No Concerns: Conjunctive Hiring and the Dilution of Exceptional-Skill Selection

Jeremy McEntire

Draft, October 2026
doi:10.5281/zenodo.23164408

Abstract

A hiring process is usually judged by how many candidates it rejects. This paper separates that quantity, *selectivity*, from *aim*: how much the process raises the share of hires who have the capability the organization wants. Under a deliberately favorable model, with skills measured without error and every rule applied correctly, we show three things. First, when a candidate must clear a threshold on each of k independent skills, enrichment for exceptional ability on any one skill equals the inverse of that skill's pass rate, and at a fixed hire rate it falls toward 1 as k grows. A ten-box process that rejects 90% of applicants raises the share of top-1% specialists from 1.00% to 1.26%. Second, once each skill is recorded as pass or fail, no rule applied to those marks can do better than the all-pass rule. The record has discarded how far above the threshold the candidate was. Third, whether this matters depends on the work. In a team-value model, selection on strength beats selection on the absence of weakness by a wide margin when payoffs are heavy-tailed, loses most of that margin when usable contribution is capped at the 95th percentile, and is the wrong rule for weakest-link work. We state the conditions under which the argument holds, the evidence against it, and the tests that would refute it.

Keywords: personnel selection, conjunctive rules, multiple hurdles, hiring, team composition, heavy-tailed productivity
JEL: M51, M12, J24, C63

1 The question

Organizations describe their hiring by its selectivity. "We hire one in ten." The number measures how many people were turned away. It does not measure what the process was looking for, or whether it found it.

Call the first quantity selectivity and the second aim. Selectivity is the share of applicants rejected. Aim is enrichment: how much more common a wanted capability is among the people hired than among the people who applied. This paper shows that the two come apart, by how much, and when the gap matters.

The method is to grant the process every advantage. Talent has many dimensions, the interview measures every one of them perfectly, and everyone follows the procedure. The only thing examined is the rule that turns the measurements into a decision.[1]

Much of the ground is old, and we claim none of it. Dawes [1964] analyzed selection on a person's least talent and on their greatest talent. Einhorn [1970] formalized conjunctive and disjunctive decision models. Sah and Stiglitz [1986] showed that requiring every evaluator's approval rejects more good projects, and that accepting any evaluator's approval admits more bad ones. Askira Gelman [2009] showed that adding criteria to a conjunctive rule raises false negatives. Ock and Oswald [2018] found by simulation that compensatory selection generally yields higher expected performance than multiple-hurdle selection.

What this paper adds is narrower:

1. An exact statement of how enrichment for exceptional ability is diluted when selectivity is spread across many mandatory thresholds (Section 3).

2. A ceiling: once evidence is reduced to pass/fail marks, no decision rule over those marks recovers what was lost (Section 3.4).

3. A team-value model in which the best rule depends on the payoff structure of the work, including a cap on usable contribution that, set at the 95th percentile, removes most of the advantage of hiring for strength (Section 5).

The derivations in Section 3 are elementary. We make no claim of priority for them.

The paper does not claim that structured interviews are defective, that specialists outperform generalists, or that judgment should replace rules. The evidence on those points runs the other way, and Section 6 says so.

2 Model

A candidate has k skills, $X = (X_1, \ldots, X_k)$, each expressed as a population percentile. The organization observes all of them exactly. A candidate is *exceptional* on skill i if they are in the top fraction r on it. We use $r = 0.01$ throughout and report two outcomes: the share of hires who are exceptional on one *specified* skill, and the share who are exceptional on *at least one*.

For an acceptance rule A, enrichment on skill i is

$$L_i = \frac{P(E_i \mid A)}{P(E_i)}. \tag{1}$$

$L_i = 1$ is what random hiring delivers. $L_i = 10$ means specialists are ten times as common among hires as among applicants.

We compare four rules, always at the same overall hire rate q:

- **All boxes** (conjunctive): hire if every skill clears its threshold. Equivalent to ranking on $\min_i X_i$. The relaxed form requires at least m of k marks.[2]

[1] The same approach, applied to forced ranking with talent as a single dimension, is in McEntire [2025].

[2] Known in the choice literature as a subset-conjunctive rule; see Kohli and Jedidi [2005].

- **Average** (compensatory): rank on the mean of the standardized skill scores, the usual compensatory composite (each percentile converted to its standard normal score).
- **Maximum** (disjunctive): rank on $\max_i X_i$.
- **Floors plus maximum:** require two skills to clear the 20th percentile, then rank on the best of the rest.

Three different maxima appear in this subject and should not be confused. A maximum over a candidate's *attributes* hires someone exceptional somewhere. A maximum over *evaluators* lets one strong assessment carry a decision. A maximum over *team members* says a team is as good as its best person at each task. Sections 3 and 4 concern the first. Section 5 concerns the third. The venture-capital evidence in Section 6 concerns the second.

3 Selectivity without aim

3.1 The identity

Let skill i be passed by a fraction p_i of applicants, and let the skills be independent. A candidate who must pass every box is accepted with probability

$$q = \prod_{i=1}^{k} p_i. \tag{2}$$

If the exceptional range on skill i lies above its threshold, then

$$P(E_i \mid A) = \frac{r_i \prod_{j \neq i} p_j}{\prod_j p_j} = \frac{r_i}{p_i}, \qquad \text{so} \qquad L_i = \frac{1}{p_i}. \tag{3}$$

Enrichment on a skill is the inverse of that skill's pass rate, and nothing else. Multiplying across skills,

$$\prod_{i=1}^{k} L_i = \frac{1}{q}. \tag{4}$$

For this rule, the hire rate fixes the total enrichment available, and adding mandatory skills divides it. With equal boxes, $p = q^{1/k}$ and

$$L = q^{-1/k} \longrightarrow 1 \quad \text{as } k \to \infty. \tag{5}$$

The identity belongs to the all-boxes rule. It is not a conservation law for selection in general, as Section 3.5 shows.

3.2 What it means in numbers

With ten boxes, the process turns away nine applicants in ten, and each box is one that four in five clear. The share of specialists moves from 1.00% to 1.26%. Tightening the hire rate does little: at a 1% hire rate with ten boxes, each box still passes 63% and enrichment is 1.58.

A middle setting does exist. The overall hire rate can take any value. What cannot be in the middle is the difficulty of the individual boxes. For half of all applicants to pass ten independent boxes, each box must pass 93%.

Table 1: All-boxes selection at a 10% hire rate. Base rate 1.00%.

Mandatory skills	Each box passes	Top-1% share among hires
1	10.00%	10.00%
5	63.10%	1.58%
10	79.43%	1.26%
20	89.13%	1.12%
100	97.72%	1.02%

3.3 The two knobs

The rule has two settings: how hard each box is, and how many must be passed.

Hold the boxes fixed at 79.43% and relax the count (Table 2). Forgiving a box lets more specialists through and lets everyone else through faster. The share of specialists falls back toward the base rate.

Table 2: Relaxing the count with each box fixed at 79.43% pass.

Required	Hire rate	Specified skill	At least one skill
10 of 10	10.00%	1.26%	11.90%
9 of 10	35.89%	1.17%	11.09%
8 of 10	66.06%	1.09%	10.43%
7 of 10	86.89%	1.04%	9.96%

Hold the hire rate at 10% and relax the count, which lets the boxes get harder (Table 3).

Table 3: Relaxing the count with the hire rate fixed at 10%.

Required	Each box passes	Specified skill	At least one skill
10 of 10	79.43%	1.26%	11.90%
9 of 10	66.32%	1.38%	13.00%
8 of 10	55.04%	1.50%	14.05%
5 of 10	26.73%	2.00%	18.38%
1 of 10	1.05%	10.00%	95.62%

Small adjustments buy little. The last row is the maximum rule: one very hard box out of ten. The two knobs can aim the process, but only by turning them all the way to the other end.

3.4 The ceiling

Suppose the decision is made from the marks alone: ten values, each pass or fail, with every box passing 79.43%. A candidate at the 22nd percentile and a candidate at the 99.9th receive the same mark.

A candidate with s passing marks is exceptional somewhere with probability

$$P(E \mid s) = 1 - \left(1 - \frac{r}{p}\right)^{s}, \tag{6}$$

which rises with s and reaches its largest value at $s = k$: **11.90%** in this example. The base rate is 9.56%.

Any decision rule that sees only the marks selects some mixture of mark patterns, and a mixture cannot exceed its best component. So no rule over these marks produces a hired group more than 11.90% exceptional. That includes unanimity, majority vote, weighted counts, and any debate among the people holding the scorecards. The all-pass rule already sits at the ceiling.

The loss does not occur at the vote. It occurs when the form records "passed" and discards the rest. Deliberation afterward works from a record in which adequate and extraordinary are the same entry.

3.5 The rules compared

Table 4: Rules at the same 10% hire rate, ten independent skills.

Rule	**Specified skill**	**At least one skill**
Applicant population	1.00%	9.56%
All boxes	1.26%	11.90%
Average	3.2%	29%
Two floors plus maximum[a]	6.40%	50.70%
Maximum	10.00%	95.62%

[a] On any of the eight non-floor skills. On the two floor skills the share is $r/0.8 = 1.25\%$.

The maximum rule is built to find maxima, so its numbers are an upper benchmark and not a recommendation. Two points stand. The same applicants and the same hire rate give 11.9% or 95.6% depending only on whether the rule takes the minimum or the maximum. And under the maximum rule every one of the ten skills is enriched tenfold at once. The budget of Section 3 constrains conjunctions, not selection.

Averaging sits between. It keeps magnitude, but a spike on one of ten independent skills moves the average by a tenth of its size.

4 Correlated skills

Skills are not independent. We repeat the comparison with correlated skills generated by a Gaussian copula: latent scores $Z \sim N(0, \Sigma_\rho)$, with unit variances and every off-diagonal entry equal to ρ, and percentiles $X_i = \Phi(Z_i)$. The parameter ρ is the correlation of the latent scores. The correlation of the percentiles themselves is $(6/\pi)\arcsin(\rho/2)$: 0.19, 0.48 and 0.79 at $\rho = 0.2$, 0.5 and 0.8. Each rule is recalibrated to a 10% hire rate. Cells in Table 5 give the share of hires in the top 1% on a specified skill. The largest value any rule can reach is 10%.

Table 5: Share of hires in the top 1% on a specified skill, with latent pairwise correlation ρ (Gaussian copula). For two floors plus maximum, the specified skill is a non-floor skill. Final column: box pass rate under the all-boxes rule.

ρ	**All boxes**	**Average**	**Two floors + max**	**Maximum**	**Box pass rate**
0.0	1.3%	3.2%	6.4%	10.0%	79.4%
0.2	3.4%	5.6%	8.5%	10.0%	69.5%
0.5	6.9%	8.4%	9.9%	10.0%	51.7%
0.8	9.6%	9.9%	10.0%	10.0%	30.7%

Correlation rescues the all-boxes rule, and substantially. At $\rho = 0.2$ it delivers about a third of the available enrichment. At 0.5, about two-thirds. At 0.8 the rules are nearly indistinguishable, because a candidate high on one skill is high on all of them, and ten boxes are close to one box asked ten times. The last column shows the same thing from the other side: as correlation rises, the boxes can be made hard again.

Dilution is therefore strongest when the dimensions describe different capabilities and shrinks as they become redundant. It does not disappear quickly: at $\rho = 0.5$ the all-boxes rule still falls 31% short of the attainable share. Under this copula, correlation sets the size of the dilution. In a real applicant population it is one input among several, including how the dimensions behave jointly in the upper tail. The correlation among the dimensions on an organization's own scorecards is the first thing to measure, not the only one.

5 When strength pays

Enrichment for exceptional skill matters only if exceptional skill is worth something to the organization. That depends on how the work combines people. Steiner's classification is the standard one: in additive tasks contributions sum, in conjunctive tasks the weakest member limits the group, and in disjunctive tasks the strongest member can carry it [Steiner, 1972]. Real engineering work contains all three.

5.1 Model

A team fills six seats from sixty applicants. Skills follow the copula of Section 4 with $\rho = 0.2$. Where work can be handed to whoever does it best, the team's value is

$$V(T) = \frac{1}{k} \sum_{j=1}^{k} \max_{i \in T} v(X_{ij}), \tag{7}$$

where v converts a skill percentile into a payoff. We use three payoff functions:

- **Linear:** $v(u) = u$. The 99th percentile is worth about twice the 50th.
- **Heavy-tailed:** $v(u) = (1-u)^{-1/\alpha}$ with $\alpha = 2.5$. Exceptional ability is worth many times ordinary ability.

- **Capped:** the heavy-tailed payoff, cut off at the value of the 95th percentile. The organization cannot use anything beyond that.

A fourth case covers weakest-link work, in which each member's output is limited by their own weakest required skill: value is the average over members of $\min_j X_{ij}$.

To the earlier rules we add one. Ranking candidates by their individual maximum can fill a team with six people who share one strength. What the team objective calls for is marginal contribution. If the team's current best on skill j is M_j, a candidate adds

$$\Delta(c \mid T) = \sum_j \max\bigl(0,\ v(X_{cj}) - M_j\bigr). \tag{8}$$

The **team-gap** rule fills each seat with the candidate who adds the most under this delegable objective, including when the resulting team is evaluated on weakest-link work. It is the only rule told the payoff function v; the others rank on skills alone.

5.2 Results

Table 6: Gain in team value over random hiring.

Payoff	**All boxes**	**Average**	**Maximum**	**Team gap**
Linear	+10%	+12%	+11%	+14%
Heavy-tailed	+59%	+96%	+126%	+137%
Heavy-tailed, capped	+21%	+27%	+24%	+33%
Weakest-link	+219%	+164%	+82%	+100%

Table 6 gives five results.

In the linear and capped cases tested here, differences among the rules are modest. Under linear payoffs all four rules land within a few points of each other, and averaging beats the individual maximum. Boundedness is not what produces this. A payoff of one above the 99th percentile and zero below it is bounded, yet in the same simulation it gives the maximum rule +631% against +214% for the all-boxes rule. What matters is the shape of the payoff and where any cap sits. There is no general theorem that hiring for strength builds better teams.

With heavy-tailed payoffs the rule matters a great deal. Hiring on strength doubles the gain of the all-boxes rule. This is the condition under which Malenko et al. [forthcoming] find that venture investors let a single partner's conviction carry an early-stage investment.

A cap removes most of the advantage. With the same applicants and the same underlying talent, cutting off usable contribution at the 95th percentile drops the maximum rule from +126% to +24%, below averaging. The identity $\max_i \min(X_i, C) = \min(\max_i X_i, C)$ shows why: the cap acts directly on the statistic the rule is trying to raise. Aiming at the team's gap keeps an edge, but a small one. How much the cap removes depends on where it sits. At the 90th percentile the maximum rule gains +11%, no more than the all-boxes rule. At the 99th it gains +67%, ahead of averaging at +59%. At the 99.9th it gains +110%.

For weakest-link work the all-boxes rule is correct. It wins by a wide margin. A role whose output is genuinely limited by its weakest required capability should be filled by a rule that

protects that bottleneck. The bottleneck does the work here, not solitude: someone can do several tasks alone without their output being set by the worst of them.

Part of the team-gap rule's lead is knowledge of the payoff. Aimed at gaps in skill percentiles instead of payoffs, it gains +114% under heavy-tailed payoffs, below the maximum rule, and +32% under the cap, still ahead of the other rules. Its lead under heavy tails depends on knowing how value scales with skill. Its lead under the cap does not.

5.3 A hypothesis about equilibrium

The capped case suggests something beyond the model. Aguinis et al. [2016] find across 229 datasets that a ceiling on productivity is associated with lighter-tailed performance distributions and fewer stars, while autonomy and job complexity are associated with heavier tails. An organization whose processes limit what any individual can contribute receives little return from hiring exceptional people. It is then rational for it to stop looking for them. A process that selects for general adequacy supplies adequate people, who go on to design work around adequate people.

If that loop exists, the all-boxes scorecard and the capped organization are a matched pair. Changing the hiring rule alone would do little. The literature supports the separate links. It does not establish the loop, and we offer it as a hypothesis.

6 Limits and contrary evidence

The model assumes the hard part. We defined "exceptional" before calculating. The organization in the model knows which skills matter and measures them perfectly. Real hiring must first recognize what is valuable, and the mathematics says nothing about how.

Fine-grained scores must be earned. One reason forms record pass or fail is that interviewers cannot rate reliably at finer resolution. The ceiling result says what is lost when magnitude is discarded. It does not show that a given interview can measure magnitude. A process that keeps magnitude needs observations able to carry it.

Unstructured aggregation is not the remedy. Kuncel et al. [2013] find a consistent loss of validity when people combine evidence in their heads instead of by a stated rule. Three limits of that evidence bear on this paper. Its job-performance outcomes are mostly supervisor ratings, which may not register the upper tail of contribution. The mechanical rules that win are linear composites, so the comparison shows that a stated rule beats holistic judgment, not that a linear rule beats one that selects on strength. And its results are averages over many judges, so they describe the typical judge rather than every judge. Their comparison also concerns only how evidence is combined. It does not show that expert judgment is dispensable in deciding what to observe or in interpreting what was observed, which is the recognition problem this model assumes solved. Sackett et al. [2022] rank structured interviews first among single selection procedures. Kahneman and Klein [2009] add that intuition becomes reliable only where regularities are stable and feedback is available. Hiring gives almost no feedback on the people it turns away. Judgment is necessary to identify and interpret relevant

evidence. Its combination should be explicit enough to inspect, challenge and evaluate. The alternative supported here is therefore a declared rule, such as a few justified floors plus one exceptional and relevant strength, applied to evidence that judgment gathered. It is not an interviewer's overall impression.

Champion rules have conditions. The venture-capital result requires heavy-tailed outcomes, evaluators who see different things, and aligned objectives. Where a champion has a private stake, the rule invites advocacy for weak cases [Malenko et al., forthcoming]. A hiring manager who needs the seat filled has a private stake.

The evidence inside a firm cannot settle the question. Rejected candidates produce no outcomes, so hiring policies cannot be compared straightforwardly on a firm's own hires [Lakkaraju et al., 2017]. Li et al. [2025] treat hiring as a problem of exploration for this reason: a firm that only selects profiles with a track record never learns about the rest. Contribution caps raise a separate problem. Measured against the existing capped job, the cap is part of the real economic outcome. It misleads only when that outcome is read as what the person could contribute under a different design of the work. Selective outcomes prevent straightforward comparison of hiring policies, and caps can conceal the value a different work design would permit. Neither, by itself, identifies the value of rejected candidates in uncapped roles.

Validation must support the method of use. U.S. guidance distinguishes screening from ranking: evidence sufficient for pass/fail use may not support ranking, particularly where ranking has greater adverse impact than an appropriate pass/fail method. Cutoff scores should normally be reasonable and consistent with expectations of acceptable proficiency [Uniform Guidelines, 1978]. A rule that ranks on strength should be validated for that use.

Independence is a baseline. Section 4 shows how far correlation weakens the result.

7 Predictions and tests

The argument predicts the following.

1. Where scorecards record magnitude and their dimensions are weakly correlated, a hire's best score will predict upper-tail contribution better than their lowest score or their mean.

2. The return to strength-based selection rises with the autonomy of the role and is near zero where work is fully prescribed.

3. Moving a person from a capped role to an uncapped one raises the realized value of exactly the strength their interview showed.

4. Teams assembled on marginal contribution outperform teams assembled on individual scores for disjunctive work, and not for weakest-link work.

Two conditions would limit where the argument applies. They would make the results for independent skills a poorer approximation for a given population, not contradict them.

- Scorecard dimensions are strongly correlated. Even at $\rho = 0.5$ the all-boxes rule falls 31% short of the attainable share, so whether the remaining dilution matters depends on what an exceptional hire is worth, not on the correlation alone.
- The payoff to exceptional skill is close to linear, or capped low by the nature of the work. In the cases tested in Section 5, differences among the rules are then modest.

The hypotheses would be rejected if the predictions above fail where their conditions hold, and most directly by a prospective trial, with evidence collection held fixed and the combining rule randomized at matched hire rates, that shows no advantage for floors plus strength in high-autonomy teams.

A retrospective study of archived scorecards that record magnitude can examine the first prediction, by comparing the minimum, mean and maximum scores within the hired sample. Under an all-pass rule, "passed every box" is the same for every hire and predicts nothing, so it cannot serve as the comparison. Such a study is exploratory: it describes relationships inside the accepted region and cannot identify which rule would have selected best. That requires hiring some people the current rule rejects, which is why the prospective trial is the test of policy.

8 Implications

The findings support a narrow set of changes, and no more.

- Structure the observation. Keep the questions, the work samples and the independent write-ups.
- Record how well, not only whether. A mark that cannot distinguish adequate from extraordinary cannot be used to find the extraordinary.
- Make each floor justify itself by what breaks without it. Keep the ones that survive.
- Combine by a stated rule chosen for the work: every box for a role whose output is limited by its weakest required capability, strength at the team's gap where work can be delegated.
- Write down what the hire is expected to contribute and check it later. Accountability for the reasoning improved interview validity in experiment. Accountability for the outcome reduced it [Brtek and Motowidlo, 2002].

Selection on minimum adequacy, on average quality and on exceptional strength are different objectives. A process should choose among them by the structure of the work. Its rejection rate does not show that it chose.

Code

All figures come from `no_concerns_simulation.py`. The code, its tests and its saved output are archived as version 1.0.0 at doi:10.5281/zenodo.23164406; `make test` reruns the calculations and checks every number quoted in the paper. Section 3 is exact. Section 4 is exact by one-dimensional quadrature over the common factor, checked by Monte Carlo. Section 5 uses Monte Carlo with a fixed seed (two hundred thousand teams). A rule's gain is the mean team value under that rule divided by the mean team value under random hiring from the same applicants, minus one. Standard errors of the gains in Table 6, by the delta method over teams, are at most 0.3 percentage points, and the figures vary by at most one point across seeds. `exact_correlated.py` computes Section 4; `robustness.py` reports seed variation, standard errors and sensitivity to the tail index, cap level, correlation and team size.

References

Herman Aguinis, Ernest O'Boyle, Erik Gonzalez-Mulé, and Harry Joo. Cumulative advantage: Conductors and insulators of heavy-tailed productivity distributions and productivity stars. *Personnel Psychology*, 69:3–66, 2016. URL https://www.hermanaguinis.com/pdf/PPsych2016a.pdf.

Irit Askira Gelman. The asymmetric nature of decision errors in multi-criteria, satisficing decisions. In *AMCIS 2009 Proceedings*, 2009. URL https://aisel.aisnet.org/amcis2009/369. Paper 369.

Mary D. Brtek and Stephan J. Motowidlo. Effects of procedure and outcome accountability on interview validity. *Journal of Applied Psychology*, 87(1):185–191, 2002.

Robyn M. Dawes. Social selection based on multidimensional criteria. *Journal of Abnormal and Social Psychology*, 68(1):104–109, 1964.

Hillel J. Einhorn. The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin*, 73:221–230, 1970.

Daniel Kahneman and Gary Klein. Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64(6):515–526, 2009.

Rajeev Kohli and Kamel Jedidi. Probabilistic subset conjunction. *Psychometrika*, 70(4): 737–757, 2005. doi: 10.1007/s11336-002-0978-8. URL https://business.columbia.edu/sites/default/files-efs/pubfiles/1182/1182.pdf.

Nathan R. Kuncel, David M. Klieger, Brian S. Connelly, and Deniz S. Ones. Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology*, 98(6):1060–1072, 2013.

Himabindu Lakkaraju, Jon Kleinberg, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. The selective labels problem: Evaluating algorithmic predictions in the presence of unobservables. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017.

Danielle Li, Lindsey Raymond, and Peter Bergman. Hiring as exploration. *Review of Economic Studies*, 2025. doi: 10.1093/restud/rdaf040.

Andrey Malenko, Ramana Nanda, Matthew Rhodes-Kropf, and Savitar Sundaresan. Catching outliers: Committee voting and the limits of consensus when financing innovation. *Journal of Finance*, forthcoming. URL https://www.hbs.edu/ris/Publication%20Files/21-131_c118f5f9-e553-4110-ab71-d29fa5bf0439.pdf. Earlier version: Harvard Business School Working Paper 21-131, 2021.

Jeremy McEntire. Tournament-based performance evaluation and systematic misallocation. arXiv:2512.06583, 2025. URL https://arxiv.org/html/2512.06583v1.

Jisoo Ock and Frederick L. Oswald. The utility of personnel selection decisions: Comparing compensatory and multiple-hurdle selection models. *Journal of Personnel Psychology*, 17 (4):172–182, 2018. doi: 10.1027/1866-5888/a000205.

Paul R. Sackett, Charlene Zhang, Christopher M. Berry, and Filip Lievens. Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, 107(11):2040–2068, 2022.

Raaj Kumar Sah and Joseph E. Stiglitz. The architecture of economic systems: Hierarchies and polyarchies. *American Economic Review*, 76(4):716–727, 1986.

Ivan D. Steiner. *Group Process and Productivity*. Academic Press, 1972.

Uniform Guidelines. Uniform guidelines on employee selection procedures, 29 CFR 1607.5(G)–(H), 1978. URL https://ecfr.gov/current/title-29/subtitle-B/chapter-XIV/part-1607/section-1607.5.