is a negation, and a negation of oneself is an affirmation. Taking «Everything I say is a lie» as an example, it can be summarized as: defining oneself by negating the whole in which it exists, so a negation of the whole is also a negation of oneself, and a negation of oneself is also an affirmation of the whole. In fact, both situations are essentially the same, differing only in the scope of self-definition (Visser, 2004). In reality, all self-referential paradoxes can be reduced to these two points, such as the Grandfather Paradox, Brain in a Vat Paradox, Socratic Paradox, Barber Paradox, and so on.

Why do paradoxes arise? Defining oneself by negating oneself is affirming oneself by negating oneself. Thus, the paradox already appears in the definition. Russell considered that since the definition is contradictory, such self-contradictory propositions cannot be used to judge other propositions or for linguistic communication; they are more like semantic logic games. What is the significance of their existence? This is a question that has always puzzled Russell. To solve this problem, a clear understanding of the vague concept of «the self» is needed (Rescher, 2001).

In mathematics and physics, everything is discussed from the perspective of mutual consistency and interaction, and the situation of only having oneself is generally not discussed. However, it cannot be denied that the existence of oneself also has its meaning in terms of existence. For example, a point on a coordinate plane may exist in

ISSN-e: 1885-5679

the Cartesian coordinate system or the complex plane; it may be a point in a linear function or a higher-order function. It has almost no definition of its position within the plane, but the point has a unique and definite position relative to itself, which is true from the perspective of position. In mathematics, discussing a point's position relative to itself does not make sense; however, self-reference involves precisely what mathematics prohibits, that is, the part that is meaningless in mathematics. (Hausman *et al*., 2021)

The origin point that is in all changes is precisely the point that is assumed to have a definite position. The same applies to fundamental particles; although they do not have a definite position from the perspective of interaction and measurement, their position relative to themselves is definite. This relative definiteness may not have meaning, but this relativity exists and cannot be denied. Self-reference ensures this, so the self is an inherent relativity of things. As long as the self exists, this relativity exists. It can even be said that as long as there is existence, what comes first is this relativity. (Cantini and Bruni, 2024)

Returning to Russell's confusion, what is the significance of defining oneself in a form that negates oneself? As we all know, the premise of self-referential paradoxes is a definite proposition. For a definite logical proposition, there are only four possible logical judgments: first, the logical proposition is correct; second, the logical proposition is incorrect; third, the logical proposition is both correct and incorrect; fourth, the logical proposition is neither correct nor incorrect. The first two situations indicate that the logical proposition is decidable, while the latter two indicate that the logical proposition is undecidable. (Barwise and Etchemendy, 1987)

An undecidable definite logical proposition implies that its certainty is inherent, innate, and not subject to external constraints. In other words, the form of negating oneself reflects the unaffected certainty of oneself. As mentioned earlier, in the classic Liar Paradox, the self is defined in a form that negates itself, so an affirmation of oneself is a negation, and a negation of oneself is an affirmation. In the paradox «Everything I say is a lie», the self is defined by negating the whole in which it exists, so a negation of the system as a whole is also a negation of oneself, and a negation of oneself is also an affirmation of the system as a whole. In summary, the affirmation and negation of oneself can coexist in form because the certainty and uncertainty of oneself coexist.

ISSN-e: 1885-5679

This is consistent with the internal relativity of the self-mentioned earlier. Thus, the relativity of the self is the hidden certainty in self-referential paradoxes; that is, the self is the root cause of paradoxes. (Landini, 1998)

From the Vicious Circle Principle, we know that the concept of the self can radiate to the whole in which it exists, and within this whole, the self is callable. Fractals, recursion, nesting, holography, and other aspects of systems are all related to the relativity of calling oneself. When we understand self-reference and self-referential paradoxes, we only emphasize the description of the self, but we do not prove and interpret what the self is. Because it seems that the self is very simple, does it not refer to the same proposition? However, this «same» is not simple; it contains a certainty that is not subject to judgment constraints. The basic concept that has always been overlooked hides the true secret of self-referential paradoxes. Therefore, the understanding of the third dimension or level is not only from a mathematical perspective but also from the perspective of the existence of all things in the universe, and the meaning of the self is richer than imagined (Landini, 1998).

Now, returning to the threat of circular measurement faced by natural sciences, social sciences, and humanities. When circular measurement problems arise, we do not have sufficient or inadequate evidence to claim that two measurement procedures are measuring the same characteristic, or that a specific procedure is indeed measuring the characteristic we are interested in. Why is that? On the surface, it is caused by the ambiguity of the concepts involved in the measurement process, but in essence, it is because the validity of the measurement (*i. e.*, the system as a whole) depends on the object it is trying to measure, such as certain concepts or theories, and cannot be independent of the object being measured, which inevitably leads to circular reasoning in the measurement results. This circularity questions the validity and accuracy of the measurement because it cannot provide objective evidence of the measured characteristic. In other words, the essence of the circular measurement problem is a logical paradox that inevitably arises in the measurement process, that is, a self-referential paradox.

ISSN-e: 1885-5679

§ 3. The solutions to the threat of circular measurement

The naturalness of the universe, that is, its spontaneity, is mirrored in the self-organization of its evolution. However, this characteristic lacks roots in classical mathematics, indicating that our understanding of the universe remains overly formalized. Mathematics, as a logical mapping of the real physical world, is effective because it avoids self-reference. Correspondingly, the validity and authenticity of the real physical world are also rooted in the avoidance of self-reference. In other words, the reality we inhabit must contain an inherent mechanism that spontaneously resists the logical paradoxes that could arise from self-reference, thereby ensuring the harmony and order of the physical world. Based on this belief, we will draw on mathematicians' approaches to the current third mathematical crisis to resolve the issue of circular measurement, that is, by clarifying the connotation and extension of concepts to avoid paradoxes of self-reference. In terms of methodology, Professor Wang Tao, the most authoritative interdisciplinary researcher in China and an advocate of the third generation of science, offers the most practical solution. Therefore, we will adopt his proposed approach to address the fundamental issues faced by this scientific research (Wang, 2017).

Concepts have two core aspects: connotation and extension. Concepts in the natural sciences are often clearly defined by a single attribute in terms of their extension, whereas concepts in the social sciences and humanities are often difficult to define in this way. However, if we look at concepts from a broader perspective that can fully encompass the more fundamental concepts of the defined concept, it becomes easier to clarify the extension of the concept. This often requires a multidisciplinary approach. While the connotation of a concept can be studied within a single discipline, other disciplines' factors are likely to influence the connotative attributes of the defined concept and its development and change, thus requiring analysis from multiple angles, across disciplines, and across historical stages. This interdisciplinary definition method includes both extensional and intensional definitions. The steps for this method are as follows:

First, find a universal concept that is one level higher than the concept to be defined. We know that the extension and connotation of a concept are generally inversely

ISSN-e: 1885-5679

proportional; the larger the extension, the smaller the connotation. Universal concepts often have the smallest or simplest connotations, making them easier to define or explain. Based on this universal concept, we can study the extension of the defined concept.

Second, classify the already defined or clearly explained universal concept, that is, study its various sub-concepts. Even concepts from other disciplines unrelated to the defined concept should be clearly studied and defined. Based on these sub-concept relationships, define the concept extensionally.

Then, consider the concept defined through extension as a universal concept and classify it at a finer level using a similar method to the one above. Use this as the basis for an intensional definition. Finally, a complete interdisciplinary definition of a concept is the synthesis of extensional and intensional definitions.

Of course, before doing the above work, we first summarize the types of interrelationships between concepts.

3. 1. The types of interrelationships between concepts

In set theory, the relationships between subsets within a universal set include mutual exclusivity, intersection, inclusion, equality, quasi-equality, complementarity, and adjacency. Similarly, concepts that are at the same hierarchical level as the defined concept also have these types of relationships (Wang, 2017). Here is a diagrammatic representation:

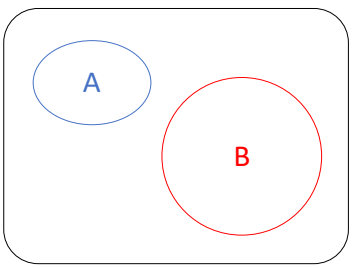


FIGURE 1. **Exclusivity**

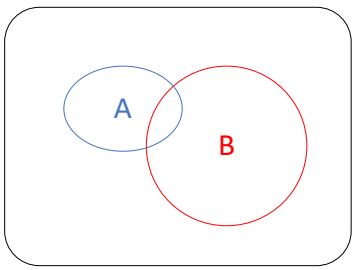


FIGURE 2. **Intersection**

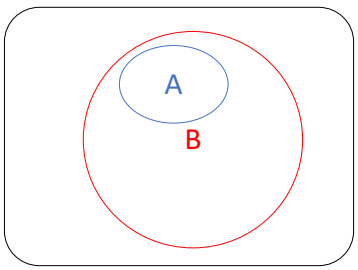


FIGURE 3. **Inclusion**

1 Mutual exclusivity. In the complete set of concepts, concept *A* and concept *B* do not intersect at all, and they are not adjacent either. They belong to an exclusive relationship. As shown in Fig.1.

2. Intersection. In the total set of concepts, if concepts *A* and *B* overlap with each other, they belong to an intersection relationship, as shown in Fig.2.

3. Inclusion. In the complete set of concepts, concept *A* and concept *B* are in a containment relationship, where *A* is contained in *B*, or *B* contains *A*. In this relationship, concept A becomes a subset of concept B, as shown in Fig.3.

4. Equality. The equivalence of concepts means they are completely the same. This is quite simple, so we don't need to discuss it in detail.

5. Quasi-equality. The connotations of the two concepts are mostly the same, but there are a few differences. There are two types of quasi-equivalence relationships: one is inclusive quasi-equivalence, and the other is intersective quasi-equivalence. In inclusive quasi-equivalence relationships, concept *B* is not identical to concept *A*, but it is very close to being identical, and *B* includes *A*, as shown in Fig.4. In this relationship, concepts *A* and *B* are often easily confused. Therefore, we need to pay special attention to this relationship. For example, architecture and construction belong to this relationship. The concepts of quasi-equivalence are easily confused with each other, especially for non-professionals. At the same time, professional scholars often like to show their professional abilities by accurately distinguishing quasi-equivalence concepts. In the intersective quasi-equivalence relationship, concept *B* and concept *A* are essentially intersective relationships, but because the vast majority of their extensions overlap, they form an intersective quasi-equivalence relationship, as shown in Fig.5.

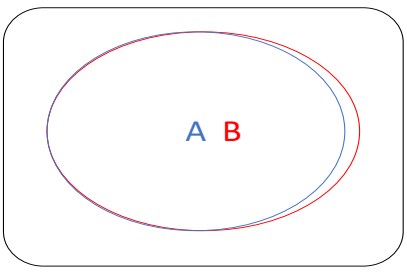


FIGURE 4. **Inclusive quasi-equivalence**

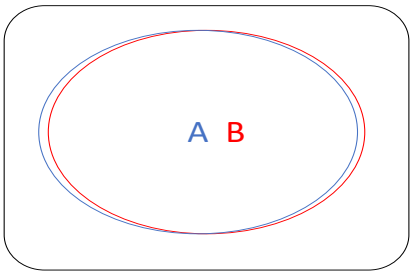


FIGURE 5. **Intersective quasi-equivalence**

6. Complementarity. In the complete set of concepts, concept *A* and concept *B* do not intersect, and their sum exactly equals the total set of concepts. At this time, it can be expressed as *B* = *non-A*, which is the complementary relationship, as shown in Fig.6.

ISSN-e: 1885-5679

7. Adjacency. Concept *A* and concept *B* do not intersect, but if we add them together to form a larger concept *C*, and consider this concept *C* as the universe, then concept *A* and concept *B* can form a complementary relationship within it. From the perspective of concept *C*, concept *A* and concept *B* are mutually exclusive. But from the perspective of concept *O*, concept *A* and concept *B* are adjacent relationships. In an adjacent relationship, those adjacent elements are easily confused with each other, and there may even be some blurry intersecting areas. If we divide them precisely and gradually reduce the intersecting area to zero, the original intersecting place will become adjacent, as shown in Fig.7.

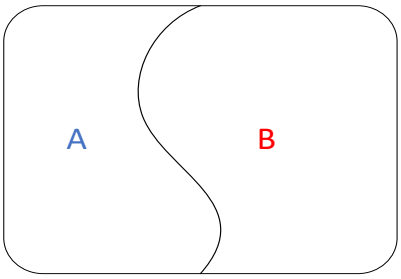


FIGURE 6. **Complementarity**

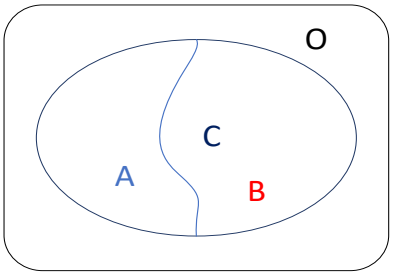


FIGURE 7. **Adjacency**

3. 2. The extensional definition method of concepts

After categorizing the relationships between concepts, it becomes clear what the extensional relationships are between a defined concept and other concepts at the same level. The method of extensional definition is somewhat similar to determining the boundaries of a country. It is difficult to fully understand where the boundaries lie just by looking from within a country. First, we need to identify the land and oceans of the entire Earth, excluding the high seas and territories without sovereignty, like Antarctica, and then remove the territories of other countries to determine the scope of a country's boundaries. Historically, the boundaries of a country are likely to change, and these changes can only be accurately understood through the historical interactions, wars, and treaties between this country and others (Wang, 2017).

Defining the extension of a concept often means determining the scope of a discipline or a field within it, and more importantly, it defines the territory or sphere

of influence of a group of scholars who claim this territory as their «sovereignty» to apply for research funding.

If one only considers the core cases of a concept, it often seems quite clear. However, when it comes to the boundaries between different concepts, disputes are likely to arise. Just as there is generally no dispute over the capitals of neighbouring countries, disputes often occur along the border lines. The emergence of a new concept is often first achieved by establishing some of the most typical cases (similar to building capital). Once the most typical case is determined, when checking its boundaries with other concepts, it becomes less clear. For example, originally, it was thought that humans only had two states: health and disease. Later, it was discovered that some states were unclear whether they belonged to health or disease. This is mainly due to excessive fatigue, staying up late, lack of sleep, and high psychological stress, leading to symptoms such as mental confusion, fatigue, lack of concentration, and memory decline, but it cannot be said that one has fallen ill. This state is thus referred to as sub-health. If it is about such typical cases, people seem to be very clear about what sub-health is. However, when it comes to precisely defining the exact boundaries between health, sub-health, and disease, and classifying certain states precisely, it becomes difficult to determine what to call some states that are on the border.

The method of extensional definition studies all concepts in the concept universe simultaneously, and in fact, it also defines them at the same time. Only in this way can we accurately understand what their mutual boundaries are.

Suppose we want to define concept B. First, find and clearly define the universal concept *O*; then conduct a next-level classification study of concept *O*, and identify concepts *A*, *B*, *C*, *D*, *E*, and *F*. *A* and *B* are mutually exclusive, *C*, *D*, and *F* are adjacent to *B*, and *E* intersects with *B*. Through previous interdisciplinary research, the results obtained may be as shown in Fig.8.

Thus, we can define *B* in this way: concept *B* is the remaining part of the universal concept *O* after excluding *A*, *C*, *D*, *F*, and the parts of *E* that do not intersect with *B*. This method of concept definition does not start from the intrinsic attributes of *B*, but rather from an external, interdisciplinary perspective that determines other concepts and uses extensional exclusion for definition. It may have nothing to do with the connotation and attributes of the defined concept *B*, but it requires a clear

ISSN-e: 1885-5679

understanding of the universal concept *O* that is different from *B*, as well as the extension and even the main connotation of all other concepts within the universal concept *O*. This extensional definition method has several advantages:

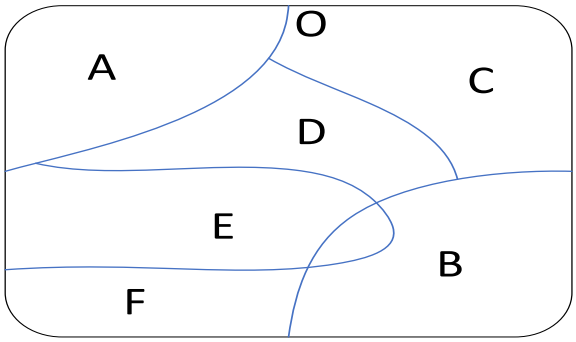


FIGURE 8. **The extensional definition method of concepts**

First, it can discuss the extensional boundaries of the defined concept very clearly because it clarifies the relationships with all other concepts at the same level. Second, the connotation of the universal concept is the most clear and simple, thus avoiding the high complexity that comes with starting from the connotation of *B*. Third, it is easy to clarify the changes in the extensional boundaries of concepts when they shift due to changes in time.

The more concepts adjacent to the defined concept, the more connotative attributes are needed to distinguish its extensional boundaries. Therefore, if one starts with connotative attributes to study, it would lead to the complexity and challenge of how various attributes of a concept play different roles under different circumstances to differentiate it from other concepts. Not only would it be unclear about the concept being studied, but it would also make the expression of the concept definition complicated and obscure.

For example, some national boundaries are defined by rivers, so an attribute definition can be used: the middle line of a certain river serves as the boundary, with the east side being country *A* and the west side being country *B*. Some are defined by lines of latitude and longitude, which can also use attribute definitions. However, in reality, many national boundaries may not have any simple attribute characteristics; they are determined by where the legal boundaries are drawn, and the final boundary is where it is drawn. In such cases, the extensional definition method is most suitable.

ISSN-e: 1885-5679

Drawing a map of the actual boundaries of all concepts is simple and clear. If one insists on using attribute definitions, it is not absolutely impossible, but it would make the expression extremely long and difficult to understand.

3. 3. The intensional definition method of concepts

After the extensional definition, the boundary scope of the defined concept is established, and then the intensional definition method can be used for a more in-depth and detailed definition (Wang, 2017).

The intensional definition method is similar to the aforementioned interdisciplinary research, where we first consider the defined concept as a universal concept and then study its classification and attributes.

Traditionally, it was thought that the connotation of a concept is its attributes. By determining the attributes of a concept, its extension and the relationships with other concepts are also determined. However, our understanding of the connotation of a concept here is broader, including:

1. Attributes; as previously mentioned, no further elaboration is needed here.

2. Further classification or sub-level classification within the concept —in different sub-level classifications, there may be many different attributes. The attributes of a concept may require the sum of all attributes of these sub-level classifications to be accurately understood. For example, to understand the concept of Chinese traditional culture, people's first reaction would be to think of the traditional culture represented by Confucianism. However, China has had some ethnic minorities throughout its history, and today, their cultures differ from the Confucian tradition. We must study the attributes of different sub-cultural subjects in Chinese history separately and then synthesize them to view them as the complete connotation of Chinese traditional culture.

3. Multi-angle classification research. The classification of a concept is not necessarily a single result; different angles of division may bring different classification results, thus obtaining different connotation attributes of sub-level concepts. For example, Chinese traditional culture can be viewed not only from the perspectives of political systems and moral thoughts but also from the angles of science and

ISSN-e: 1885-5679

technology, food culture, literature, painting, music, drama, and other arts, agriculture, and industrial production culture, marriage and funeral culture, folklore, etiquette, commercial culture, martial arts, and other specific cultures for classification, analysis, and understanding;

4. Interaction and influence with other concepts —the attributes, extension, and classification of the defined concept may be changed by the changes in other concepts;

5. Historical evolution —the connotation of a concept may change. Especially due to innovation and technological development, objects in the social field will undergo increasingly rapid development and changes. This requires a timely understanding and cognition of the connotation of the defined concept. For example, to define the object of the world's highest mountain, Mount Everest, if viewed from a span of a few years, Mount Everest is relatively static, with the main changes being the flora and fauna on the mountain surface and the streams in the mountain gullies. However, if viewed from a span of hundreds of millions of years, the body of Mount Everest itself will change greatly due to the Earth's mountain-building movements, and it may even be unrecognizable. Some concept objects are relatively easy to determine their starting point in time, while others are vague. For example, the United States of America was established on July 4, 1776, and this time point is very clear. However, if we want to study the history of the United States clearly, we cannot only look at this time point; in fact, it started a long time before, at least from the early 17th century during the colonial period, and even earlier to understand it properly. If there are specific historical events as markers, the starting time point is relatively easy to define precisely. However, if there is no specific time point, the process of change is slow and gradual, making it difficult to determine the starting time point. For example, the starting time point of pristine civilizations is often difficult to determine because they evolve gradually on their own. In contrast, the starting time point of new civilizations established by foreign civilization invasions is relatively easier to determine because they often correspond to events of conquering their core areas through war.

Therefore, the connotation of a concept is not necessarily singular and pure but may be extremely rich and diverse and subject to development and change. This intensional definition is more in line with the realities of fields such as social sciences and humanities.

ISSN-e: 1885-5679

§ 4. Conclusion and discussion

The introduction of this article lists the circular measurement issues encountered in different fields of natural science, social science, and humanities, showing how circular measurements threaten our interpretation and trust in measurement results. To avoid this adverse effect, modern biomedicine has proposed the method of randomized double-blind experiments and the concept of evidence-based medicine based on Koch's postulates. However, the method of randomized double-blind experiments has inherent limitations, and even the term «evidence-based medicine» itself is perplexing. Below, I will quote Professor Wang Tao, a leading interdisciplinary researcher in China and an advocate of the third generation of science, to briefly explain this issue (Wang, 2017).

The underlying principle of randomized double-blind experiments is to use the cancellation of random errors through equal-precision redundant measurement[2]. Additionally, by analyzing the results of comparative tests under the same conditions, some systematic errors can be offset. Therefore, only by delving deeper into the principles of equal-precision redundant measurements, random errors, and systematic errors can we have a clear understanding of whether randomized double-blind experiments are effective and what problems they may have.

The purpose of randomization and blinding is to ensure randomness, especially to exclude human interference. However, does merely employing a randomized double-blind design always ensure randomness and the cancellation of random errors? Not necessarily. For example, a company has developed a new drug, and underlying diseases such as heart disease can affect the efficacy of the drug. To shield against this influence, we hope that through random double-blind drawing, there will be an equal number of heart disease patients with the same condition in both the placebo group and the experimental group. However, due to the limitations of heart disease patients, among 1600 subjects, only 10 are heart disease patients. The placebo group drew 7 people, and the drug experimental group had 3, 4 fewer than the placebo group. Two

[2] Equal-precision redundant measurement refers to the process of conducting multiple repeated measurements of the same quantity under the same measurement conditions, and on this basis, increasing the number of observations or observation points to enhance the accuracy and reliability of the measurement results. The core of this measurement method lies in reducing the impact of random errors by increasing the number of observations and using statistical methods to process the measurement data, thereby obtaining more precise measurement results.

ISSN-e: 1885-5679

people died in the placebo group. Since very few people actually die in phase three drug-randomized double-blind experiments, these two heart disease-related deaths in the placebo group can have a significant impact on the experimental results. Since there may be many underlying diseases that affect drug experiments, have these randomized double-blind experiments been fully studied for the randomness of various underlying diseases? It is difficult to see relevant content in many papers.

Once the principles of equal-precision redundant measurements and random errors are understood, it is not necessary to have randomized double-blind experiments to conduct very effective experiments. For example, one could specifically seek out 100 heart disease patients and fully understand whether the subjects have other underlying diseases. Divide them into two groups as evenly as possible according to age, gender, and other underlying disease conditions, with one group as the placebo group and the other as the drug experimental group. Such an experiment can specifically study the impact of drug use on the results in patients with underlying heart disease. Although this is not entirely randomized and double-blind and does not require as many samples, this experiment, entirely based on measurement science, will yield more valuable results. Generally, obtaining more samples increases randomness, but at the same time, more natural samples also introduce more factors. For example, the more types of underlying diseases in the samples, the more influencing factors there are, and the more difficult data analysis becomes.

Moreover, equal-precision redundant measurements can effectively offset random errors but cannot completely offset systematic errors. For example, assume that the ambient temperature has a certain impact on the drug. If the temperature in the area where the randomized double-blind experiment is conducted is generally high, causing a significant decrease in drug efficacy, then the randomized double-blind experiment conducted there, although the experimental process is entirely qualified, cannot obtain accurate results due to the uneliminated systematic errors.

The purpose of double-blinding is to eliminate the psychological influence of subjects and experimenters. If subjects know they have taken medication, they may have complex psychological effects, leading to biases in their verbal reports of the effects of the medication. This method is also reflected in other similar experiments. For example, find a group of loyal users of a certain drinking water brand who firmly

ISSN-e: 1885-5679

believe that this brand of water has a special taste, is very delicious, and is significantly different from other brands of water and tap water. The experimenter provides this brand's water, another brand's water, and tap water, but removes all identifiers, allowing all subjects to taste one by one and state which is their favorite brand, which is tap water, and which is the other brand's water. The result is that most people got it wrong, and even the highest-rated turned out to be tap water (Mastinu *et al*., 2022).

Experimenters may also be biased; for example, if they tend to support or deny the efficacy of the drug in advance, it will cause them to interfere with the experimental process, data acquisition, and data processing. Research in related fields generally uses P-values to judge whether experimental results are acceptable, without the most basic concepts of measurement science, such as random errors, Gaussian distribution, systematic errors, and equal-precision redundant measurements (Wang, 2017). Whether it is experimental papers and reports published in general professional journals or those published in Nature, Science, Cell, The Lancet, etc., most of them are seriously lacking in measurement literacy. How can this ensure the scientific nature of experimental design and results? If there is no analysis of the distribution of interfering factors, whether random or not, random errors, systematic errors, etc., in the experimental papers, no matter which journal they are published in, they are considered garbage. The above is not a criticism of a particular researcher but a common problem in the entire industry or field.

The term evidence-based medicine itself is somewhat perplexing; does it imply that medicine before this approach did not require scientific evidence to treat patients? The key issue is how to obtain truly scientific evidence and how to apply it. Metrology is the only foundational theory and knowledge required. There is no need for any field to close itself off and ponder in isolation, as no field is special in this regard. Pondering in isolation for a long time will only lead to half-baked results. Establishing medical metrology based on metrology is sufficient. Do we need to create evidence-based physics, evidence-based chemistry, evidence-based astronomy, evidence-based naval architecture...?

The positive rate, false positive rate, negative rate, and false negative rate, often used in fields such as molecular biology, are inherently incorrect concepts. For example, when it is said that the positive rate of an antigen reagent is a certain

ISSN-e: 1885-5679

percentage, how is this measured? It is done by testing within a positive group to see how many results are positive. However, as a measurement tool, its core indicator is sensitivity (which can be called the detection limit because it can be strictly defined using the sensitivity of metrology). If all the sample pathogen concentrations in a positive group are above the sensitivity, the positive rate can reach 100%. If all the sample pathogen concentrations are below the sensitivity, there will be no positive results, and the positive rate will be 0. We can only say what the actual proportion of positives measured by the highest sensitivity tool in an actual group is; how could there be an indicator of the positive rate for a measurement tool? These are also common problems in the current biopharmaceutical industry.

This is also the reason why the proportion of retracted papers in this field is particularly high. Many times, retraction is not due to personal ethical issues of the paper's authors but rather a serious lack of basic metrology literacy (Wang, 2017).

In short, cyclic measurement problems are a type of epistemological issue that may arise during the measurement process, namely the circularity issue in cognition. This issue involves how to prove that certain measurement procedures can indeed identify the quantities or characteristics of interest, especially in the absence of independent methods to assess these procedures. Specifically, this circularity issue involves two dimensions of measurement: the material dimension and the representational dimension. Measurement is not only an interaction of physical processes but also involves the representational relationship between the measured item and the target attribute. The risk of this circularity issue is greater when the theoretical understanding of the quantities or characteristics of interest is weak, as the answers to What counts as a measurement of *X*? and What is *X*? often seem to presuppose each other. This means that the risk of the threat of cyclic measurement is more likely to occur when the empirical relationship between the representative and the represented is still forming (Luchetti, 2024). This article analyzes the origins and nature of cyclic measurement problems and proposes possible solutions, which may provide some inspiration for scientific research in different fields. Of course, due to cognitive abilities and levels, this article may still have some shortcomings, and we hope to receive criticism and correction from fellow researchers.

ISSN-e: 1885-5679

References

Alexandrova, Anna; and Daniel M. Haybron (2016), «Is Construct Validation Valid?», *Philosophy of Science* 83 (5), pp. 1098-1109.

Barlas, Yaman; and Stanley Carpenter (1990), «Philosophical roots of model validation: Two paradigms», in *System Dynamics Review* 6 (2), pp. 148-166.

Barwise, Jon; and John Etchemendy (1987), *The Liar: An Essay on Truth and Circularity*. Oxford University Press.

Bodea, Marcel (2013), «Some Operations in the Class Algebra of NBG to Make Different Mathematical Objects - A Philosophical Approach», in *Logos Architekton. Journal of Logic and Philosophy of Science* 7 (01), pp. 7-40, <https://api.semanticscholar.org/CorpusID:172132180>, [25/06/2026].

Borsboom, Denny; and Keith A. Markus (2013), «Truth and Evidence in Validity Theory», in *Journal of Educational Measurement* 50 (1), pp. 110-114.

Cantini, Andrea; and Riccardo Bruni (2024), «Paradoxes and Contemporary Logic», in *The Stanford Encyclopedia of Philosophy (Winter 2024 Edition)* edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University, <https://plato.stanford.edu/archives/win2024/entries/paradoxes-contemporary-logic>, [25/06/2026].

Chang, Hasok (2004), *Inventing Temperature: Measurement and Scientific Progress*. Oxford University Press.

Dummett, Michael (2006), «The Vicious Circle Principle», in *Cambridge and Vienna: Frank P. Ramsey and the Vienna Circle*, edited by Maria Carla Galavotti, 29-33. Dordrecht, Springer Netherlands.

Felgner, Ulrich (2006), *Models of ZF-set theory*. Vol. 223. Springer.

Garaway, G. B. (1997), «Evaluation, validity, and values», in *Evaluation and Program Planning* 20 (1), pp. 1-5.

Ghilardi, Giampaolo; and Flavio Keller (2012), «Epistemological Foundation of Biometrics», in *Second Generation Biometrics: The Ethical, Legal and Social Context*, edited by Emilio Mordini, *et al*. Dordrecht, Springer Netherlands, pp. 23-47

Gödel, Kurt (1992), *On formally undecidable propositions of Principia Mathematica and related systems*. New York, Courier Corporation.

Goertz, Gary; and James Mahoney (2012), «Concepts and measurement: Ontology and epistemology», in *Social Science Information* 51 (2), pp. 205-216.

Hausman, Alan, *et al*. (2021), *Logic and Philosophy: A Modern Introduction* (thirteenth edition). Indianapolis, Indiana, Hackett Publishing Company, Inc.

Hood, S. Brian (2009), «Validity in Psychological Testing and Scientific Realism», in *Theory & Psychology* 19 (4), pp. 451-473.

James-Brabham, Ella, *et al*. (2023), «How do socioeconomic attainment gaps in early mathematical ability arise?», in *Child Development* 94 (6), pp. 1550-1565, <https://doi.org/10.1111/cdev.13947>, [20/06/2026].

Landini, Gregory (1998), *Russell's Hidden Substitutional Theory*. Oxford University Press.

Luchetti, Michele (2024), «Epistemic circularity and measurement validity in quantitative psychology: insights from Fechner's psychophysics», in *Frontiers in Psychology* 15, <https://doi.org/10.3389/fpsyg.2024.1354392>, [20/06/2026].

ISSN-e: 1885-5679

Luchetti, Michele (2020), «From successful measurement to the birth of a law: Disentangling coordination in Ohm's scientific practice», in *Studies in History and Philosophy of Science Part A* 84: pp. 119-131.

Mari, Luca (2005), «The problem of foundations of measurement», in *Measurement* 38 (4), pp. 259-266.

Mari, Luca (2003), «Epistemology of measurement», in *Measurement* 34 (1), pp. 17-30.

Markus, Keith A. (1998), «Science, measurement, and validity: Is completion of Samuel Messick's synthesis possible?», in *Social Indicators Research* 45 (1), pp. 7-34.

Markus, Keith A., and Denny Borsboom (2013), *Frontiers of Test Validity Theory: Measurement, Causation, and Meaning (1st ed.)*. New York, Routledge.

Mastinu, Mariano, *et al*. (2022), «A Simple Taste Test for Clinical Assessment of Taste and Oral Somatosensory Function-The "Seven-iTT"», in *Life (Basel)* 13 (1).

Micheli, Pietro, and Luca Mari (2014), «The theory and practice of performance measurement», in *Management Accounting Research* 25 (2), pp. 147-156.

Monteiro, Renan P., *et al*. (2022), «The Efficient Assessment of Self-Esteem: Proposing the Brief Rosenberg Self-Esteem Scale», in *Applied Research in Quality of Life* 17 (2), pp. 931-947, <https://doi.org/10.1007/s11482-021-09936-4>, [20/06/2026].

Parsons, Charles (1974), «The liar paradox», in *Journal of Philosophical Logic* 3 (4), pp. 381-412.

Priest, Graham (1984), «Logic of paradox revisited», in *Journal of Philosophical Logic* 13 (2), pp. 153-179.

Punniyamoorthy, M.; and M. Prasanna Mohan Raj (2007), «An empirical model for brand loyalty measurement», in *Journal of Targeting, Measurement and Analysis for Marketing* 15 (4), pp. 222-233, <https://doi.org/10.1057/palgrave.jt.5750044>, [26/06/2026].

Rescher, Nicholas (2001), *Paradoxes: Their Roots, Range, and Resolution*. Chicago, Open Court.

Strauss, Milton E., and Gregory T. Smith (2009), «Construct Validity: Advances in Theory and Methodology», in *Annual Review of Clinical Psychology*, vol. 5, pp. 1-25.

Tal, Eran (2013), «Old and New Problems in Philosophy of Measurement», in *Philosophy Compass* 8 (12), pp. 1159-1173.

Traini, Luca, *et al*. (2024), «AI-driven Java Performance Testing: Balancing Result Quality with Testing Time», in Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering, <https://doi.org/10.1145/3691620.3695017>, [25/06/2026].

Vallin, Robert W. (2013), *The Elements of Cantor Sets: With Applications*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118548745>, [25/06/2026].

Van Dalen, D. (1972), «Set theory from Cantor to Cohen», in *Sets and Integration: An outline of the development*, 1-74. Dordrecht, Springer Netherlands.

Visser, Albert (2004), «Semantics and the Liar Paradox», in *Handbook of Philosophical Logic*, edited by D. M. Gabbay, *et al*., Dordrecht, Springer Netherlands, pp. 149-240.

Wang, Tao (2017), *Experiment, measurement, and science (in Chinese)*. Oriental Publishing House.

Whitehead, Alfred North, and Bertrand Russell (2019), *Principia Mathematica (Volume III)*. Alpha Edition.

Yao, Buwen, *et al*. (2020), «Comparing Shape Categorization to Circularity Measurement in the Evaluation of Median Nerve Compression Using Sonography», in *J Diagn Med Sonogr* 2020, <https://doi.org/10.1177/8756479319898471>, [205/06/2026].

Zhao, Kino (2023), «Measuring the Nonexistent: Validity before Measurement», in *Philosophy of Science* 90 (2), pp. 227-244, <https://doi.org/10.1017/psa.2023.3>, [20/06/2026].

ISSN-e: 1885-5679

ISSN-e: 1885-5679