

The Hidden Proxy Problem: When the Measured Quantity Is Not the Question

Juan Seguro

3Dresyns by Resyner Technologies S.L.

© 2026 Juan Seguro. All Rights Reserved.

Correspondence: info@3Dresyns.com · www.3Dresyns.com

HEB-004 · Independent Technical Article · Version v0.2 · 22 September 2026

Abstract

A measurement can be precise, repeatable and technically well executed while still failing to answer the scientific question if the quantity measured is only an imperfect proxy for the intended measurand. This article develops the hidden proxy problem as a general failure of scientific inference. The Joint Committee for Guides in Metrology defines the measurand as the quantity intended to be measured and treats measurement models as the relations that connect input quantities to that measurand. That distinction becomes critical whenever an indicator, surrogate, screening metric or intermediate state is silently substituted for the target quantity. Four examples are examined. In photopolymer networks, degree of conversion, free residual monomer, extractables or leachables and biological exposure are related but non-equivalent quantities. In photoinitiation, optical absorption is necessary for photon capture but does not by itself quantify productive radical or ionic initiation. In mechanical characterisation, indentation hardness can be useful for screening and local comparison but does not become fracture toughness or damage tolerance by correlation alone. More generally, a process indicator can support a decision only inside the domain in which its relationship to the target has been demonstrated. The article proposes a practical proxy-validation architecture based on explicit target definition, a stated measurement or causal model, domain-bounded calibration, uncertainty, perturbation testing and independent confirmation. The central conclusion is methodological: indirect measurement is not intrinsically weak, but the inferential bridge between proxy and target must itself be treated as part of the experiment.

Keywords: measurand; proxy measurement; surrogate endpoint; measurement model; degree of conversion; residual monomer; extractables; photoinitiator; absorbance; hardness; fracture toughness; validation; scientific inference

1. Purpose and evidence boundary

This article examines a recurring scientific error: measuring one quantity competently and then interpreting it as though a different quantity had been measured. The problem is not restricted to poor instruments or noisy data. It can occur in technically sophisticated work with excellent repeatability, calibrated equipment and statistically precise results. The failure lies in the inferential step between the measured quantity and the scientific question.

The analysis is methodological. It does not claim that proxy measurements are undesirable. Much of science depends on indirect measurement. Temperature, concentration, molecular conversion, stress, dose and biological exposure are frequently inferred from intermediate signals through validated measurement models. The distinction developed here is therefore between a justified proxy and an unvalidated substitution.

HEB-001 introduced the narrower example that degree of conversion, free residual monomer, extractables or leachables and biological exposure are not interchangeable measurands [1]. HEB-004 generalises that observation across measurement domains. Related Engineering Reviews retain ownership of application-specific qualification requirements; this article addresses the reasoning error that occurs before such requirements are designed.

The article advances one central proposition:

A proxy becomes scientifically informative only through a demonstrated relationship to the quantity the question actually asks about.

2. The measurand comes before the instrument

The International Vocabulary of Metrology defines the measurand as the quantity intended to be measured [2]. That wording is more demanding than it first appears. It requires the experimenter to specify not merely an instrument output, but the quantity, state and relevant conditions that make the intended quantity scientifically meaningful. The same vocabulary defines a measurement model as the mathematical relation among the quantities known to be involved in a measurement, with the measurand as the output quantity inferred from the inputs [3].

These definitions separate three things that are often collapsed in practice: the indication produced by an instrument, the quantity actually observed or calculated from that indication, and the target quantity required by the scientific question. Sometimes the three are almost identical. Often they are not.

A useful minimal representation is:

signal → measured quantity X → model or evidence bridge → target measurand Y

The scientific burden increases as the distance between X and Y increases. If X is a direct realization of Y under a well-characterised procedure, the bridge may be short. If X is a surrogate chosen because Y is difficult, destructive, expensive, slow or inaccessible, the bridge becomes a substantive part of the evidence. Precision in X cannot compensate for an unknown relationship between X and Y.

This suggests a distinction between measurement uncertainty and proxy uncertainty. Measurement uncertainty concerns how well X has been determined. Proxy uncertainty concerns how well X constrains Y in the stated domain. The second can dominate even when the first is very small.

3. Correlation is not identity

A proxy is often adopted because it correlates with a target under some conditions. Correlation can be useful, but it does not make two quantities identical. The relation may depend on composition, geometry, temperature, time, processing history, population, instrument geometry or another latent variable. A calibration built in one domain can therefore fail when transferred to another.

Suppose a proxy X is related to a target Y through a model $Y = f(X, Z)$, where Z represents additional state variables. Replacing that relation with $Y = f(X)$ is justified only if the relevant dependence on Z is negligible, controlled or explicitly incorporated for the intended use. Otherwise two specimens with the same X can have different Y values because they differ in Z .

The strongest warning sign is semantic compression: words such as “cure”, “strength”, “biocompatibility”, “efficiency”, “resolution” or “stability” are used as if each denoted a single measurable property. In reality, each may contain several distinct quantities and endpoints. Before selecting a proxy, the target must be decompressed into the quantity actually needed for the decision.

A correlation can support inference. It cannot erase the distinction between the variables being correlated.

4. Case study I: conversion is not residual monomer, extractables or exposure

Vat-photopolymerised networks provide a clear example because several chemically related quantities are often compressed into the informal idea of “how cured” a material is. Degree of conversion measured spectroscopically reports the fraction of monitored reactive groups that have reacted. It does not directly report the mass fraction of free residual monomer, the composition of material that can be extracted under a defined protocol, the amount that can leach under service conditions, or the biological dose reaching cells or tissue.

HEB-001 used a 2025 study of a vat-photopolymerised UDMA-based material to illustrate the distinction [1,4]. Post-cure conditions changed the measured degree of conversion, but the unreacted C=C fraction cannot be interpreted numerically as the same fraction of free monomer. In a multifunctional molecule, one reactive group may have joined the network while another remains unreacted. The molecule is then chemically incorporated even though residual unsaturation remains detectable.

The converse problem also matters. A low-molar-mass monofunctional species can remain mobile if its reactive group does not react, but whether it becomes extractable or biologically available also depends on network affinity, diffusion distance, washing, post-cure history, specimen geometry and the extraction or exposure environment. Jin and co-workers, for example, measured residual-monomer elution, degree of conversion and mechanical properties as distinct responses when studying postwashing effects in a printed dental material [5]. Treating any one of those responses as a complete substitute for the others would discard part of the experimental state.

The distinction can be written compactly as:

$$\text{degree of conversion} \neq \text{free residual monomer} \neq \text{extractables/leachables} \neq \text{biological exposure}$$

The quantities can be causally connected without being numerically interchangeable. Application-specific Engineering Reviews already own conversion metrology, extractables/leachables qualification and residual-species attribution. The HEB question is more general: at what point did an intermediate chemical indicator become silently substituted for the quantity that the conclusion actually requires?

5. Case study II: absorption is not productive photoinitiation

A second example comes from photochemistry. An absorption spectrum answers whether, and how strongly, a chromophore absorbs at particular wavelengths under specified conditions. Productive photoinitiation asks a later question: how effectively absorbed photons generate the reactive species and subsequent chemistry required for polymerisation.

Between those two quantities lie excited-state formation, internal conversion, intersystem crossing or electron transfer, bond cleavage where relevant, cage escape, quenching, oxygen or inhibitor reactions, radical or ionic reactivity toward the formulation, and propagation. Eibel, Fast and Gescheidt showed explicitly that photoinitiator performance depends on a combination of absorption, dissociation quantum yield, light-source properties, radical reactivity and side reactions rather than on absorption alone [6]. Carroll and co-workers later demonstrated directly that wavelength-resolved photochemical reactivity can diverge from absorptivity, showing why an absorption spectrum cannot be treated as a universal action spectrum [7].

Therefore a spectrum can be an excellent proxy for photon capture and a poor proxy for productive initiation if the downstream photochemistry changes. Two initiators can absorb comparable numbers of photons while generating different numbers of useful reactive species. Conversely, a compound with weaker absorption at the lamp wavelength may perform well if its absorbed photons are used more efficiently.

This is why ER-202 separates spectral photon delivery, absorbed-photon rate, photochemical events, reactive-species formation and formulation response rather than labelling all of them “efficiency” [8]. HEB-004 does not reproduce that qualification framework. Its methodological point is that the location of a measurement along a causal chain must match the location of the conclusion.

Absorption measures entry into a photophysical pathway. It does not, by itself, measure successful exit from the pathway as polymerisation.

6. Case study III: hardness is not fracture resistance

Mechanical testing offers another persistent proxy problem. Hardness measurements are attractive because they can be rapid, local and economical. Instrumented indentation can characterise resistance to local deformation and can support material screening under a defined method. ISO/TS 19278, for example, specifies instrumented micro-indentation hardness for plastics as a reproducible method useful for quality control, material selection and screening [9].

Those virtues do not make hardness identical to toughness, fatigue resistance, tear resistance or structural damage tolerance. Fracture toughness concerns resistance to crack extension under a defined fracture-mechanics configuration. Hardness concerns a response to indentation. Both may depend on network state and microstructure, and correlations can exist within a constrained material family, but the quantities are physically different.

The indentation literature itself demonstrates the distinction. Lawn and Cook describe sharp indentation as a powerful probe capable of interrogating multiple properties while also emphasising size effects, deformation and fracture mechanisms, rate effects and microstructure [10].

Quinn and Bradt showed that Vickers-indentation estimates of fracture toughness involve a crack system and arrest condition fundamentally different from standard fracture-toughness tests [11]. The lesson is not that indentation is invalid. It is that an inferred toughness value requires a model and validation appropriate to the material and geometry; a hardness number alone is not a toughness measurement.

In engineering work this distinction matters because a screening metric can rank formulations correctly within one family and still fail as a transferable design property. ER-070 retains the application-specific ownership of fracture toughness and damage tolerance in vat-printed thermosets [12]. HEB-004 asks the prior epistemic question: what evidence permits a convenient local response to stand in for the failure mode that matters?

7. The proxy ladder: indication, state, performance and decision

Many experimental programmes contain several layers of proxy rather than one. A photopolymer example might proceed from infrared absorbance to conversion, from conversion to network state, from network state to extractables or modulus, and from those quantities to a service or biological decision. Each arrow can be reasonable. The danger arises when the arrows disappear from the report and the first observable is described using the name of the final decision variable.

The same architecture occurs outside photopolymers. An accelerated ageing condition is a proxy for service history; a biomarker can be a proxy for a clinical endpoint; a coupon can be a proxy for a component; a short-term creep test can be a proxy for long-duration deformation; a colour change can be a proxy for chemical conversion; and a single geometric feature can be a proxy for printer capability. None is automatically invalid. Each requires a domain, a model and evidence that the relation remains fit for the intended decision.

A useful four-level ladder is:

indication → material or system state → functional performance → decision

Errors become more consequential as conclusions travel upward without evidence. An instrument may measure an indication accurately, yet the decision can still be wrong because the mapping between levels is unstable or incomplete.

8. When a proxy is legitimate

A defensible proxy does not need to be perfect. It needs to be fit for a declared purpose. Six questions are usually sufficient to expose whether the inferential bridge has been tested.

First, what is the target measurand? The target should be stated as a quantity or endpoint, not a vague material adjective. Second, what exactly is the proxy? The measured quantity, sampling depth, geometry, time point and analytical method must be explicit. Third, why should the proxy constrain the target? This requires a measurement model, causal mechanism or empirical calibration rather than a verbal resemblance. Fourth, over what domain has the relation been demonstrated? Composition, temperature, thickness, process state, population and instrument conditions define the valid region. Fifth, what uncertainty enters through the proxy-to-target relation in addition to measurement uncertainty in the proxy itself? Sixth, what independent observation can falsify or verify the inferred target?

A practical validation sequence is therefore:

define Y → measure X → establish X–Y relation → perturb the domain → quantify uncertainty → confirm Y independently

The perturbation step is important. A proxy that merely tracks the target because both respond to a third variable can fail when that variable changes. Deliberately varying plausible confounders reveals whether the relation is structural or accidental.

A proxy is not wrong because it is indirect. It becomes misleading when the inferential bridge is assumed rather than demonstrated.

9. Multiple proxies and triangulation

When the target is difficult to measure directly, multiple partially independent proxies can be stronger than a single convenient indicator. Spectroscopic conversion, extractable-species analysis and mechanical response, for example, interrogate different aspects of a cured network. Agreement among them does not make them identical, but convergence can constrain the plausible state more strongly than any one measurement alone.

Triangulation is most informative when the proxies fail differently. Two methods built from the same signal processing or the same underlying assumption can reproduce the same bias. Independent physical principles, distinct sampling depths or different response stages provide more useful cross-checks.

Disagreement is also information. If hardness increases while fracture resistance decreases, or conversion increases while extractable concentration does not fall as expected, the mismatch identifies a missing state variable or an inadequate model. The correct response is not to choose the preferred proxy. It is to investigate why the proxies have decoupled.

10. Proxy failure modes

Several recurring patterns deserve explicit suspicion. Range failure occurs when a calibration is extrapolated beyond the conditions used to establish it. Composition failure occurs when a relation developed for one material family is transferred to another with different chemistry or microstructure. Scale failure occurs when surface or small-volume measurements are treated as bulk properties without demonstrating representativeness. Time failure occurs when a proxy and target evolve at different rates. Mechanism failure occurs when the causal pathway connecting proxy and target changes. Decision failure occurs when a useful screening metric is treated as sufficient evidence for a higher-level safety, reliability or performance claim.

A further danger is proxy optimisation. Once a convenient metric becomes a target, experimental development can improve the metric while leaving the intended outcome unchanged or even degraded. The issue is not statistical gaming alone; it can happen in good-faith engineering whenever the measured variable is easier to optimise than the property that originally motivated it.

The antidote is periodic re-anchoring: return to the intended measurand and verify that improvement in the proxy still predicts improvement in the target.

11. Relationship to the Engineering Review corpus

The HEB series addresses failures of reasoning that can recur across scientific domains. The ER series addresses engineering evidence and qualification questions in defined technical contexts. This article therefore does not prescribe acceptance limits, extraction protocols, photoinitiator ranking methods, hardness thresholds or fracture-design values. Those belong to the corresponding application-specific reviews.

The boundary can be expressed simply: HEB asks whether a conclusion follows from what was measured; ER asks what evidence is required to qualify a particular engineering claim. HEB-004 names the inferential failure. ER-027, ER-070, ER-086, ER-092 and ER-202 retain ownership of conversion metrology, fracture and damage tolerance, extractables/leachables, residual-species attribution and photoinitiator action-spectrum/quantum-yield metrology respectively [8,12-15].

12. Conclusions

The hidden proxy problem begins when a convenient observable inherits the name and authority of a different quantity. The result can look scientifically strong because the proxy may be measured with excellent precision, replication and instrumental control. None of those qualities validates the proxy-to-target relation by itself.

The main conclusions are:

1. The measurand is the quantity the scientific question requires, not merely the quantity the instrument happens to output.
2. Indirect measurement is scientifically valid when the relation between proxy and target is demonstrated for the intended domain.
3. Measurement uncertainty in the proxy and uncertainty in the proxy-to-target relation are different contributions and should not be conflated.
4. Correlation can support a proxy but does not make two quantities identical or guarantee transfer outside the calibration domain.
5. Degree of conversion, free residual monomer, extractables or leachables and biological exposure are distinct quantities even when causally connected.
6. Optical absorption is not equivalent to productive photoinitiation because downstream excited-state and reaction steps determine useful chemical output.
7. Hardness is not fracture toughness or damage tolerance; a screening relationship requires its own model and validation.
8. Multiple independent proxies can constrain a target more strongly, especially when their failure modes differ.
9. A proxy should be deliberately perturbed across plausible confounders before it is trusted as a transferable surrogate.
10. The final scientific statement should name the quantity actually supported by the evidence and stop before the inferential bridge becomes untested.

The most dangerous proxy is not an imperfect one. It is the proxy whose imperfection has disappeared from the language of the conclusion.

References

- [1] Segurola, J. The Hidden Error Budget: Measurement-System Failure, False Precision and Experimental Design in Laboratory Science. HEB-001. 3Dresyns Independent Technical Articles (2026). DOI: 10.5281/zenodo.22862642.
- [2] Joint Committee for Guides in Metrology. International Vocabulary of Metrology - Basic and General Concepts and Associated Terms (VIM), 3rd ed., JCGM 200:2012, §2.3, measurand. DOI: 10.59161/JCGM200-2012.
- [3] Joint Committee for Guides in Metrology. International Vocabulary of Metrology - Basic and General Concepts and Associated Terms (VIM), 3rd ed., JCGM 200:2012, §2.48, measurement model. DOI: 10.59161/JCGM200-2012.
- [4] Cheadle, A. M. G.; Maier, E.; Palin, W. M.; Tomson, P. L.; Poologasundarampillai, G.; Hadis, M. A. The impact of modifying 3D printing parameters on mechanical strength and physical properties in vat photopolymerisation. Scientific Reports 15 (2025) 12592. DOI: 10.1038/s41598-025-97294-8.
- [5] Jin, G.; Gu, H.; Jang, M.; Bayarsaikhan, E.; Lim, J.-H.; Shim, J.-S.; Lee, K.-W.; Kim, J.-E. Influence of postwashing process on the elution of residual monomers, degree of conversion, and mechanical properties of a 3D printed crown and bridge materials. Dental Materials 38(11) (2022) 1812-1825. DOI: 10.1016/j.dental.2022.09.017.
- [6] Eibel, A.; Fast, D. E.; Gescheidt, G. Choosing the ideal photoinitiator for free radical photopolymerizations: predictions based on simulations using established data. Polymer Chemistry 9 (2018) 5107-5115. DOI: 10.1039/C8PY01195H.
- [7] Carroll, J. A.; Pashley-Johnson, F.; Klein, M.; Stephan, T.; Pandey, A. K.; Walter, M.; Unterreiner, A.-N.; Barner-Kowollik, C. Microenvironments as an Explanation for the Mismatch between Photochemical Absorptivity and Reactivity. Journal of the American Chemical Society 147(30) (2025) 26643-26651. DOI: 10.1021/jacs.5c06961.
- [8] Segurola, J. Photoinitiator Action-Spectrum and Quantum-Yield Metrology: Absorbed Photons, Reactive-State Formation and Cross-Laboratory Comparability. ER-202. 3Dresyns Engineering Publications (2026). Zenodo record 22842462.
- [9] ISO/TS 19278:2019. Plastics - Instrumented micro-indentation test for hardness measurement. International Organization for Standardization.
- [10] Lawn, B. R.; Cook, R. F. Probing material properties with sharp indenters: a retrospective. Journal of Materials Science 47(1) (2012) 1-22. DOI: 10.1007/s10853-011-5865-1.
- [11] Quinn, G. D.; Bradt, R. C. On the Vickers Indentation Fracture Toughness Test. Journal of the American Ceramic Society 90 (2007) 673-680. DOI: 10.1111/j.1551-2916.2006.01482.x.
- [12] Segurola, J. Fracture Toughness and Damage Tolerance in Vat-Photopolymerised Thermosets: Crack Preparation, Cure State, Anisotropy and Qualification. ER-070. Zenodo (2026). Concept DOI: 10.5281/zenodo.22123345.

- [13] Segurolo, J. Conversion Metrology and Spatial Cure Heterogeneity in Vat Photopolymerisation: From Exposure to Network State. ER-027. Zenodo (2026). Concept DOI: 10.5281/zenodo.22017523.
- [14] Segurolo, J. Analytical Extractables and Leachables from Vat-Printed Photopolymers: Specimen State, Extraction Design, Analytical Coverage and Cross-Study Comparability. ER-086. Zenodo (2026). Concept DOI: 10.5281/zenodo.22123380.
- [15] Segurolo, J. Residual Reactants, Impurities and Manufacturing Conversion as Hidden Drivers of Photopolymer Hazard. ER-092. 3Dresyns Engineering Publications (2026). Concept DOI: 10.5281/zenodo.22156128.

Knowledge boundary

The metrological definitions in this article are grounded in JCGM VIM terminology. The photopolymer examples use published studies and existing 3Dresyns Engineering Reviews only to demonstrate distinctions among measurands and stages of inference. HEB-004 does not re-evaluate the full datasets behind those Engineering Reviews, prescribe application-specific acceptance criteria, or infer safety, biological performance or structural reliability from the proxy examples presented here.

Limitations

The article is a methodological synthesis rather than a quantitative meta-analysis of proxy validity. No universal numerical threshold is proposed for deciding when a proxy is adequate. The strength of a proxy is necessarily conditional on the target measurand, calibration domain, uncertainty, mechanism and intended decision. Examples from photopolymers, photochemistry and indentation are illustrative and should not be transferred as material-specific qualification rules without the corresponding domain evidence.