

# DEMOCRATIC-TRANSFORMATIVE ASSESSMENT AND EVALUATION IN VETERINARY, ANIMAL, AND HEALTH SCIENCES EDUCATION

An activity-theoretical alternative to authoritarian, individualistic, competitive, and grade-based evaluation

Federico de la Colina Flores<sup>1</sup>, Heriberto Rodríguez Frausto<sup>2</sup>, Tzitzí de la Colina García<sup>3</sup>, Paul Alexis de la Colina García<sup>4</sup>.

<sup>1</sup> Universidad Autónoma de Zacatecas. México. <https://orcid.org/0000-0002-8890-2863>

<sup>2</sup> Universidad Autónoma de Zacatecas. México. <https://orcid.org/0000-0001-9608-5843>

<sup>3</sup> Colegio Edison A.C. Guadalupe, Zac. México. <https://orcid.org/0000-0001-8502-3903>

<sup>4</sup> Colegio Edison A.C. Guadalupe, Zac. México. <https://orcid.org/0009-0000-1666-1469>

## ABSTRACT:

*This article develops a model of assessment (formative) and evaluation (summative) for the education of veterinarians, animal scientists, and health professionals that rejects authoritarianism, individualism, competition, numerical grading, and categorical ranking. The proposal is grounded in cultural-historical activity theory and in a modified activity system in which the classical element of rules is replaced by a persuasive tool. Assessment is therefore reconceived as a collective, dialogic, and accountable process through which participants study the transformations produced by shared work. The article draws normative inspiration from Richard D. Wolff's conception of democracy at work, Grace Blakeley's emphasis on community power and democratic economic reconstruction, and Alfie Kohn's critique of rewards, grades, competition, and homework. It also engages Anna Stetsenko's radical-transformative agency, Wolff-Michael Roth's activity-theoretical account of emotion and work, and Igor Arievitch's agentive activity perspective. The central argument is that assessment and evaluation in professional education should judge neither isolated individuals nor abstract task completion, but the quality of transformative work across the whole activity system: the object, the subject, knowledge, technical tools, persuasive tools, the community, and the division of labor. A democratic-transformative assessment architecture is proposed and applied to four major educational domains: data analysis, workshops, surgical procedures, and design processes. The article concludes that summative decisions can be made without grades through narrative, evidence-based authorization statements that specify demonstrated transformations, current conditions for responsible participation, and the next developmental horizon.*

**Keywords:** Cultural-historical activity theory, democratic assessment, formative assessment, grade-free evaluation, transformative work,

## INTRODUCTION

Assessment and evaluation are some of the most politically consequential dimensions of education. It is where institutions decide what counts as worthwhile learning and development, who is recognized as competent, which forms of conduct are rewarded, and how authority circulates. In veterinary medicine, animal science, and the health professions, this question is especially serious because assessment helps shape those who will make decisions affecting living beings, public health, ecological systems, and professional communities. Yet many prevailing appraisal systems remain tied to forms of judgement that are authoritarian, individualistic, competitive, and narrowly productivist. They rank learners through numerical or categorical grades, reward strategic compliance, and often isolate evaluation from the social and ethical purposes of professional work.

The present article argues that such arrangements are educationally and politically inadequate. They are educationally inadequate because they reduce complex professional activity to fragments that can be scored,

often losing the meaning of the work when it is translated into numbers. They are politically inadequate because they reproduce relations of command, comparison, and scarcity that are incompatible with genuinely collaborative and community-accountable professional formation. These limitations are well recognized in parts of the health professions assessment literature. Programmatic assessment and evaluation have emerged partly in response to the inadequacy of modular, grade-driven systems, emphasizing the value of multiple low-stakes observations, rich feedback, and holistic decision-making over isolated scoring events <sup>[1]</sup>. Narrative approaches have further argued that grades and ratings are a poor fit for the complex, situated performance required in clinical work <sup>[2-3]</sup>.

This article develops a stronger alternative. Building on activity theory as a pedagogical foundation <sup>[4]</sup>, it proposes democratic-transformative assessment and evaluation. In this model, appraisal is not the post hoc measurement of isolated performance. It is a structured inquiry into the transformations produced by collective work. The assessment object is therefore expanded. It includes not only the transformation of the immediate object of activity, such as a dataset, a patient, a surgical field, or a design prototype, but also the transformation of the subject, knowledge, technical tools, persuasive tools, the community, and the division of labor.

This proposal is normatively informed by three lines of critique. From Richard D. Wolff comes the insistence that democratic work requires those who do the work to participate in directing and evaluating it, rather than leaving judgement to a separate managerial layer <sup>[5]</sup>. From Grace Blakeley comes the insistence that durable democratic change depends on building power in communities, workplaces, and local institutions rather than waiting for reform to descend from above <sup>[6]</sup>. From Alfie Kohn comes a sustained critique of grades, rewards, competition, and homework as mechanisms that displace intrinsic engagement and distort the purposes of learning <sup>[7-9]</sup>. The article also draws on Anna Stetsenko's radical-transformative conception of agency in education <sup>[10]</sup>, Wolff-Michael Roth's treatment of emotion, motivation, and identity in activity <sup>[11]</sup>, and Igor Arieievitch's account of mind, learning, and teaching as agentic activity rather than information transfer <sup>[12-13]</sup>.

The thesis defended here is that assessment in veterinary, animal, and health sciences should be organized as a democratic, dialogic, and evidence-rich process oriented to transformative work and future participation rather than to grade production and competitive sorting.

### **ACTIVITY THEORY AND THE PROBLEM OF ASSESSMENT**

We consider activity theory as the pedagogical foundation to design educational interventions grounded in six principles: activity and consciousness, object orientation, the hierarchy of activity, action, and operations, mediation, internalization and externalization, and development <sup>[4]</sup>. These principles make it possible to rethink assessment and evaluation at their roots.

First, if consciousness develops in and through activity, then assessment and evaluation cannot be restricted to the measurement of private possession of knowledge. It must attend to participation in meaningful work. Second, if activities are defined by their objects, then appraisal must ask what object is being transformed and why it matters. Third, if activity is hierarchically organized, then evaluators must distinguish among motives, goal-directed actions, and context-dependent operations. Fourth, if all activity is mediated, then assessment must examine not only subjects but also the tools, symbols, and organizational forms through which work is carried out. Fifth, if internalization and externalization are central developmental processes, then assessment must rely on visible, discussable, revisable externalizations of understanding rather than on hidden inferences from scores alone. Sixth, if development proceeds through contradictions and transformation, then assessment must register change over time and treat tensions as pedagogically productive rather than as mere failure <sup>[14-16]</sup>.

This framework has an important consequence. Assessment and evaluation are not neutral mirrors of learning and development. They are themselves mediational practices that organize motives, relations, and possibilities for development. When they are reduced to grades, they become technologies of comparison and control. When they are reconstructed as collective studies of transformative work, they become parts of the educative process.

### **DEMOCRACY AT WORK, COMMUNITY POWER, AND ANTI-REWARD PEDAGOGY**

The democratic ambition of the present proposal is not metaphorical. Wolff's conception of democracy at work centers on the claim that those who collectively produce the social surplus should also participate in directing the enterprise, rather than having decisions monopolized by owners or managers <sup>[5]</sup>. Transposed into education, the relevant point is that those who collectively perform and support educational work should not be excluded from the processes that define and judge its value. Evaluation should therefore be a matter of shared deliberation among teachers, learners, technical staff, peers, and, where relevant, affected community members.

Blakeley's recent work on democratic reconstruction makes a complementary point. She argues that democratic innovations such as community wealth building cannot succeed unless power is built from below in communities, workplaces, and public institutions <sup>[6]</sup>. Educational assessment and evaluation, if they are to be democratic, must therefore be community-facing and not merely faculty-facing. It must examine whether work contributes to collective capacities and shared infrastructures, not only whether it satisfies internal institutional expectations.

Kohn's work is essential because it identifies how assessment and evaluation are corrupted when they are fused with rewards, competition, and coercive extension of school time. He argues that grades tend to diminish interest, encourage avoidance of intellectual risk, and reduce the quality of thinking by shifting attention from the work itself to performance outcomes <sup>[7]</sup>. His broader critique of rewards is consistent with a large meta-analysis showing that expected, tangible, performance-contingent rewards tend to undermine intrinsic motivation, while informative feedback can enhance it <sup>[17]</sup>. Experimental work by Butler and Nisan likewise found that task-related comments supported intrinsic motivation and performance more effectively than numerical grades <sup>[18]</sup>. Kohn's critique of homework is also pertinent: routine after-hours assignments often extend coercion and busywork into domestic life rather than supporting meaningful learning <sup>[9, 19]</sup>.

Together, these lines of thought suggest that democratic assessment in professional education should minimize ranking, external incentives, and decontextualized take-home compliance. Instead, it should maximize meaningful participation, collective responsibility, and informative feedback.

### **RADICAL-TRANSFORMATIVE AGENCY, EMOTION, AND AGENTIVE ACTIVITY**

The move from grades to transformative assessment also requires a richer account of agency. Stetsenko's radical-transformative stance rejects educational models that socialize learners into passive adaptation to the existing order and instead treats education as participation in inventing the future through present activity <sup>[10]</sup>. This is highly relevant to assessment. If students are to become transformative agents, they cannot be evaluated as mere recipients of standards. They must be assessed as contributors to changing practices, relations, and futures.

Stetsenko and Arieivitch's work on the self in cultural-historical activity theory reinforces this point by reclaiming the unity of the individual and the social and by insisting that agency emerges through participation in socially situated activity rather than through isolated mental properties <sup>[20]</sup>. Arieivitch's later account of an agentive activity perspective goes further, arguing that learning is not information storage and retrieval but the development of mind through consciously arranged and self-regulated activity <sup>[12]</sup>. In educational terms, this means that assessment should focus on how students orient, regulate, and expand activity, not merely on what information they can reproduce.

Roth's contribution is equally important. His analysis of emotion at work argues that emotion, motivation, and identity are not external to activity but constitutive of it <sup>[11]</sup>. Assessment and evaluation systems that ignore the emotional dimension of work often intensify anxiety, conceal uncertainty, and reward defensive compliance. A democratic-transformative model should instead create conditions for trust, dialogic interpretation, and collective responsibility. This aligns with dialogic feedback research, which emphasizes that effective feedback has cognitive, social-affective, and structural dimensions and that trust and mutual reasoning are central to students' uptake of feedback <sup>[21]</sup>.

## **WHY NUMERICAL AND CATEGORICAL GRADES ARE INCOMPATIBLE WITH TRANSFORMATIVE ASSESSMENT**

In complex professional activity, grades are weak informational devices. They compress situated judgements into symbols that often hide the very features that matter most. Hanson, Rosenberg, and Lane argue that in clinical education the main problem with numerical ratings is not merely unreliability but that the meaning of observed performance is lost when it is translated into numbers <sup>[2]</sup>. Ginsburg, van der Vleuten, and Eva show that narrative comments contain significant value and can support reliable judgements when treated seriously rather than as surplus text around ratings <sup>[3]</sup>.

This does not mean that any narrative will do. Judgement still requires structure, transparency, and cumulative evidence. Programmatic assessment addresses this by emphasizing multiple low-stakes observations, rich feedback, and a holistic synthesis process for consequential decisions <sup>[1]</sup>. Yet even programmatic assessment often retains categories or decision rules tied to institutional progression. The present proposal takes a further step. It argues that the pedagogical core of summative judgement should not be a grade at all. It should be a narrative authorization statement describing what transformations have been demonstrated, under what conditions the student can participate responsibly, what support remains necessary, and what the next developmental horizon is.

This does not eliminate accountability. It changes its form. Accountability is preserved through evidence, deliberation, and written public reason, not through reduction to scalar marks.

## **A DEMOCRATIC-TRANSFORMATIVE APPRAISAL ARCHITECTURE**

The model proposed here treats assessment as collective inquiry into transformation. The basic evaluative question is not “How well did this individual perform relative to others?” but “What transformations has this work produced across the activity system, and what forms of further responsibility does that justify?” The relevant transformations are sevenfold.

The first concerns the object. Has the work improved a dataset, a clinical situation, a surgical outcome, a design prototype, a workshop process, a welfare condition, or a community-facing problem? The second concerns the subject. Has the student or team developed greater judgement, initiative, self-regulation, confidence, and ethical responsibility? The third concerns knowledge. Has the work generated clearer definitions, better explanations, stronger interpretations, or more adequate problem framings? The fourth concerns technical tools. Have methods, procedures, code, instruments, or protocols been used, adapted, or improved? The fifth concerns the persuasive tool. Has the work produced better explanations, better consent processes, better argumentation, or more legitimate forms of communication? The sixth concerns the community. Have clients, peers, patients, producers, or publics been better served, included, or empowered? The seventh concerns the division of labor. Has the work reorganized participation, clarified roles, reduced domination, or expanded cooperative capability?

These seven registers supply the frame for both formative and summative judgement. They also reflect the modified activity-system premise that educational work should transform not only objects but the whole network of relations and mediators through which objects are pursued.

## **ASSESSMENT**

Assessment in this model is not a stream of teacher comments appended to products. It is an organized process of collective study while the work is unfolding. It begins with an intention conference in which students, supervisors, and relevant partners clarify the object of the work, the needs it addresses, the expected beneficiaries, the available tools, the likely contradictions, and the criteria by which transformation will be recognized. This stage corresponds to the orienting basis of action emphasized in the Galperinian tradition taken up by Arievidtch and Haenen <sup>[13]</sup>.

The second moment is witnessed work. Evidence is gathered during authentic activity rather than mainly after it. In data analysis this may include code reviews, analytic memos, decision logs, version histories, and team interpretation sessions. In workshops it includes facilitation moves, collective problem framing,

responsiveness to participants, and the quality of shared artifacts. In surgery it includes preparation, asepsis, tissue handling, team communication, welfare-sensitive conduct, and post-procedure reflection. In design it includes framing, iteration, prototype testing, and responsiveness to user and community critique.

The third moment is dialogic debriefing. Following dialogic feedback theory, feedback is treated as a process in which participants think together, rather than as one-way transmission of corrections<sup>[21]</sup>. The focus is on sense-making, contradiction analysis, and planning the next modification of the work. The teacher is not the sole voice. Peers, technicians, patients or clients, community partners, and the student all contribute distinct evidence and interpretations when appropriate. Smeets and colleagues' work on interprofessional assessment is useful here because it shows that complex collaborative competence cannot be judged through single instruments alone and that assessment tasks, assessors, and procedures must be designed as an integrated whole<sup>[22]</sup>.

The fourth moment is revision in shared time. Here Kohn's critique of homework becomes relevant. If revision is displaced into unpaid, isolated, and weakly supported after-hours work, then assessment reproduces inequity and individualizes responsibility. In a democratic-transformative model, the primary site of revision is collective work time with access to peers, tools, and supervision. Out-of-session work is not banned absolutely, but it is reserved for self-chosen inquiry or indispensable continuation of shared activity, not for routine compliance or compensation for poor institutional design.

## **EVALUATION**

Evaluation in this model is a collective synthesis of evidence leading to a narrative authorization decision. It does not culminate in a number, letter, or category. It culminates in a written judgement produced by a small committee or review group composed according to the nature of the work. In some cases, this may include faculty, peers, technical staff, and community representatives. The group reviews a portfolio of transformative evidence accumulated across time. That portfolio includes artifacts, observation notes, students' reflections, peers' statements, community responses, and records of revisions.

The resulting judgement does three things. It identifies demonstrated transformations across the seven registers. It specifies the current conditions under which the student may responsibly undertake similar work in the future. It identifies the next developmental horizon, including what new kinds of activity can now be attempted and what support remains necessary.

The point of this decision is not to sort students competitively. It is to authorize responsible expansion of participation. This is consistent with the most productive aspects of programmatic assessment while avoiding the reintroduction of scalar ranking at the final moment<sup>[1]</sup>. If an external bureaucracy later requires a grade or code for transcript purposes, that conversion should occur after and outside the pedagogical judgement. It should not structure the educational process itself.

## **APPLICATIONS TO MAJOR FORMS OF PROFESSIONAL WORK**

Below are some educational activities to which this model can be applied, along with a description of how they may look in practice: data analysis, workshops, surgical procedures, and design processes.

### **DATA ANALYSIS**

Data analysis in veterinary and health sciences is rarely an individual calculation exercise. It is a collective, tool-mediated process involving data cleaning, documentation, model choice, interpretation, uncertainty communication, and reproducible reporting. Community-led reproducibility initiatives show that training is most effective when it treats reproducibility as a workflow ecology rather than as a set of isolated technical tips<sup>[23]</sup>. Democratic-transformative assessment and evaluation in data analysis should therefore examine whether the work improved not only a result but also the knowledge base, the toolchain, the communicative clarity of claims, and the team's analytic division of labor.

Formative evidence may include analytic plans, code commits, peer code review, notebook annotations, versioned Quarto or R Markdown reports, and reflective memos explaining why certain choices were made and

how uncertainty was handled. Summative judgement should state what kinds of analytic work the learner or team can now undertake responsibly, under what data conditions, with what levels of independence and collaboration. The object of evaluation is not statistical correctness alone. It is the transformation of analytic practice into something more transparent, reproducible, collaborative, and useful.

### **WORKSHOPS**

Workshops are often assessed through attendance or end-point products. This is too thin. A workshop is an activity system with its own object, participation structure, and persuasive tools. Its educational value depends on how well participants are drawn into shared problem-solving, how knowledge is externalized, and how the division of labor supports rather than suppresses participation.

Assessment should therefore attend to facilitation, co-construction of aims, responsiveness to participants, quality of discussion, and the development of shared artifacts. Evaluation should ask what the workshop changed in the participants, in the object of work, in the community's capacity to continue, and in the tools and routines available for future action. Such an approach resonates with the literature on cooperative and interprofessional learning and development, which warns that appraisal focused only on individuals or only on teams can produce either competition or free-riding and recommends integrated designs that address both collaborative process and individual contribution <sup>[22]</sup>.

### **SURGICAL PROCEDURES**

Surgical education is often seen as the paradigmatic domain for checklists and scores. Yet even here the logic of democratic-transformative assessment is strong. Clinical skills laboratories and flipped approaches in veterinary surgery have shown value for building self-efficacy and practical competence through structured preparation and authentic practice <sup>[24]</sup>. But surgical competence is not exhausted by technical metrics. It also includes humane orientation to the patient, communication with the team, anticipation of complications, adaptation under changing conditions, and the responsible use of instruments.

Formative assessment in surgery should therefore include observed procedural work, collective debriefing, review of preparatory planning, and attention to affective regulation and teamwork. Roth's insistence that emotion and identity are constitutive of activity is particularly relevant here <sup>[11]</sup>. Summative judgement should not rank students by surgical elegance. It should authorize forms of participation in surgery under explicitly stated conditions. A narrative authorization might state that a learner can now undertake particular steps of a procedure with specified supports, and that they have shown reliable transformations in planning, aseptic conduct, tissue respect, communication, and reflective revision.

### **DESIGN PROCESSES**

Design processes in veterinary and health contexts include building protocols, communication tools, extension materials, clinical pathways, devices, dashboards, and service models. These are inherently iterative and participatory. Assessment should therefore attend not only to the final artifact but to how the design process reorganized knowledge, tools, relations with users, and the local division of labor.

Formative evidence should include problem framing, stakeholder interviews, draft iterations, critique sessions, and documented changes after feedback. Summative judgement should evaluate whether the design has increased community capability, clarified work, improved communication, or opened a more democratic relation between professionals and those affected by the work. Because design processes often make the persuasive tool especially visible, this domain is well suited to examining how learners justify decisions, negotiate constraints, and produce artifacts that others can understand and trust.

### **GOVERNANCE OF ASSESSMENT AND EVALUATION**

If assessment and evaluation are to be democratic, their governance cannot remain teacher-sovereign. Inspired by Wolff's democracy at work, evaluative authority should be redistributed so that those who do the work and

those affected by it participate in judging its value<sup>[5]</sup>. This does not mean that expertise disappears. It means that expertise is exercised in public, accountable, and shared ways.

A workable governance arrangement is a layered one. Day-to-day formative processes are conducted by those closest to the activity. Periodic summative syntheses are handled by small committees that include relevant voices and that are trained in narrative judgement, evidence review, and bias awareness. Community participation should be appropriate to the work. In some cases, it may mean clients or patient representatives. In others it may mean producer partners, peers, technicians, or local organizations. Blakeley's insistence that democratic change requires power in communities, workplaces, and local institutions is directly relevant here<sup>[6]</sup>. Appraisal should not be a sealed academic function. It should be part of the institution's democratic relation to its publics.

The persuasive tool replaces rules here as well. Assessment criteria, review memos, deliberation templates, public rationales, and authorization letters are the governance artifacts through which judgement is made visible, debatable, and revisable. In a democratic-transformative model, the main question is not whether a criterion has been mechanically satisfied, but whether the reasons offered for judgement are compelling, transparent, and fair.

### CONCLUSION

Assessment and evaluation in veterinary, animal, and health sciences should no longer be organized around numerical or categorical grades that rank isolated individuals and obscure the meaning of work. An activity-theoretical perspective shows that education is the development of participation in collective, object-oriented, mediated activity. On that basis, assessment and evaluation should be reconstructed as democratic and transformative inquiries into what work changes across the whole activity system.

The model proposed here is inspired by democracy at work, community-centered democratic reconstruction, and anti-reward pedagogy. It is strengthened by activity-theoretical accounts of radical-transformative agency, emotion in work, and agentic learning and development. Its central claim is that formative and summative assessment should focus on transformations in the object, the subject, knowledge, technical tools, persuasive tools, the community, and the division of labor. Summative decisions can therefore take the form of narrative authorization statements rather than grades. These statements synthesize evidence, specify current conditions for responsible participation, and open the next developmental horizon.

In professional fields where evaluation shapes future caretakers, clinicians, analysts, designers, and researchers, this is not merely a pedagogical refinement. It is a democratic necessity.

### REFERENCES

- [1] Vleuten C van der, Lindemann I, Schmidt L. *Programmatic assessment: The process, rationale and evidence for modern evaluation approaches in medical education*. *Medical Journal of Australia* 2018; 209:386–8. <https://doi.org/10.5694/mja17.00926>.
- [2] Hanson JL, Rosenberg AA, Lane JL. *Narrative descriptions should replace grades and numerical ratings for clinical performance in medical education in the united states*. *Frontiers in Psychology* 2013; 4:668. <https://doi.org/10.3389/fpsyg.2013.00668>.
- [3] Ginsburg S, Vleuten CPM van der, Eva KW. *The hidden value of narrative comments for assessment: A quantitative reliability analysis of qualitative data*. *Academic Medicine* 2017; 92:1617–21. <https://doi.org/10.1097/ACM.0000000000001669>.
- [4] Rodríguez Frausto H, Colina Flores F de la, Colina García T de la, et al. *Activity theory as a pedagogical foundation for educational models*. *International Journal of Scientific Research and Modern Technology* 2026; 5:72–8. <https://doi.org/10.38124/ijrsmt.v5i1.1139>.
- [5] Wolff RD. *Democracy at work: A cure for capitalism*. Chicago, IL: Haymarket Books; 2012.
- [6] Blakeley G. *Democratized finance is the future*. *Tribune* 2025.
- [7] Kohn A. *The case against grades* n.d.
- [8] Kohn A. *Punished by rewards: The trouble with gold stars, incentive plans, a's, praise, and other bribes*. Boston, MA: Houghton Mifflin; 1993.
- [9] Kohn A. *The homework myth: Why our kids get too much of a bad thing*. New York, NY: Da Capo Press; 2006.
- [10] Stetsenko A. *Radical-transformative agency: Continuities and contrasts with relational agency and implications for education*. *Frontiers in Education* 2019; 4:148. <https://doi.org/10.3389/feduc.2019.00148>.

- [11] Roth W-M. *Emotion at work: A contribution to third-generation cultural-historical activity theory*. *Mind, Culture, and Activity* 2007;14:40–63. <https://doi.org/10.1080/10749030701307705>.
- [12] Arieivitch IM. *Beyond the brain: An agentic activity perspective on mind, development, and learning*. Rotterdam: Sense Publishers; 2017. <https://doi.org/10.1007/978-94-6351-104-9>.
- [13] Arieivitch IM, Haenen JPP. *Connecting sociocultural theory and educational practice: Galperin's approach*. *Educational Psychologist* 2005; 40:155–65. [https://doi.org/10.1207/s15326985ep4003\\_2](https://doi.org/10.1207/s15326985ep4003_2).
- [14] Leontiev AN. *Activity, consciousness, and personality*. Englewood Cliffs, NJ: Prentice-Hall; 1978.
- [15] Engeström Y. *Expansive learning at work: Toward an activity-theoretical reconceptualization*. *Journal of Education and Work* 2001; 14:133–56. <https://doi.org/10.1080/13639080020028747>.
- [16] Roth W-M, Lee Y-J. "Vygotsky's neglected legacy": Cultural-historical activity theory. *Review of Educational Research* 2007; 77:186–232. <https://doi.org/10.3102/0034654306298273>.
- [17] Deci EL, Koestner R, Ryan RM. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 1999; 125:627–68. <https://doi.org/10.1037/0033-2909.125.6.627>.
- [18] Butler R, Nisan M. *Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance*. *Journal of Educational Psychology* 1986; 78:210–6.
- [19] Kohn A. *Rethinking homework* n.d.
- [20] Stetsenko A, Arieivitch IM. *The self in cultural-historical activity theory: Reclaiming the unity of social and individual dimensions of human development*. *Theory & Psychology* 2004; 14:475–503. <https://doi.org/10.1177/0959354304044921>.
- [21] Yang M, Carless D. *The feedback triangle and the enhancement of dialogic feedback processes*. *Teaching in Higher Education* 2013; 18:285–97. <https://doi.org/10.1080/13562517.2012.719154>.
- [22] Smeets HWH, Sluijsmans DMA, Moser A, et al. *Design guidelines for assessing students' interprofessional competencies in healthcare education: A consensus study*. *Perspectives on Medical Education* 2022; 11:316–24. <https://doi.org/10.1007/s40037-022-00728-6>.
- [23] Auer S, Kissel L, Ojo A, et al. *A community-led initiative for training in reproducible research*. *eLife* 2021;10: e64719. <https://doi.org/10.7554/eLife.64719>.
- [24] Decloedt A, Franco D, Martlé V, et al. *Development of surgical competence in veterinary students using a flipped classroom approach*. *Journal of Veterinary Medical Education* 2021; 48:281–8. <https://doi.org/10.3138/jvme.2019-0060>.