

Article

A Level-Based Master Plan for Strengthening Research Projects

Adilbek K. Bisenbaev 

Institute of Philosophy, Political Science and Religious Studies, Ministry of Science and Higher Education of the Republic of Kazakhstan, Str. Kurmangazy, 29, Almaty 050010, Kazakhstan; eldarknar@gmail.com

Abstract

This paper proposes a level-based scientific maturation master plan (SMMP) for strengthening research projects prior to manuscript submission. A weak manuscript is often not simply a weak text but an immature project that has been translated too early into publication form. Contemporary research management is better at registering deadlines, deliverables, resources, and visible publication signals than at diagnosing the internal maturity of a scientific object. The result is false readiness: a project may have a topic, structure, literature, methodological vocabulary, and a polished manuscript but still lack a mature problem, a coherent conceptual architecture, a testable design, sufficient evidence, and a disciplined contribution. To address this gap, this paper proposes a nine-level SMMP, moving from thematic impulses to peer review and publication readiness. The model integrates noncompensatory gates, evidence packages, red flags, maturation debt, bottlenecks, the publication maturation gap, and peer-review readiness. Methodologically, the paper is a conceptual design study supplemented by a proof-of-concept documentary application to publicly available CORDIS project biographies. The framework shows how to distinguish publication polish from scientific maturation and how to translate expert criticism into concrete presubmission actions.

Keywords: scientific maturation; research projects; publication readiness; peer review; research governance; false maturity

1. Introduction

A scientific article rarely becomes weak only at the moment of writing. More often, it reaches the manuscript stage already weakened. Its problem was too broad, insufficiently problematized, or weakly differentiated from the assumptions of the existing literature (Alvesson & Sandberg, 2011). Its concepts had not yet become analytical instruments. Its method was chosen by habit rather than by the inner necessity of the question. Its data were collected before the boundaries of the claim had been clarified. Its contribution was formulated after the text, not before the inquiry. For this reason, publication quality cannot be reduced to language, structure, or editorial preparation. Publication weakness is often the late symptom of an earlier condition. The project acquired textual form before it achieved scientific maturity.

Here lies a central blind spot in contemporary research culture. Projects are governed through schedules, budgets, deliverables, milestones, reports, and texts. These elements are necessary. However, they describe the movement of a project, not its scientific maturation. A project may move quickly and remain immature. It may have a plan, a team, a presentation, a proposal, a literature base, and even a manuscript but still lack a mature scientific core. Delay is sometimes not a sign of weakness but a moment of genuine maturation: the



Academic Editor: Andrew Kirby

Received: 15 June 2026

Revised: 8 July 2026

Accepted: 9 July 2026

Published: 14 July 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

reformulation of a question, the abandonment of a false hypothesis, the refinement of operationalization, or the restriction of excessive claims.

The problem is intensified by audit culture and metricized research governance, which can impose substantial emotional and organizational burdens on researchers and research institutions (Watermeyer et al., 2022). Evaluation regimes do not merely measure science. They also shape what becomes valuable, visible, and institutionally rewarded (Power, 1997; Strathern, 2003; Shore, 2008; Hicks et al., 2015). The research project learns to appear ready before it becomes scientifically viable. It produces visible signals of maturity: persuasive rhetoric, extensive literature, tidy tables, methodological vocabulary, article structure, and presentational smoothness. However, visible readiness is not the same as maturity. Moreover, the stronger the institutional pressure for reporting and publication is, the greater the likelihood that scientific maturation will be replaced by publication simulation.

Classical project-management tools solve only part of this problem. Stage-gate approaches structure transitions between development phases and reduce the risk of chaotic movement toward a product (Cooper, 1990, 2008). Readiness models provide a language of levels, thresholds, and accumulated evidence. In research contexts, such models have already been associated with scientific readiness, technology readiness, and service readiness (Hughes et al., 2021; Knar, 2025). However, a research project is not a standard product cycle. Its object is not always stable at the outset. Its question may change during inquiry. Its negative results may indicate intellectual maturity rather than failure (Fanelli, 2012). Its method does not merely serve a predefined object; it may sometimes reorganize the very logic of the inquiry (Law, 2004; Rheinberger, 1997). A scientific project therefore requires not only project management but also maturation governance.

This problem is especially important for scholarly communication. Editors and reviewers usually encounter not the project itself but its late textual trace. They see the manuscript, but rarely the trajectory of its maturation. Peer review must assess novelty, method, evidence, limitations, and contribution. If these elements have not matured before writing, review becomes a late diagnosis of structural defects. It may identify weakness, but it cannot always return the project to the point at which that weakness should have been corrected. Prepublication governance is therefore not an external bureaucracy but a condition of more responsible scholarly communication.

Research on reproducibility and open science shows that publication quality depends not only on how results are presented but also on earlier decisions concerning question formulation, design strength, data transparency, analytical discipline, interpretive boundaries, and scientific incentives (Altman, 1994; Ioannidis, 2005; Nosek et al., 2015; Munafò et al., 2017). In this sense, a publication is the final packaging of a longer epistemic trajectory. If that trajectory is immature, the text cannot fully compensate for its defects. Editorial polishing may make weakness less visible, but it cannot eliminate it.

In this paper, we propose the level-based scientific maturation master plan (SMMP) as an instrument for designing this trajectory. Our aim is not to add yet another checklist to an already overburdened research system. SMMP replaces late and diffuse criticism with earlier, more precise, and more productive diagnosis. It asks which levels a project must pass before its article is not only written but also scientifically mature enough for publication.

SMMP should be distinguished from the Scientific Maturation Plan (SMP), on which it builds. SMP diagnoses whether a project has reached sufficient maturity at a given stage, separating scientific maturity from administrative progress and visible readiness. SMMP extends this logic from stage diagnosis to trajectory design. It does not merely ask whether a project can pass a gate. It specifies the route by which a project becomes sufficiently mature to withstand publication and review. The unit of analysis therefore shifts from stage sufficiency to the prepublication maturation trajectory.

In this paper, we develop a level-based master plan for the scientific maturation of research projects up to the stage of publication and review resilience.

First, we shift the unit of analysis from the manuscript to the project as a maturing epistemic object.

Second, we distinguish publication readiness, scientific maturity, and peer-review readiness.

Third, we introduce the maturation debt, maturation bottleneck, and publication maturation gap indicators.

Fourth, we propose a noncompensatory gate architecture in which strong rhetoric, an extensive bibliography, or stylistic smoothness cannot compensate for an immature problem, an unsuitable method, or an undisciplined claim.

Operational Definitions and Scope of the Construct

Because the language of maturity and maturation is used as its central conceptual vocabulary, these terms must be defined operationally. In this paper, scientific maturation is not treated as a fully transparent inner state that can be exhausted by formal indicators. Mature scholarship also depends on tacit knowledge, disciplinary judgment, mentoring, criticism, collaboration, and intellectual communities (Polanyi, 1966; Becher & Trowler, 2001; Bourdieu, 2004; Collins, 2010). SMMP therefore does not measure the whole living process of maturation itself. It formalizes the observable traces through which that process becomes assessable before publication.

The operational definitions used in the model are presented in Table 1.

Table 1. Operational definitions used in the revised manuscript.

Concept	Operational Definition
Scientific maturity	The degree to which a research project has achieved internal alignment among problem, concepts, question or proposition, method, evidence, limitations, and contribution. It is not identical to textual completion, administrative progress, or rhetorical persuasiveness
Maturation	The process by which an initial thematic impulse is transformed into a disciplined research object capable of supporting a defensible claim. Maturation includes conceptual, methodological, evidential, and contribution-related development
Publication readiness	The condition in which a manuscript appears formally submittable: it has structure, language, sections, references, and visible scholarly form. Publication readiness may occur before scientific maturity
False readiness	A condition in which visible signs of readiness—polished prose, administrative completion, extensive bibliography, or formal manuscript structure—create the impression that the project is mature when key scientific nodes remain immature
Pseudomaturity	A more stable and consequential form of false readiness in which external maturity systematically exceeds actual project maturity across one or more mandatory levels
Maturation debt	The deficit between the actual maturity of a level and the threshold required for responsible transition to the next level
Publication maturation gap	The distance between manuscript readiness and the minimum maturity of the underlying project. A positive gap indicates that textual readiness has outpaced project maturity
Peer-review readiness	The capacity of the project and manuscript to withstand expected reviewer questions concerning novelty, problem clarity, concept logic, method, evidence, limitations, contribution, transparency, and journal fit.

These definitions also clarify the scope of the model. SMMP is not a psychology of creativity, a sociology of mentoring, or a universal theory of discovery. It is an operational architecture for diagnosing the visible maturity of a research project before it becomes a manuscript submitted to peer review.

2. Materials and Methods

Our study is designed as a conceptual design with a retrospective documentary proof-of-concept demonstration component. This design is appropriate when the task is not the statistical testing of an already existing scale but the construction, justification, and preliminary application of a new operational architecture. In our case, the object of construction is a master plan of scientific maturation. It must be theoretically dense, applicable in research practice, and sufficiently formalized for reproducible expert discussion while remaining explicit about the limits of formalization.

In developing SMMP, we relied on five research traditions. First: Stage-gate logic. From this tradition, the model takes transition thresholds, go/no-go decisions, and the possibility of stopping a project when a critical deficit remains.

Second: Readiness and maturity modeling. Here, maturity is understood as movement across levels, supported by evidence of readiness.

Third, the theory of change and the literature on complex interventions, where causal assumptions, context, iterativity, and evidence of mechanism are central, should be considered (Mayne, 2015; Skivington et al., 2021).

Fourth, research evaluation studies show that evaluation regimes are not neutral and can reconfigure the behavior of research organizations (Thomas et al., 2020; Smit & Hessels, 2021; Williams, 2020; Sørensen et al., 2022).

Fifth, the literature on peer review, reproducibility, and research integrity has linked publication quality to transparency, verifiability, bias control, and responsible assessment (Tennant et al., 2017; Moher et al., 2020).

We synthesized these traditions not through the mechanical accumulation of terms. Our synthesis was directed by a managerial and epistemic problem that none of them solved alone. Stage-gate tools control movement but do not always distinguish scientific maturity. Readiness models assess readiness but do not always specify a route of maturation. Peer-review frameworks assess a prepared text but rarely return the project to its early epistemic nodes. Research integrity approaches formulate norms of responsible science but do not always translate them into a level-based project trajectory. SMMP integrates these elements into a single architecture of prepublication maturation governance.

2.1. Model Construction Logic

The model was constructed through theory-driven synthesis rather than empirical induction from the two CORDIS cases. The cases were used after model construction to demonstrate whether the rubric can generate distinguishable diagnoses from open documentary traces. They were not used to derive the nine levels, select the criteria, or set the threshold values.

The criteria were selected according to four design principles. First, each criterion had to identify a necessary maturation node rather than a merely desirable feature. Second, each criterion had to be visible through an artifact, such as a problem statement, conceptual map, method-design fit sheet, evidence dossier, contribution memo, or reviewer-risk matrix. Third, each criterion had to correspond to a recognizable class of reviewer objection: unclear novelty, weak problem, unstable concepts, method–design mismatch, insufficient evidence, excessive claims, or unclear contribution. Fourth, the criterion had to be general enough

for cross-disciplinary use while remaining adaptable to disciplinary standards of evidence, interpretation, and contribution.

The nine levels were therefore not intended as a universal chronology of discovery. They represent an ordered dependency structure. A project can move iteratively, and some levels may overlap, but later maturity normally presupposes earlier forms of stabilization. A contribution cannot be integrated responsibly if the problem is still diffuse; evidence cannot discipline a claim if the method does not fit the question; and peer-review readiness cannot be achieved if the manuscript masks unresolved deficits in problem, method, evidence, or contribution.

2.2. Scoring Calibration and Transition Thresholds

The 0–4 scale is ordinal and diagnostic. It is not intended to produce false numerical precision. Its purpose is to anchor expert judgment in reproducible categories. The transition threshold is set at 3 because this score denotes sufficient maturity for responsible movement to the next level. A score of 4 denotes not only sufficiency but also stable, documented, and transferable maturity. The distinction is important: a project can be allowed to proceed at 3, but it becomes robustly review-resilient only when its key nodes approach 4.

The calibration anchors for the maturity scale are presented in Table 2.

Table 2. Calibration anchors for the 0–4 maturity scale.

Score	Anchor	Observable Meaning	Decision Rule
0	Absent	The feature is not present or cannot be reconstructed from available documents.	Do not proceed. Create the missing artifact or return to the previous level.
1	Intuitive or implicit sketch	The feature exists only as an intuition, slogan, broad intention, or undeveloped claim.	Do not proceed unless the feature is made explicit and connected to the project logic.
2	Partial articulation	The feature is visible but incomplete, weakly connected, or insufficiently documented.	Proceed only for exploratory work; transition to the next level remains premature.
3	Transition sufficiency	The feature is sufficiently specified, documented, and connected to the project to justify movement to the next level.	Proceed, but record remaining risks and maturation debt.
4	Stable maturity	The feature is robust, documented, internally coherent, and likely to withstand informed expert questioning.	Proceed; the level is not merely passable but review-resilient.

For example, in L1, a score of 3 means that the knowledge gap and boundaries are sufficiently clear to support a research question; a score of 4 means that the gap is not only clear but also sharply positioned against prior work and robust against likely reviewer objections. In L4, a score of 3 means that the method fits the question and an analysis plan exists; a score of 4 means that the operationalization, robustness logic, transparency, and limitations are already stable enough for expert scrutiny. In L7, a score of 3 means that the contribution can be stated; a score of 4 means that the contribution is distinguishable, situated, and defensible within the field.

The operational logic of the model uses an ordinal scale from 0 to 4. A score of 0 indicates the absence of the feature. A score of 1 indicates an intuitive sketch. A score of 2 indicates partial articulation. A score of 3 indicates sufficiency for transition. A

score of 4 indicates stable maturity. The scale does not claim false precision. Its function is to distinguish absence, emergence, incomplete assembly, threshold sufficiency, and stable maturity.

Each SMMP level is defined through mandatory criteria, red flags, an evidence package, and a transition rule. The criteria indicate what must mature. Red flags identify deficits incompatible with transition. The evidence package specifies the minimal artifacts through which maturity becomes visible. The transition rule translates expert judgment from a general assessment into a reproducible decision. This is necessary because scientific maturity should not dissolve into the author's charisma, the rhetorical force of the topic, or the general sympathy of the evaluator.

Our key methodological principle is noncompensatoriness. In aggregated assessments, strong elements can conceal a critical defect. For a research project, this is dangerous. Strong literature does not compensate for a weak question. An elegant design does not compensate for the absence of access to data. Rich empirical material does not compensate for unsuitable operationalization. A well-written text does not compensate for the absence of contribution. SMMP therefore uses a minimum function: the maturity of a level is determined not by the average brilliance of the project but by its weakest mandatory node. This principle applies only to mandatory blocking criteria and not to every desirable feature of a project. Discipline-specific adaptation is therefore necessary: In qualitative, interpretive, historical, or theoretical research, the relevant mandatory node may be an interpretive claim, source logic, conceptual necessity, or argumentative warrant rather than a statistical hypothesis. The noncompensatory rule is universal only at the level of logic—a blocking defect cannot be hidden by strengths elsewhere—while the concrete content of the blocking criteria must be calibrated to the discipline and research design.

We deliberately present the documentary component as a proof-of-concept demonstration, not as a full empirical validation. It does not claim final predictive, construct, or psychometric validity. It tests a more modest but important question: can SMMP distinguish different maturation biographies on the basis of open documentary traces of completed projects? Publicly available CORDIS/Horizon materials were used because they may include early project descriptions, reporting pages, results, deliverables, and publication trails (European Commission, n.d.-a). This type of evidence cannot replace full validation, but it can show whether the model remains merely rhetorical or can generate distinguishable diagnoses of project trajectories.

The demonstration cases were selected according to three criteria.

First, the project had to be completed so that its early formulation could be compared with later outcomes.

Second, the public record had to include an early source from which the problem, knowledge gap, boundaries, and initial design logic could be reconstructed.

Third, the late layer had to have sufficient documentary density, including reports, deliverables, publications, indicators, or other publicly visible results. IMPACT-EV and ySKILLS were selected because both have strong late-output trials but differ in the visibility and continuity of the early architecture of maturity.

In the coding procedure, we followed a fixed sequence. First, the early project source was read and coded for problem specificity, knowledge-gap clarity, boundary clarity, conceptual visibility, and methodological formalization. The late layer of results was then read and coded for evidence density and field integration. Apparent maturity was then assessed in terms of rhetorical maturity, administrative maturity, and manuscript/output polishing. Finally, stage sufficiency, false maturity gap, publication maturation gap, and peer-review readiness were interpreted on the basis of these scores. The coding was conducted as a single-coder proof-of-concept exercise. Ambiguous cases were evaluated

conservatively. If a feature was plausible but not clearly visible in the public record, a lower score was assigned. This rule is important because our model diagnoses maturity from documentary traces rather than reconstructing an idealized internal state of the project.

To improve reproducibility, the revised model treats each score as an anchored expert judgment rather than an impressionistic evaluation; coders must cite the documentary artifact that supports each score and must use the conservative rule when evidence is ambiguous. A full validation study would require multiple coders, interrater agreement, adjudication of disagreements, and testing of the relationship between early maturity profiles and later outcomes.

The main limitation of our procedure is that public documents are not identical to the internal life of a project. A large gap may indicate real pseudomaturity, but it may also point to documentary asymmetry. Early project maturation may have occurred without being fully reflected in the public records. The test presented here should therefore be interpreted as a test of coding discipline and discriminatory capacity, not as a final empirical calibration of thresholds.

3. Results

The main result of this paper is the architecture of the Level-Based Scientific Maturation Master Plan. It moves a research project from the state of thematic impulses to the state of a peer-review-resilient scientific contribution. Unlike a conventional work plan, SMMP is organized not by calendar time but by levels of scientific maturity. Each level asks a concrete question: what must become mature before the project can responsibly move further?

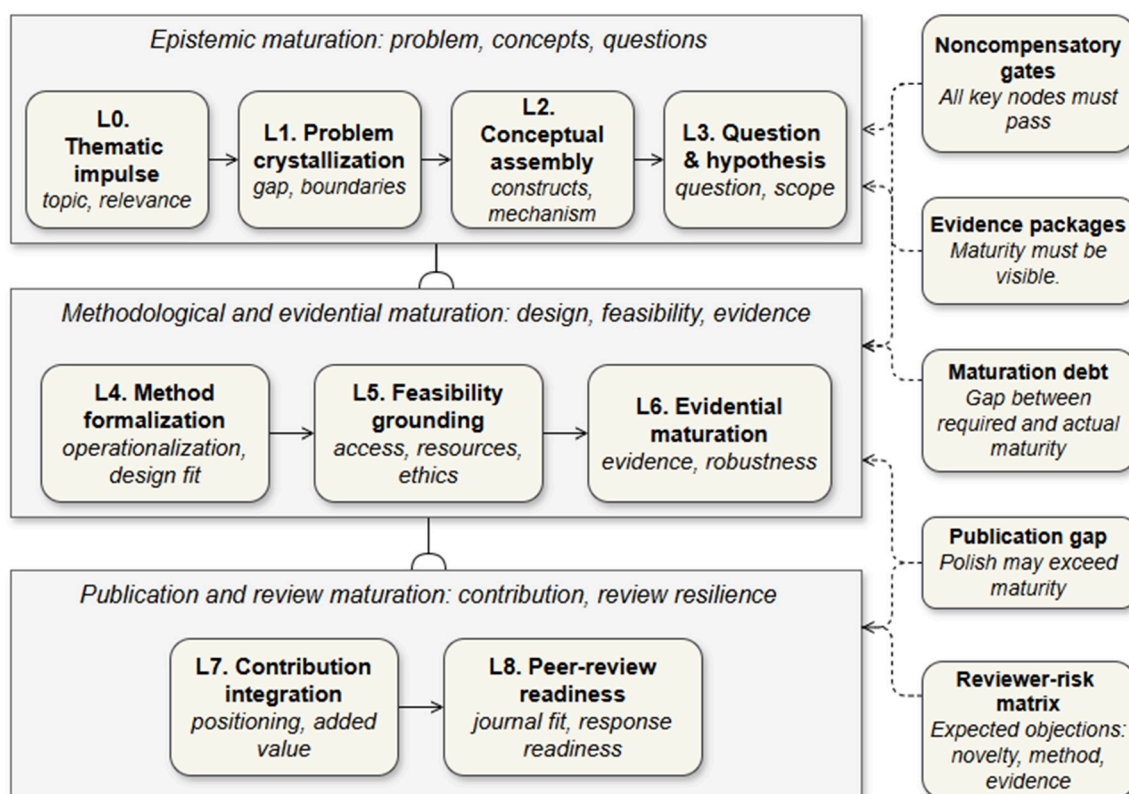
SMMP comprises nine levels. They should not be understood as rigid linear staircases for all disciplines. They are better understood as a disciplined map of maturation. In some projects, levels may partially overlap. However, movement toward publication without passing through these nodes creates the risk of premature publication readiness. This risk is especially high when a topic is socially timely, the literature is abundant, the language is persuasive, and the manuscript can quickly be assembled into formally acceptable text.

To avoid reducing SMMP to a simple sequence of stages, it must be presented as an architecture of maturation. In this architecture, each level performs not only an isolated but also a transitional function: it not only records the current condition of the project but also defines the condition for its further scientific movement. The level-based master plan should therefore be understood not as a calendar ladder but as a system of epistemic, methodological, evidential, and publication filters. It is precisely this system that makes it possible to see where a project is genuinely maturing and where it is merely acquiring external signs of readiness.

The logical relationships among the nine levels can be summarized as three successive transformations. L0–L3 constitute epistemic maturation: a thematic impulse is converted into a problem, the problem receives a conceptual architecture, and that architecture becomes a research question, hypothesis, proposition, or interpretive claim. L4–L6 constitute methodological and evidential maturation: the question is translated into a design, the design is tested against feasibility, and the resulting material becomes disciplined evidence rather than merely available data. L7–L8 constitute publication and review maturation: the result is integrated as a contribution, and the contribution is tested against foreseeable reviewer objections.

The levels are therefore sequential in terms of dependency but not mechanically linear in terms of practice. A project may return from L6 to L4 if evidence reveals a design defect or from L7 to L1 if the claimed contribution reveals that the original problem was too broad. Iteration does not contradict the level architecture; it is the mechanism through which maturation debt is repaid.

Figure 1 visualizes the general logic of SMMP as a movement from thematic impulses to peer-review readiness. The upper track captures the epistemic maturation of the project. A topic must be transformed into a problem. The problem into a conceptual assembly. Conceptual assembly into a stable question or hypothesis. The middle track translates this epistemic construction into a methodological and evidential form. The project must acquire a design, feasibility, and a sufficient evidential base. The lower track shows publication and review maturation. The result must be integrated as a contribution, and the manuscript must become resilient to the expected questions of peer review.



A manuscript may look ready before a project is scientifically mature. SMMP identifies where maturation is still needed before submission.

Figure 1. Master Plan for Level-based Scientific Maturation: From Thematic Impulsiveness to Peer-Review Readiness.

The right-hand side of the figure shows that movement between levels cannot be automatic. It is regulated by noncompensatory gates, packages of evidential artifacts, maturation debt, the publication maturity gap, and a reviewer-risk matrix.

Overall, Figure 1 emphasizes the model's central thesis: a manuscript may appear ready before the project becomes scientifically mature. The external completion of the text is not a sufficient basis for submission if even one mandatory node of the project has not reached threshold maturity.

Figure 1 also connects the internal logic of the project with the future test of peer review. A reviewer's question about novelty usually stems from the immaturity of problem crystallization or contribution integration. A question about the method arises from weak design fit. Questions about claims are linked to insufficient evidential maturation. A question about limitations shows that the project has not passed the discipline of restricting inference. SMMP therefore not only describes stages of project development but also translates potential reviewer objections in advance into diagnosable zones of maturation.

After the visual presentation of the overall architecture, each level of the model must be operationalized. For this purpose, Table 3 presents SMMP not only as a movement scheme but also as a working matrix for managerial and expert decision-making. For each level, the guiding question, mandatory maturity criteria, evidence package, and blocking red flags are specified. The table transforms the figure from a conceptual map into an applicable instrument for presubmission assessment, doctoral supervision, grant development, and internal research governance.

Table 3. Level-Based Scientific Maturation Master Plan.

Level	Guiding Question	Mandatory Maturity Criteria	Evidence Package	Blocking Red Flags
L0. Thematic impulse	Is there only a topic, or already the germ of a knowledge deficit?	Research area; initial relevance; preliminary intuition of the problem	Topic memo; preliminary relevance note	The topic replaces the problem; relevance is presented as novelty
L1. Problem crystallization	Which precise knowledge gap does the project address?	Problem specificity; clear knowledge deficit; boundaries; scientific significance	Problem statement; gap memo; boundary note	Broad topic; blurred boundaries; absence of an autonomous question
L2. Conceptual assembly	With which concepts does the project think?	Definitions; mechanism logic; linkage between constructs; distinction from the field	Conceptual map; definitions sheet; novelty note	Terms decorate the text but do not function analytically
L3. Question and hypothesis stabilization	What exactly is being tested, explained, or reconceptualized?	Research question; hypothesis or proposition; scope of application; expected type of inference	Question/hypothesis protocol; assumptions memo	The question follows convenient data; the hypothesis is not falsifiable
L4. Methodological formalization	Does the method genuinely fit the question?	Operationalization; design fit; analysis plan; falsifiability; transparency	Research protocol; method-design fit sheet; analysis plan	The method is chosen by fashion; the data do not answer the question; there is no analysis plan
L5. Feasibility grounding	Can the project actually be carried out?	Access; resources; competencies; ethics; risk plan	Feasibility dossier; access plan; ethics and risk matrix	The project is feasible only on paper; access, time, or ethical constraints are ignored
L6. Evidential maturation	Is the evidential base sufficient for the claims?	Evidence density; consistency; robustness; limitation discipline; negative-results logic	Evidence dossier; robustness memo; limitations note	Claims are broader than the data; weak robustness checking; limitations are concealed
L7. Contribution integration	What exactly does the project add to the field?	Contribution statement; field positioning; theoretical and/or empirical added value; relation to prior work	Contribution memo; positioning matrix; article skeleton	Results exist, but the contribution is indistinguishable; the article repeats the field
L8. Peer-review and publication readiness	Can the work withstand peer review?	Reviewer-risk resilience; transparent reporting; claim discipline; journal fit; response readiness	Reviewer-risk matrix; submission package; response map	The manuscript appears ready but cannot withstand questions about novelty, method, evidence, and limitations

Through Table 3, we show that publication maturity does not arise at the end of a project. It is assembled across the entire trajectory. If a project has not passed L1, it brings an unclear problem into the manuscript. If it has not passed L2, it results in conceptual overload without an explanatory architecture. If it has not passed L4, it leads to methodological mimicry. If it has not passed L6, it results in inflation. If it has not passed

L7, it brings a text without contribution. L8 is therefore not merely an editorial stage. It is a test of whether all previous levels have matured into a form capable of withstanding peer review.

For what follows, we note that the formal part of the model is needed not for mathematization as an external display but to protect the project from false compensation.

Let l denote the level of the master plan and j denote the mandatory criterion within that level. The score x_{lj} takes values from 0 to 4. Level sufficiency can then be defined as the minimum score across all mandatory criteria of the level:

$$SS_\ell = \min_{j \in M_\ell} x_{\ell j} \quad (1)$$

where

SS_l is the sufficiency of level l ;

x_{lj} is the score of mandatory criterion j at level l ;

M_l is the set of mandatory criteria of level l .

The meaning of Formula (1) is that a level is mature only to the extent that its weakest mandatory element is mature. Formula (1) protects the project from a common illusion: average quality may appear acceptable while one critical node remains immature. In a scientific project, such a node can destroy the entire publication architecture. The choice of the minimum function is therefore a conservative safety rule rather than a claim that all the criteria are equally important in every discipline. Alternative weighted or averaged models may be useful for portfolio description, but they are less appropriate for transition decisions because they allow a critical weakness to be diluted by unrelated strengths. SMMP uses the bottleneck rule only when the criterion has been designated as mandatory for the given project type; nonmandatory or discipline-specific refinements can be reported separately without blocking transition.

The second variable captures apparent maturity. It does not measure science directly. It measures the strength of the external impression. The minimal model includes rhetorical maturity, administrative maturity, and manuscript polishing:

$$AM_\ell = \frac{R_\ell + A_\ell + MP_\ell}{3} \quad (2)$$

According to Formula (2), a project may have a high AM_l if it sounds persuasive, is administratively well organized, and appears textually polished. This does not prove maturity. We therefore define the false maturity gap as follows:

$$FMG_\ell = AM_\ell - SS_\ell \quad (3)$$

If FMG_l is small or negative, external readiness does not outpace internal maturity. If FMG_l increases, the project enters the zone of pseudomaturity. The most dangerous situation occurs when the manuscript already appears ready while the minimum level of maturity remains low.

For the master plan, the indicator of maturation debt is especially important:

$$MD_\ell = \max(0, \theta_\ell - SS_\ell) \quad (4)$$

Here, θ_l is the normative transition threshold, usually equal to 3. Maturation debt points not to an abstract weakness of the project but to the precise deficit that must be corrected. If the problem has $SS = 2$, the debt is 1. If the method has $SS = 1$, the debt is 2. The project does not need general criticism. It needs repayment of the debt at the level where the blocking deficit arises.

Finally, the publication maturation gap (*PMG*) describes the distance between manuscript readiness and project maturity:

$$PMG = \max(0, MP - \min SS_0, SS_1, SS_2, \dots, SS_8) \quad (5)$$

In formula (5), *PMG* shows the extent to which textual readiness has outpaced scientific maturation. It is a central indicator for presubmission review. If the manuscript is stylistically ready while the minimum level of maturity remains low, submission is premature. In such a situation, editing merely masks the deficit. The project must return to the level that has not matured.

The formulas proposed here are not separate mathematical blocks but the operational language of SMMP. We see their task as making expert judgment more reproducible and less dependent on a general impression of the project. If the model were limited to a description of levels, it would remain a conceptual map. If, however, each level receives its own indicator of sufficiency, apparent maturity, maturation debt, and review resilience, the SMMP becomes an instrument of managed diagnosis. The next step is therefore to bring the key indicators of the model together in a single analytical Table 4.

Table 4. Key Indicators of the Scientific Maturation Master Plan.

Indicator	Definition	Formula/Decision Logic	Interpretive Function
Stage/Level Sufficiency (SS)	Actual sufficiency of the level according to mandatory criteria	$SS_l = \min\{x_{lj}\}$	Identifies the weakest mandatory node of maturity
Apparent Maturity (AM)	Externally perceived readiness of the project	$AM_l = (R_l + A_l + MP_l)/3$	Separates the impression of readiness from scientific validity
False Maturity Gap (FMG)	Gap between visible and actual maturity	$FMG_l = AM_l - SS_l$	Diagnoses pseudomaturity and performative readiness
Maturation Debt (MD)	Deficit between actual maturity and the transition threshold	$MD_l = \max(0, \theta_l - SS_l)$	Shows the precise maturation debt at a given level
Maturation Bottleneck	Criterion blocking transition of the entire level	$\operatorname{argmin}\{x_{lj}\}$ within M_l	Translates general weakness into a concrete object of correction
Publication Maturation Gap (PMG)	Gap between manuscript readiness and project maturity	$PMG = \max(0, MP - \min SS)$	Warns against premature manuscript submission
Peer-Review Readiness (PRR)	Capacity of the project to withstand expected reviewer questions	$PRR = \min(\text{problem, concept, method, evidence, limitations, contribution, transparency})$	Assesses not the beauty of the text, but the review resilience of the contribution

Table 4 shows that SMMP is not reducible to a single final index. In our model, we intentionally decompose project maturity into several related but distinct dimensions. Stage/level sufficiency captures the minimum actual sufficiency of a level. The apparent maturity reflects the strength of the external impression. The false maturity gap detects the dangerous outpacing of actual maturity by visible readiness. Maturation Debt translates project weakness into a precise maturation deficit. Maturation Bottleneck indicates the specific blocking criterion. The publication maturation gap shows whether the manuscript has become textually ready before the project has become scientifically mature. Peer-

review readiness connects the internal architecture of the project with the future questions of reviewers.

A weak manuscript usually contains not a single defect but a mismatch among several regimes of readiness. The text may be written, but the contribution may not be integrated. The method may be named but not connected to the question. The data may be presented, but the claims remain broader than the evidence. SMMP therefore requires not a general assessment of “ready/not ready” but the precise identification of the type of maturity that has not reached the threshold. Expert criticism then ceases to be an external judgment of quality and becomes a mechanism of directed project maturation.

We argue that this also changes the very nature of expert criticism. Ordinary criticism is often formulated as a general judgment of quality: the article is weak, the method is insufficient, and the contribution is unclear. Our master plan requires a different form of judgment: At which level did the deficit arise, which mandatory criterion failed to reach the threshold, and which artifact must be created for the transition to become justified? In this way, the model transforms reviewer negativity into a managed maturation trajectory.

Of special importance here is the typology of premature publication readiness. It describes regimes in which a project begins to resemble an article before it becomes a scientific contribution. These are not only individual authorial errors. They are structural regimes of contemporary publication culture. They emerge because the scientific system rewards visible readiness, speed of submission, and the capacity to package results rapidly.

However, indicators alone do not yet show the forms in which premature publication readiness arises. For the practical use of SMMP, it is important not only to measure the maturity gap but also to recognize its type. One project may be rhetorically persuasive but immature at the problem level. Another may have rich literature but no contribution of its own. A third may display methodological vocabulary but lack design fit. The next step in the model is therefore to typologize the main regimes of premature publication readiness, as presented in Table 5.

Table 5 presents premature publication readiness as a family of different pseudomaturity states. Its purpose is not to list typical manuscript weaknesses but to show their project-level origin. Rhetorical readiness indicates immaturity in problem crystallization. Bibliographic readiness without contribution indicates incomplete conceptual assembly. A formal hypothesis without falsifiability shows weakness in question stabilization. Methodological readiness without design fit means that the method has not yet become an internal answer to the research question. The presence of data without evidential discipline points to a gap between evidence and claims. A completed manuscript without contribution integration shows that textual form has outpaced scientific positioning.

We argue that the main virtue of this typology is that it makes criticism correctable. If a reviewer writes that the contribution is unclear, SMMP translates this comment into L7 and requires a contribution memo or positioning matrix. If the comment concerns method, the deficit is localized in L4 and requires the reconstruction of operationalization. If the problem is too broad, the project returns to L1. The table therefore does not merely classify errors. It specifies a route back from publication pseudomaturity to real project maturation.

After the typology is constructed, the next methodological question arises: can the model distinguish such regimes not only in abstract logic but also in documentary biographies of completed research projects? To answer this question, we applied SMMP to two publicly available CORDIS/Horizon cases. This application is intentionally limited. It is not offered as proof that the model is already valid in the strong empirical sense. It is a preliminary documentary demonstration of discriminatory capacity, showing whether the framework can separate different visible trajectories of maturation where ordinary output-based evaluation tends to capture only late productivity.

Table 5. Typology of Premature Publication Readiness.

Type	How It Appears in the Manuscript	Underlying Immaturity	Maturation Action
Rhetorical readiness without problem maturity	The introduction sounds persuasive, but the knowledge gap remains broad or banal	L1 has not reached the threshold	Compress the problem; write a gap memo; define the boundaries
Bibliographic readiness without conceptual contribution	The literature review is rich, but does not generate an independent analytical position	L2 is immature	Build a conceptual map; define construct logic; distinguish the project from the field
Hypothesis readiness without falsifiability	A hypothesis formally exists, but cannot be tested or refuted	L3 is immature	Rewrite the question; limit the scope; clarify the expected evidence
Methodological readiness without design fit	Methods are named, but do not fit the question or the data	L4 is immature	Create a method-design fit sheet; reconstruct the operationalization
Feasibility fiction	The project appears ambitious, but ignores access, time, competencies, and ethics	L5 is immature	Conduct feasibility reduction; create an access and risk dossier
Data readiness without evidential discipline	Data exist, but claims are broader than the evidence	L6 is immature	Prepare a robustness memo; limit the claims; state limitations explicitly
Manuscript readiness without contribution integration	The text is complete, but it remains unclear what it adds to the field	L7 is immature	Write a contribution memo; build a positioning matrix
Submission readiness without peer-review resilience	The article is submitted, but cannot withstand standard reviewer questions	L8 is immature	Create a reviewer-risk matrix; prepare a response map; check journal fit

The demonstrative documentary application of SMMP shows that the model can distinguish not only strong and weak projects but also different types of maturation trajectories. In open CORDIS/Horizon project biographies, the key relation is that between the early architecture of the problem and the late output trial. If one evaluates only publications, deliverables, and outputs, different projects may appear equally strong. The master plan asks another question: was late evidential density the continuation of a mature early architecture, or did it emerge with a less publicly visible early assembly of the project?

As a proof of concept, we compared two open project biographies: IMPACT-EV and ySKILLS (European Commission, n.d.-b, n.d.-c, n.d.-d). IMPACT-EV demonstrates a coherent maturation trajectory. The early formulation of the problem, operational goals, and late results form a relatively continuous line. ySKILLS demonstrates asymmetrically visible maturation. Its late layer of reports, indicators, longitudinal outputs, and publications is very dense, whereas the publicly visible early architecture of the problem and concepts is presented more compactly. This does not mean that the project is weak. It points to a distinction between the maturity of the project itself and the public visibility of its early maturation.

For the demonstration coding, we selected two cases that appear strong at the level of late results but differ in the visibility of their early maturation architecture. This comparison is important for testing the sensitivity of SMMP. If our model merely repeated output-based evaluation, both projects would receive almost identical interpretations as successful and mature. SMMP, however, asks a stricter question: how strongly is the late output trail

linked to the reconstructable early maturity of the problem, conceptual logic, boundaries, and methodological formalization?

Therefore, in Table 6, we present not an assessment of “better/worse” but a preliminary distinction between two types of documentarily visible trajectories.

Table 6. Preliminary Demonstration Coding of Two CORDIS/Horizon Cases.

Indicator	IMPACT-EV	ySKILLS	Interpretive Note
Problem specificity	4	3	IMPACT-EV makes the evaluation problem maximally explicit; ySKILLS is sufficient, but more compact in the early public layer
Knowledge-gap clarity	4	3	Both projects identify a gap, but the early visible differentiation is denser in IMPACT-EV
Boundary clarity	4	4	Both cases set clear boundaries of scope, program context, and target field
Conceptual/mechanism visibility	4	3	The evaluation and impact mechanism is more explicitly visible in IMPACT-EV; ySKILLS is more strongly elaborated in the late layer
Methodological formalization	4	4	Both project biographies show strong design formalization
Evidence density	4	4	Both projects have a strong late output trail
Field integration/projection	4	3	IMPACT-EV demonstrates more explicitly visible integration into evaluation systems; ySKILLS shows strong but more distributed integration
Apparent maturity (AM)	3.3	3.3	Both projects appear highly mature in public presentation and output visibility
Early layer sufficiency (SS)	4.0	3.0	The early public architecture is more fully reconstructable in IMPACT-EV
Preliminary FMG = AM – SS	−0.7	0.3	IMPACT-EV shows no signal of false maturity; ySKILLS shows a moderate signal of documentary asymmetry
Publication maturation gap (PMG)	0.0	0.3	No significant gap is visible in IMPACT-EV; in ySKILLS there is a small gap because late public visibility is stronger than early public visibility
Peer-review readiness (PRR)	4	3	Both projects are strong, but IMPACT-EV is more continuously visible from early formulation to late integration

The data in Table 6 indicate that both projects have high late evidential density but not the same structure of visible maturity. IMPACT-EV demonstrates a maximally reconstructable early architecture. In particular, problem specificity, knowledge gap clarity, boundary clarity, conceptual visibility, and methodological formalization receive high scores. Its preliminary FMG is therefore negative. Visible maturity does not outpace the actual reconstructable maturity of the early layer. In other words, late productivity appears as the continuation of an already dense early formulation.

ySKILLS demonstrates not weakness but a different regime of maturity. Its evidence density and methodological formalization are also high, but early public architecture is presented more compactly. This produces a moderately positive FMG and a small publication maturation gap. This is not a diagnosis of pseudomaturity in the strict sense. It is rather a signal of documentary asymmetry: the late layer of the project is more visible than its early maturation. This interpretation is important because SMMP should not automatically penalize a project for incomplete public documentation. It must distinguish real immaturity from asymmetry of documentary visibility.

Overall, the data in Table 6 demonstrate the minimal discriminatory capacity of the model. SMMP can distinguish coherent maturation from asymmetrically visible maturation even when both projects appear strong according to late outputs. This allows us to move from individual indicators to a more generalized interpretation of project-trajectory types.

To avoid overloading the interpretation with individual scores, we then aggregated the coding results into a more compact trajectory matrix. Table 7 shows which type of maturity is identified in each case, which early maturity signal is visible in public documentation, which late evidence signal is presented in the results, and what final interpretation SMMP provides.

Table 7. Demonstration Interpretation of Maturation Trajectories in Two CORDIS/Horizon Cases.

Case	Visible Trajectory Type	Early Maturity Signal	Late Evidence Signal	SMMP Interpretation
IMPACT-EV	Coherent maturation	Dense early formulation of the problem, goals, and evaluation logic	Case studies, indicators, integrated results, and publication trail	Early architecture and late integration appear coherent
ySKILLS	Asymmetrically visible maturation	Strong, but more compactly presented early problem	Very dense layer of reports, indicators, longitudinal outputs, and articles	Late evidential density is more visible than the early public architecture of maturation

Table 7 shows that the distinction between coherent maturation and asymmetrically visible maturation has not only descriptive but also managerial significance. For research governance, this distinction is fundamental. Expert bodies almost never work with the full internal life of a project. They work with documentary traces. An instrument must therefore distinguish not only pseudomaturity but also documentary asymmetry. Otherwise, a project with a rich late output trail will automatically be regarded as mature across the entire trajectory, which may be analytically mistaken.

These results are limited but important. They do not prove that SMMP is already empirically calibrated. They show that the model produces different diagnoses where ordinary output-based evaluation tends to see only late productivity. This is sufficient to justify the next phase of research: extended documentary validation on a corpus of completed projects with known downstream outcomes, interrater reliability testing, and predictive validity assessment.

The main practical result of SMMP is that it makes the project governable before the manuscript stage. An author, supervisor, research office, or grant team can use the master plan not to punish immaturity but to localize it. The guiding question then changes: not “is this a good project?”, but “which level has not matured, which criterion blocks transition, which artifact must be created, and which reviewer objection will be removed by this maturation action?”

4. Discussion

Taken together, our results show that publication readiness is not scientifically mature. Publication readiness concerns the form of presentation. Scientific maturity concerns the internal organization of knowledge. Peer-review readiness concerns the ability of that organization to withstand external examination. These three regimes often coincide in a strong article, but they should not be analytically conflated. Their conflation creates the illusion that a well-written text is already mature science.

This conclusion must be understood with important qualifications. Scientific maturation is not only a property of the project as an isolated object. It emerges through the interaction of researchers, mentors, collaborators, disciplinary norms, institutional incentives, critical feedback, and tacit judgment (Merton, 1973). The same topic can mature differently in different intellectual communities. SMMP does not deny this emergent character. Rather, it asks which traces of that emergent process must become visible before a manuscript is responsibly submitted.

This distinction also clarifies the relationship between the construct and its indicators. Problem crystallization, conceptual assembly, methodological alignment, evidential discipline, limitation control, and contribution integration do not exhaust scientific maturation. They are observable indicators and necessary traces of maturation in a publication-oriented project. In the language of validity theory, the model therefore requires an interpretive argument: scores are meaningful only if the observed artifacts justify inferences about the underlying construct (Cronbach & Meehl, 1955; Messick, 1995; Kane, 2013). The revised SMMP should thus be read as a structured representation of maturity evidence, not as a complete ontology of scientific creativity.

At its present stage, SMMP should therefore be understood as a conceptual and operational architecture rather than as a fully validated measurement system. The model specifies what should be observed, documented, scored, and discussed when assessing the maturation of a research project, but it does not yet claim psychometric calibration, predictive validity, or universal threshold stability across disciplines. Its current function is diagnostic and formative: it translates diffuse judgments about manuscript weakness into explicit maturation levels, evidence packages, bottlenecks, and reviewer-risk zones. Full measurement use would require subsequent validation, including multicoder application, interrater reliability testing, disciplinary calibration, and empirical examination of whether early maturity profiles predict later publication, review, reproducibility, or implementation outcomes.

In this sense, SMMP radicalizes the question of publication quality. The quality of an article does not begin with the abstract, introduction, methods, or results. It begins earlier: where a topic becomes a problem, concepts become a mechanism, methods become answers to a question, data become evidence, limitations become discipline, and results become contributions. If this trajectory is not passed, the reviewer will almost inevitably discover late symptoms of early immaturity.

This is especially important for research integrity. Scientific irresponsibility does not always take the form of direct fabrication. It often appears more subtly through overstated conclusions, methodological disproportion, selective reporting, weak limitation discipline, imitative novelty, and publication haste. Such practices may not be deliberate deception. However, they produce epistemic harm because they admit into scholarly communication texts whose external readiness exceeds the maturity of their foundations. SMMP responds to this not with moralization but with an architecture of early diagnosis.

In this respect, our model is fully consistent with contemporary reforms in open science and reproducibility. The Transparency and Openness Promotion guidelines emphasize the importance of disclosing data, materials, analysis, and design (Nosek et al., 2015). The

literature on reproducible science links the quality of knowledge to methods, reporting, evaluation, and incentives (Munafò et al., 2017). Registered reports shift part of the evaluative attention to the question and method before the results are known (Chambers, 2013). SMMP moves in the same direction but applies this logic more broadly: not only to an individual study or manuscript but also to the maturation of the project as a whole.

The difference between SMMP and ordinary manuscript checklists is also fundamental. A checklist asks whether elements are present. The master plan asks whether they have matured. A checklist can confirm the presence of a method. SMMP asks about design fit. A checklist can confirm the presence of a limitations section. SMMP asks whether those limitations actually discipline the claims. A checklist can confirm the presence of a literature review. SMMP asks whether an autonomous problem core has emerged from it. The master plan therefore does not compete with reporting guidelines. It operates earlier and more deeply.

SMMP also occupies a distinctive position in relation to readiness models. Readiness levels usually assess whether an object is ready for application, implementation, or further development. SMMP assesses not only readiness but also epistemic assembly. A scientific project may be ready for execution but immature for publication. It may be ready for reporting but immature as a contribution. It may be ready for presentation but immature for peer review. Maturity here is therefore not merely readiness. It is the disciplined transformation of an initial epistemic impulse into a defensible contribution.

The model can be applied in at least five practical regimes. First, we discuss doctoral supervision. A supervisor can use the levels not as a dissertation checklist but as a map for the maturation of problems, methods, and contributions.

Second, we discuss grant development. A proposal can be assessed not only by relevance and work packages but also by the maturity of its problem-concept-method architecture.

Third, we review internal milestones. A research office can identify where a project is stuck in terms of method, evidence, feasibility, or contribution.

Fourth, a presubmission review was performed. An author team can determine whether the manuscript has outpaced its own maturation.

Fifth, we discuss portfolio governance. An organization can track not only the number of projects but also the distribution of maturity bottlenecks across the portfolio.

The concept of maturation debt has particular value. It allows us to move away from the crude language of failure. A project is not necessarily bad. It may have a specific maturation debt. If the debt lies in L1, the project needs problematization. If it lies in L4, it needs methodological reconstruction. If it lies in L6, it needs evidential discipline. If it lies in L7, the contribution must be redefined. This language changes the culture of evaluation. It reduces personal stigmatization and increases the precision of managerial intervention.

Of course, the model also has a certain degree of risk. Any instrument created against simulation can itself become part of simulation. If the SMMP is used in a mechanical bureaucratic form, it becomes a new variety of audit culture. Authors will learn to fill in maturity sheets just as they currently learn to produce external signals of readiness. To avoid this, the master plan must be used as a formative governance tool. Its purpose is not paper compliance but project maturation. An artifact matters only when it changes the scientific structure of the project.

Several limitations should be noted.

These limitations are not peripheral to the model but define its current developmental status: SMMP is proposed here as a structured conceptual protocol and preliminary documentary demonstration, not as a finalized assessment instrument.

First, the model is conceptual and design oriented. Its thresholds are analytically justified but have not yet been calibrated on large corpora.

Second, the 0–4 scale requires interrater reliability testing.

Third, different disciplines define the maturity of evidence, feasibility, and contribution in different ways.

Fourth, public documentary biographies of projects are incomplete and may distort internal maturation trajectories.

Fifth, the predictive validity of the model must be tested through the relationship between early maturity profiles and later outcomes: publications, citations, reproducibility, implementation, policy use, or grant continuation.

Future research should develop in three directions.

First, extended retrospective documentary testing should be performed on 50–100 completed projects from different programs and disciplines.

Second, expert calibration of the scale should be conducted with measurement of interrater agreement.

Third, a digital SMMP toolkit should be developed so that maturation debt, bottlenecks, and reviewer-risk matrices can be used by authors and research offices before submission.

Only after this will the model be able to move from a conceptual protocol to a fully validated governance instrument.

5. Conclusions

This article has argued that a mature publication is not produced by textual polishing alone. It is produced by a project that has passed through visible and defensible scientific maturation: a topic becomes a problem, concepts become an analytical architecture, the method becomes an answer to the question, data become evidence, limitations discipline inference, and results become contribution.

The Level-Based Scientific Maturation Master Plan translates this trajectory into an operational architecture. Its main theoretical contribution is the distinction between publication readiness, scientific maturity, and peer-review readiness. Its practical contribution is a set of levels, mandatory criteria, evidence packages, red flags, noncompensatory gates, and indicators of maturation debt that can be used before submission rather than after rejection.

The revised model also clarifies its own limits. SMMP does not claim to capture the whole tacit, social, and emergent process of scientific maturation. It formalizes the observable traces through which such maturation becomes assessable in research governance, supervision, grant development, and presubmission review. The two CORDIS cases should therefore be read as a proof-of-concept documentary demonstration, not as final validation.

This distinction is essential: the article introduces the architecture of a future reproducible assessment tool, but the tool itself remains to be empirically calibrated, tested for reliability, and validated across disciplinary and institutional contexts.

The broader implications are simple but demanding: scholarly communication should be governed not only by movement toward publication but also by maturation before publication. This is especially important in accelerated and AI-mediated research environments, where coherent text can be produced faster than the scientific object behind it can mature (Bisenbaev, 2026). Future work should calibrate the scale across disciplines, test interrater reliability, and examine whether early maturity profiles predict later research outcomes.

Funding: This research is funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP26199077).

Data Availability Statement: The documentary materials analyzed in this study are publicly available through the European Commission's CORDIS Projects & Results portal. Materials concerning the

IMPACT-EV project are available at <https://cordis.europa.eu/project/id/613202/reporting> (accessed on 8 July 2026), while the ySKILLS project materials are available at <https://cordis.europa.eu/project/id/870612> (accessed on 8 July 2026) and <https://cordis.europa.eu/project/id/870612/results> (accessed on 8 July 2026). The coding framework, assessment scores, and derived results supporting the conclusions of this study are included in the article, particularly in Tables 4 and 5. Further inquiries may be directed to the corresponding author.

Conflicts of Interest: The author declares that there are no competing interests.

Abbreviations

The following abbreviations are used in this manuscript:

SMMP	Scientific Maturation Master Plan
SMP	Scientific Maturation Plan
SS	Stage/Level Sufficiency
AM	Apparent Maturity
FMG	False Maturity Gap
MD	Maturation Debt
PMG	Publication Maturation Gap
PRR	Peer-Review Readiness

References

- Altman, D. G. (1994). The scandal of poor medical research. *BMJ*, 308, 283–284. [CrossRef] [PubMed]
- Alvesson, M., & Sandberg, J. (2011). Generating research questions through problematization. *Academy of Management Review*, 36(2), 247–271. [CrossRef]
- Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (p. 238). McGraw-Hill Education.
- Bisenbaev, A. K. (2026). Scientific artificial intelligence: From a procedural toolkit to cognitive coauthorship. *Philosophies*, 11(1), 12. [CrossRef]
- Bourdieu, P. (2004). *Science of science and reflexivity* (p. 168). University of Chicago Press.
- Chambers, C. D. (2013). Registered reports: A new publishing initiative at Cortex. *Cortex*, 49, 609–610. [CrossRef] [PubMed]
- Collins, H. (2010). *Tacit and explicit knowledge* (p. 200). University of Chicago Press.
- Cooper, R. G. (1990). Stage-gate systems: A new tool for managing new products. *Business Horizons*, 33(3), 44–54. [CrossRef]
- Cooper, R. G. (2008). Perspective: The stage-gate idea-to-launch process—Update, what’s new, and NexGen systems. *Journal of Product Innovation Management*, 25(3), 213–232. [CrossRef]
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281–302. [CrossRef] [PubMed]
- European Commission. (n.d.-a). *CORDIS: Projects & results*. Available online: <https://cordis.europa.eu/projects> (accessed on 10 June 2026).
- European Commission. (n.d.-b). *Evaluating the impact and outcomes of European SSH research (IMPACT-EV): Reporting page*. Available online: <https://cordis.europa.eu/project/id/613202/reporting> (accessed on 10 June 2026).
- European Commission. (n.d.-c). *Youth skills (ySKILLS): Project/fact sheet*. Available online: <https://cordis.europa.eu/project/id/870612> (accessed on 10 June 2026).
- European Commission. (n.d.-d). *Youth skills (ySKILLS): Results page*. Available online: <https://cordis.europa.eu/project/id/870612/results> (accessed on 10 June 2026).
- Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90(3), 891–904. [CrossRef]
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429–431. [CrossRef] [PubMed]
- Hughes, J., Lennon, M., Rogerson, R. J., & Crooks, G. (2021). Scaling digital health innovation: Developing a new service readiness level framework of evidence. *International Journal of Environmental Research and Public Health*, 18(23), 12575. [CrossRef] [PubMed]
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124. [CrossRef] [PubMed]
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. [CrossRef]
- Knar, E. (2025). Integrating scientific readiness and added scientific value: A multiplicative conceptual model for research assessment. *Publications*, 13(4), 57. [CrossRef]
- Law, J. (2004). *After method: Mess in social science research*. Routledge. [CrossRef]

- Mayne, J. (2015). Useful theory of change models. *Canadian Journal of Program Evaluation*, 30(2), 119–142. [\[CrossRef\]](#)
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago Press.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. [\[CrossRef\]](#)
- Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., Coriat, A. M., Foeger, N., & Dirnagl, U. (2020). The Hong Kong principles for assessing researchers: Fostering research integrity. *PLoS Biology*, 18(7), e3000737. [\[CrossRef\]](#) [\[PubMed\]](#)
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behavior*, 1, 0021. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., ... Yarkoni, T. (2015). Promoting an open research culture. *Science*, 348(6242), 1422–1425. [\[CrossRef\]](#) [\[PubMed\]](#)
- Polanyi, M. (1966). *The tacit dimension* (p. 108). Doubleday, ISBN 038506988X.
- Power, M. (1997). *The audit society: Rituals of verification*. Oxford University Press.
- Rheinberger, H.-J. (1997). *Toward a history of epistemic things: Synthesizing proteins in the test tube*. Stanford University Press.
- Shore, C. (2008). Audit culture and illiberal governance: Universities and the politics of accountability. *Anthropological Theory*, 8(3), 278–298. [\[CrossRef\]](#)
- Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore, L. (2021). New framework for developing and evaluating complex interventions: Update of medical research council guidance. *BMJ*, 374, n2061. [\[CrossRef\]](#) [\[PubMed\]](#)
- Smit, J. P., & Hessels, L. K. (2021). The production of scientific and societal value in research evaluation: A review of societal impact assessment methods. *Research Evaluation*, 30(3), 323–335. [\[CrossRef\]](#)
- Sørensen, O. H., Bjørner, J., Holtermann, A., Dyreborg, J., Sørli, J. B., Kristiansen, J., & Nielsen, S. B. (2022). Measuring societal impact of research: Developing and validating an impact instrument for occupational health and safety. *Research Evaluation*, 31(1), 118–131. [\[CrossRef\]](#)
- Strathern, M. (Ed.). (2003). *Audit cultures: Anthropology studies in accountability, ethics and the academy*. Routledge. [\[CrossRef\]](#)
- Tennant, J. P., Dugan, J. M., Graziotin, D., Jacques, D. C., Waldner, F., Mietchen, D., Elkhatab, Y., Collister, L. B., Pikas, C. K., Crick, T., Masuzzo, P., Caravaggi, A., Berg, D. R., Niemeyer, K. E., Ross-Hellauer, T., Mannheimer, S., Rigling, L., Katz, D. S., Greshake Tzovaras, B., ... Colomb, J. (2017). A multidisciplinary perspective on emergent and future innovations in peer review. *F1000Research*, 6, 1151. [\[CrossRef\]](#) [\[PubMed\]](#)
- Thomas, D. A., Nedeva, M., Tirado, M. M., & Jacob, M. (2020). Changing research on research evaluation: A critical literature review to revisit the agenda. *Research Evaluation*, 29(3), 275–288. [\[CrossRef\]](#)
- Watermeyer, R., Derrick, G. E., & Borrás Batalla, M. (2022). Affective auditing: The emotional weight of the research excellence framework. *Research Evaluation*, 31(4), 498–506. [\[CrossRef\]](#)
- Williams, K. (2020). Playing the fields: Theorizing research impact and its assessment. *Research Evaluation*, 29(2), 191–202. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.