Intellectual Breadth and Success: What the Evidence Actually Shows - and What It Means for Education Policy

Executive Summary

- **Breadth and success are reliably correlated, but the causal arrow is mostly unproven.** Across creativity, science, forecasting, and elite sport, people with wider interests and cross-domain experience show better creative and long-run outcomes. But the best-identified studies suggest much of this association is driven by third variables (general intelligence *g*, the personality trait Openness, and curiosity) and by reverse causation (ability and success enabling exploration), not purely by breadth causing success. **Confidence: strong (correlation); weak-to-moderate (causation).**
- **Deliberately broadening your knowledge does NOT reliably raise general intelligence.** The "far transfer" literature - brain training, working-memory training, music, chess - shows near-zero transfer to untrained cognitive abilities once placebo effects and publication bias are controlled. Breadth builds a portfolio of domain-specific skills and knowledge, not a higher *g*. **Confidence: strong.**
- **Where breadth clearly helps is in "wicked," complex, changing environments** - innovation, forecasting, long-run career resilience - and much less in "kind," stable, rule-bound domains (chess, classical music, golf, surgery) where early depth dominates. **Confidence: moderate.**
- **The strongest mechanism evidence is for cross-domain recombination/analogy (novel-but-grounded idea combinations) and network "brokerage,"** followed by career match-quality and flexibility. "Learning to learn" as a general transferable capacity is the weakest-supported mechanism. **Confidence: moderate.**
- **Policy bottom line: favor breadth ("sampling") in childhood and early stages, with depth layered in progressively and specialization deferred rather than rushed.** The controlled evidence for the *timing* claim is genuinely thin, but what exists - late-tracking studies, the Malamud higher-education natural experiment, and athlete-development meta-analysis - points the same way: early specialization buys a short-lived head start that later disappears, while breadth supports better matches and more sustainable long-run excellence. **Confidence: moderate, with explicit uncertainty.**

1. Correlation: What the Data Show, With Effect Sizes

How "breadth" is operationalized. The literature has no single definition; studies use (a) **Openness to Experience** scores (Big Five), (b) **intellectual curiosity** measures (e.g., Typical Intellectual Engagement), (c) **number of arts/crafts avocations or hobbies**, (d) **number of scientific fields or disciplines** spanned (interdisciplinary publication/citation patterns), (e) **number of sports practiced** in athletic development, and (f) **breadth of academic coursework** before specialization. These are related but not interchangeable, which is a core interpretive caveat - an Openness score is not the same construct as a count of hobbies or a measure of interdisciplinary citations.

Personality/disposition correlates of creativity and achievement: - **Openness <-> creativity:** Openness is the most consistent Big Five correlate of creativity. Feist's (1998) foundational meta-analysis found Openness among the strongest personality discriminators of creative scientists and artists; a recent meta-analysis on divergent thinking analyzed 156 effect sizes for Openness, confirming a positive association. Openness also correlates with intelligence at roughly r = .30-.45 depending on the component (e.g., ~.20 fluid, ~.29 verbal/crystallized). **Correlational.** - **Curiosity <-> creativity:** Schutte & Malouff's (2020) meta-analysis of 10 studies (N = 2,692) found a weighted r = .41 (95% CI [.27, .54], p = .0001)

between curiosity and creativity. Importantly, the effect was largely driven by self-report-to-self-report measures (r = .52); when curiosity was correlated against *rated* creativity it fell to r = .16 - a caution that common-method bias inflates the headline number. **Correlational.** - **Curiosity <-> academic performance:** von Stumm, Hell & Chamorro-Premuzic's (2011) "Hungry Mind" meta-analysis (~200 studies, ~50,000 students) found intellectual curiosity is a medium-sized predictor of academic performance; conscientiousness and intellectual curiosity together accounted for 25.7% of variance in academic performance, such that "the additive predictive effect of the personality traits of intellectual curiosity and effort rival that [of] the influence of intelligence." **Correlational.**

Breadth of hobbies and eminence (science): - **Root-Bernstein et al. (2008):** Nobel laureates in science were ~3x more likely to have arts/crafts avocations than typical scientists (Sigma Xi members), and the relationship was monotone up the eminence ladder (Nobel > National Academy/Royal Society > Sigma Xi > general public). The breakdown versus typical scientists is striking: Nobel laureates were **22x more likely to perform, sing, or act**, **12x more likely to write fiction, plays, poetry, or short stories**, **~7x more likely to do designing, painting, drawing, or sculpting**, and **~5x more likely to do crafting, woodworking, mechanics, or glassblowing**. **Correlational, with self-selection and data-comparability caveats the authors themselves flag (e.g., the comparison datasets were collected in different decades and by different methods).**

Breadth of knowledge combinations and scientific impact: - **Uzzi et al. (2013), Science:** Analysis of 17.9 million papers found the highest-impact papers combine **mostly conventional knowledge with a minority of atypical (cross-domain) combinations**. In the authors' words: "Papers of this type were twice as likely to be highly cited works. Novel combinations of prior work are rare, yet teams are 37.7% more likely than solo authors to insert novel combinations into familiar knowledge domains." Pure novelty was NOT rewarded - the sweet spot is breadth grounded in convention. **Correlational/observational, very large sample.** - **Interdisciplinarity is "high-risk, high-reward":** Leahey, Beckman & Stanko (2017) studied ~900 scientists and ~32,000 articles: interdisciplinary research carried a **productivity penalty** (fewer papers; standardized direct effect ~ -0.06) but a **reception benefit** (more citations; ~ +0.04 direct). The productivity penalty slightly dominated, making the total effect on citations modestly negative in the short run. Later work (Nature Communications Physics, 2021; 44,419 UK grants) found interdisciplinary researchers dominate collaboration networks and **outperform in long-term funding** despite short-run citation penalties. **Correlational; consistent "short-run cost, long-run benefit" pattern.**

Forecasting: - **Tetlock's Expert Political Judgment** (2005): 284 experts made roughly 28,000 predictions over 1984-2003. "Foxes" (broad, eclectic, multi-perspective thinkers) reliably beat "hedgehogs" (single-big-idea specialists); notably, hedgehogs performed *less well, especially on long-range forecasts within the domain of their own expertise*. Superforecasters in the later Good Judgment Project scored high on open-mindedness and curiosity. **Correlational/observational.**

Elite sport (the best quantitative breadth data): - **Güllich, Macnamara & Hambrick (2022),** *Perspectives on Psychological Science*, 17(1), 6-29 (DOI 10.1177/1745691620974772): PRISMA-compliant meta-analysis of 51 study reports, 477 effect sizes, drawn from 150 independent samples, **N = 6,096 athletes including 772 world-class (404 adult international medalists; 209 gold medalists)**. Adult **world-class athletes** (vs. national-class) engaged in **more childhood/adolescent multisport practice** (other-sports coach-led practice d = 0.52, 95% CI [0.36, 0.68]; early other-sports practice up to age 15 d = 0.53, [0.36, 0.70]), **started their main sport later** (d = 0.33, [0.14, 0.51]), **accumulated less main-sport practice** (d = -0.27, [-0.46, -0.09]), and **progressed more slowly early** (reached milestones later, d = 0.45, [0.25, 0.64]). Crucially, the pattern **reverses for junior success**: junior elites started earlier, did more main-sport practice, did *less* other-sports practice, and progressed faster. Youth-led play had negligible effects at both levels. **Correlational meta-analysis; the authors explicitly state "we cannot draw causal conclusions."**

Higher-education breadth and earnings: - **Malamud (2010/2011)**, natural experiment exploiting England (early specialization) vs. Scotland (late specialization): early specializers earn slightly more right after college, but late specializers catch up because better field-match reduces costly later switches; **average earnings did not differ significantly** between systems over time. **Quasi-experimental (the strongest design in this subsection).**

Bottom line on correlation: The association between breadth (however measured) and creative/long-run success is real and appears across independent literatures (confidence: strong). Effect sizes are mostly small-to-moderate (r ~ .2-.5 for dispositional measures; d ~ 0.3-0.5 in sport), and almost all are observational.

2. Causation: Evidence for Each Hypothesis, and a Verdict

The four competing hypotheses:

(a) Talent/success -> breadth (reverse causation). Strong support. The behavioral-genetics "Wilson effect" shows heritability of intelligence *rises* with age (Briley & Tucker-Drob, 2013 meta-analysis of 23 longitudinal twin/adoption studies; Haworth et al., 2010, ~11,000 twin pairs), driven by **gene-environment correlation**: as people gain autonomy, they *select and build* environments matching their genetic propensities. Able and successful people have more slack (time, money, security) to explore widely. This means breadth is partly a *consequence* of ability. **Confidence: moderate-to-strong.**

(b) Breadth -> talent/success (the popular "Range" claim). Weak direct causal support. The claim is mostly built on observational and anecdotal patterns (Epstein's *Range*). The athletic meta-analysis (Güllich et al.) and Uzzi's combination data are consistent with breadth helping, but are correlational. There are essentially no randomized trials showing that deliberately broadening interests raises later professional success. **Confidence: weak (as a causal claim).**

(c) Third variable - g, Openness, curiosity. Strong support as a *partial* explanation. Openness and curiosity are substantially heritable and predict both breadth-seeking behavior AND achievement; *g* predicts performance across nearly all domains. When Openness is entered into regressions predicting creativity, the predictive weight of intelligence drops (e.g., beta from .43 to .26 in one study), indicating shared variance. So part of the breadth-success link is a spurious product of common causes. **Confidence: moderate-to-strong.**

(d) Reinforcing feedback loop. Plausible and theoretically favored. Gene-environment correlation models (Tucker-Drob, Briley, Harden) describe exactly this: a propensity (curiosity/Openness/ability) leads to environmental selection (exploring fields), which builds skills/knowledge, which enables more exploration. The rising-heritability finding is the clearest empirical fingerprint of such transactions. **Confidence: moderate.**

Verdict on direction: The honest reading is that **breadth is more a marker and a consequence of underlying ability and disposition than a proven independent cause of success**, with a likely genuine but modest causal contribution in specific (wicked, innovative) contexts operating through recombination and networks. Claims that "generalists triumph" because breadth *makes* people successful run well ahead of the controlled evidence. **Overall confidence in any strong causal breadth->success claim: weak.**

3. Transfer: Does Broadening Raise Intelligence? - Qualified NO

This is the clearest verdict in the whole literature.

- **Sala & Gobet second-order meta-analysis** (Collabra: Psychology, 2019; meta-analysis of meta-analyses, up to k = 233 samples across working-memory training, video games, music, chess,

exergames): **near transfer is real, far transfer is essentially zero.** "When placebo effects and publication bias were controlled for, the overall effect size and true variance equaled zero." The lack of generalization is described as "an invariant of human cognition." **Confidence: strong.**

- **Music training:** Sala & Gobet (2017/2020) multilevel meta-analysis (initially 38 studies, 118 effect sizes, ~3,085 participants; later more than doubled): overall d ~ 0.16, shrinking to ~zero in studies with active controls and randomization. RCTs (e.g., Mehr, Schachner, Katz & Spelke, 2013, Harvard) found no consistent non-musical cognitive benefit. A dissent (Bigand & Tillmann, 2021) re-analyzed and found a small positive effect, so this is not 100% settled, but the weight is clearly toward null. **Confidence: strong (qualified).**
- **n-back/working-memory:** two meta-analyses found small significant transfer to fluid intelligence, but these effects largely vanish under stricter controls; the consensus is near transfer only. **Confidence: moderate.**

Bottom line: Broadening raises **domain-specific knowledge and skill**, and gives you more conceptual building blocks to recombine - but it does **not** raise general cognitive ability (*g*). The value of breadth is not "it makes you smarter in general"; it is "it gives you more, and more diverse, material to think with." **Confidence: strong.**

4. Mechanisms, Ranked by Strength of Evidence

1. **Cross-domain recombination / atypical combination (strongest).** Uzzi et al.'s 17.9-million-paper finding that high impact = conventional core + atypical (cross-field) intrusion is the best large-scale evidence. Analogical-innovation experiments (Kittur et al., PNAS 2019) show distant-domain analogies causally increase the rate of creative idea generation - but with a "too far" boundary (very distant analogies stop helping). This is the one mechanism with experimental backing. **Confidence: moderate-to-strong.**
2. **Network brokerage / structural holes.** Burt (2004, AJS) showed managers whose networks bridge structural holes (i.e., span diverse groups) get better performance reviews, faster promotions, higher pay, and more ideas rated valuable. Breadth of people/fields -> information arbitrage. Largely correlational but replicated across settings. **Confidence: moderate.**
3. **Career match-quality and flexibility/resilience.** Malamud's natural experiment shows late specializers make better field-matches and recover from switches; the athletic data show multidisciplinary backgrounds support more *sustainable* long-run development. Mechanism = sampling reduces costly mismatch and builds adaptable skill sets. **Confidence: moderate.**
4. **"Learning to learn" / meta-learning (weakest).** The far-transfer literature directly undercuts the idea of a general transferable learning capacity built by breadth; Sala & Gobet explicitly list "learning to learn" among theories their data contradict for far transfer. There may be metacognitive/self-regulation benefits, but the strong version is not supported. **Confidence: weak.**

Boundary condition (cross-cutting): Hogarth's **kind vs. wicked learning environments** (Hogarth 2001; Hogarth, Lejarraga & Soyer, 2015) is the key moderator. In **kind** environments (stable rules, fast clear feedback - chess, golf, classical performance, some surgery), deep early specialization and deliberate practice pay off. In **wicked** environments (changing, ambiguous, delayed/misleading feedback - most of modern knowledge work, innovation, leadership, forecasting), breadth and cross-domain transfer matter more. This single distinction reconciles most of the apparently conflicting findings. **Confidence: moderate.**

5. The Deliberate-Practice Context (Why Depth Alone Is Insufficient)

The breadth debate is partly a reaction against the "10,000-hour"/deliberate-practice view (Ericsson). The corrective evidence: - **Macnamara, Hambrick & Oswald (2014),** *Psychological Science* meta-analysis: deliberate practice explained **26% of variance in games, 21% music, 18% sports, 4% education, and <1% professions.** Most variance in performance is explained by other factors (starting age, working-memory capacity, genes). **Confidence: strong**, though Ericsson disputed the definition/methods (the Macnamara-Ericsson exchange is unresolved; a later chess analysis by Burgoyne et al. found >50% of peak-rating variance attributable to practice-related variables, so the exact figure is contested).

The implication: since depth/practice leaves most performance variance unexplained - especially in complex professions - there is room for other factors (ability, disposition, and breadth-enabled recombination) to matter.

6. Policy: Depth vs. Breadth, With an Age/Stage Roadmap

Overall stance: Invest in **breadth early, depth progressively, and defer (don't rush) specialization** - while treating the timing evidence as suggestive rather than settled. The reasoning rests on three convergent bodies of evidence, none individually decisive:

1. **Tracking/streaming timing (strongest policy evidence).** Hanushek & Woessmann (2006) used an international differences-in-differences design (comparing primary vs. secondary outcomes across tracked vs. comprehensive systems; 6 assessments, 18-26 country comparisons): **early tracking increases educational inequality and tends to lower mean performance** - "there does not appear to be any equity-efficiency trade-off." A 75-country replication (2020, 20 years of PISA/PIRLS/TIMSS) confirmed early tracking widens social achievement gaps. **Confidence: moderate-to-strong for "don't track early."**
2. **Higher-education specialization timing.** Malamud: deferring specialization improves field-match without a lasting earnings penalty. Maastricht liberal-arts evidence: broad undergraduate curricula did not impair later specialized master's performance (similar dropout, GPA, thesis grades). **Confidence: moderate.**
3. **Athletic development (analogy, not direct schooling evidence).** Güllich et al.: early sampling/multidisciplinary practice predicts senior world-class status; early specialization predicts junior success that does not sustain ("senior world-class athletes who began their main sport early and specialized are the exception, not the rule"). Exception: sports with very early peak ages (women's gymnastics, figure skating, diving) where early specialization is more defensible. **Confidence: moderate, domain-specific.**

Age/stage roadmap:

- **Early childhood-primary (~ ages 4-11): favor breadth strongly.** Wide exposure (multiple sports, arts, sciences, play). Rationale: heritability of *g* is lower and environment more influential at this stage (Wilson effect); early tracking harms equity with no efficiency gain; sampling builds the diverse base later recombination draws on. Do NOT expect enrichment to raise *g* (far-transfer null) - justify breadth by interest-discovery, foundational skills, and motivation, not IQ gains. **Confidence: moderate.**
- **Lower secondary (~ 12-15): keep options open; delay binding tracking.** This is the developmental window the tracking literature most directly addresses - early selection here entrenches inequality. Begin "T-shaping": a broad base with emerging areas of relative depth. **Confidence: moderate-to-strong (anti-early-tracking).**
- **Upper secondary / early tertiary (~ 16-20): progressive, reversible specialization with sampling.** Malamud's window: let students sample before committing; preserve cheap switching. Match quality matters more than head-start. **Confidence: moderate.**

- **Tertiary and professional (~ 20+): build depth, but maintain deliberate breadth at the edges.** In kind/stable domains, prioritize deep deliberate practice. In wicked domains (innovation, leadership, research, entrepreneurship), cultivate cross-domain exposure and network brokerage; encourage the "conventional core + atypical combination" pattern that predicts high impact. **Confidence: moderate.**

Benchmarks that would change these recommendations: - If RCTs of curriculum breadth showed lasting downstream achievement/earnings costs -> shift earlier toward depth. - If the tracking literature's inequality findings were shown to be confounded by unobserved country differences (a real concern with cross-country designs) -> weaken the anti-early-tracking stance. - In any domain empirically shown to be "kind" with an early peak-performance age -> favor earlier specialization.

7. Open Questions and Weakest Links

- **Causation is the central gap.** Almost no randomized or strongly quasi-experimental evidence isolates breadth->success. The field leans on observational data vulnerable to reverse causation and third-variable confounding.
- **"Breadth" is inconsistently operationalized**, so effect sizes are not strictly comparable across literatures (Openness != hobby count != interdisciplinary citations != number of sports).
- **The depth-vs-breadth *timing* question has thin controlled evidence.** Tracking studies are cross-country (confound-prone); Malamud is a single natural experiment; athletic findings are correlational and domain-specific. We should not overstate confidence in any precise "optimal age."
- **Publication bias and replication problems** are documented in the cognitive-training literature (and the music subset), and the deliberate-practice variance estimates remain actively contested (Macnamara/Hambrick vs. Ericsson).
- **Popular-press overreach:** *Range* and Medici-effect-style narratives often present selected anecdotes and correlational studies as if they established that generalism causes success. They identify a real and useful pattern, but the strong causal/universal claim outruns the controlled evidence.
- **Boundary conditions need mapping:** we lack a clean empirical taxonomy of which real-world domains are "kind" vs. "wicked," which is exactly what would make the policy advice precise.

8. Key References

- Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognitive ability across development. *Psychological Science*, 24(9), 1704-1713. doi:10.1177/0956797613478618
- Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, 110(2), 349-399. doi:10.1086/421787
- Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality and Social Psychology Review*, 2(4), 290-309.
- Güllich, A., Macnamara, B. N., & Hambrick, D. Z. (2022). What makes a champion? Early multidisciplinary practice, not early specialization, predicts world-class performance. *Perspectives on Psychological Science*, 17(1), 6-29. doi:10.1177/1745691620974772
- Hanushek, E. A., & Woessmann, L. (2006). Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries. *Economic Journal*, 116(510), C63-C76.
- Hogarth, R. M., Lejarraga, T., & Soyer, E. (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science*, 24(5), 379-385.

doi:10.1177/0963721415591878

- Kittur, A., et al. (2019). Scaling up analogical innovation with crowds and AI. *PNAS*, 116(6). doi:10.1073/pnas.1807185116
- Leahey, E., Beckman, C. M., & Stanko, T. L. (2017). Prominent but less productive: The impact of interdisciplinarity on scientists' research. *Administrative Science Quarterly*, 62(1). doi:10.1177/0001839216665364
- Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, 25(8), 1608-1618. doi:10.1177/0956797614535810
- Malamud, O. (2010). Breadth versus depth: The timing of specialization in higher education. *Labour*, 24(4), 359-390. doi:10.1111/j.1467-9914.2010.00489.x
- Mehr, S. A., Schachner, A., Katz, R. C., & Spelke, E. S. (2013). Two randomized trials provide no consistent evidence for nonmusical cognitive benefits of brief preschool music enrichment. *PLoS ONE*, 8(12), e82007.
- Root-Bernstein, R., et al. (2008). Arts foster scientific success: Avocations of Nobel, National Academy, Royal Society, and Sigma Xi members. *Journal of Psychology of Science and Technology*, 1(2), 51-63.
- Sala, G., & Gobet, F. (2019). Near and far transfer in cognitive training: A second-order meta-analysis. *Collabra: Psychology*, 5(1), 18. doi:10.1525/collabra.203
- Sala, G., & Gobet, F. (2020). Cognitive and academic benefits of music training with children: A multilevel meta-analysis. *Memory & Cognition*, 48. doi:10.3758/s13421-020-01060-2
- Schutte, N. S., & Malouff, J. M. (2020). A meta-analysis of the relationship between curiosity and creativity. *Journal of Creative Behavior*, 54(4), 940-947. doi:10.1002/jocb.421
- Tetlock, P. E. (2005). *Expert Political Judgment: How Good Is It? How Can We Know?* Princeton University Press.
- Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157), 468-472. doi:10.1126/science.1240474
- von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science*, 6(6), 574-588. doi:10.1177/1745691611421204

Note on evidence quality: This review prioritized meta-analyses and large observational/quasi-experimental studies. The single largest limitation across the entire field is the scarcity of randomized or natural-experimental designs that can isolate the causal effect of intellectual breadth from the abilities and dispositions that produce it. Confidence levels above reflect this: claims about correlation and about the absence of far transfer are well-supported; claims about breadth causing *success, and about the optimal* timing *of depth vs. breadth, rest on thinner, largely correlational foundations and should be held provisionally.*