Received: 31 March 2017 | Revised: 18 June 2018 | Accepted: 3 October 2018
DOI: 10.1111/jems.12295

ORIGINAL ARTICLE

WILEY Journal of Economics & Management Strategy

Optimal team composition for tool-based problem solving

Jonathan Bendor[1] | **Scott E. Page**[2]

[1]Department of Political Economy, Stanford University Graduate School of Business, Stanford, California

[2]Center for the Study of Complex Systems, Departments of Political Science and Economics, University of Michigan, Ann Arbor, Michigan, Santa Fe Institute

Correspondence
Jonathan Bendor, Department of Political Economy, Stanford University Graduate School of Business, Stanford, CA 94305.
Email: jonathan.bendor@stanford.edu

Funding information
Army Office of Research, Grant/Award Number: W911NF1010379

Abstract

In this paper, we construct a framework for modeling teams of agents who apply techniques or procedures (tools) to solve problems. In our framework, tools differ in their likelihood of solving the problem at hand; agents, who may be of different types, vary in their skill at using tools. We establish baseline hiring rules when a manager can dictate tool choice and then derive results for strategic tool choice by team members. We highlight three main findings: First, that cognitively diverse teams are more likely to solve problems in both settings. Second, that teams consisting of types that master diverse tools have an indirect strategic advantage because tool diversity facilitates coordination. Third, that strategic tool choice creates counterintuitive optimal hiring practices. For example, optimal teams may exclude the highest ability types and can include dominated types. In addition, optimal groups need not increase setwise. Our framework extends to cover teamwork on decomposable problems, to cases where individuals apply multiple tools, and to teams facing a flow or set of problems.

1 | INTRODUCTION

Over the past 35 years the number of people used as cognitive, nonroutine workers has doubled to 60 million. The job classifications for this category of workers include managing, designing, performing basic research, investing, strategic consulting, engineering, and providing legal advice and medical care.[1] The tasks carried out by cognitive, nonroutine workers consist in large part of solving problems. Biomedical researchers isolate molecules. Financial analysts build portfolios. Consultants develop reorganization plans. Engineers design batteries. Equally relevant to our analysis, most problem solving is now done in teams. Therefore, the study of problem solving is also the study of teams and teamwork.[2]

In this paper, we construct a framework for studying team performance on problem solving. Specifically, we analyze the proability of success at problem solving given team composition and derive optimal hiring rules. Our framework assumes problem solvers who possess skill of varying degrees at applying tools. Better problem solvers know more tools and are more adept at applying them. Given that problem solving is done primarily in teams, our framework focuses on teamwork. We assume that a team either succeeds or fails depending upon whether any member of the group finds a solution. We do not constrain the definition of a problem, so the binary nature of outcomes—success or failure—does not greatly limit the scope of our framework. Any task that requires constraint satisfaction (with or without an optimization criterion), such as reducing the rate of defects in a manufacturing process by a certain percentage or designing an internal combustion engine that exceeds environmental standards while maintaining a torque profile, is admissible.

The difficulty of developing an optimal hiring policy stems from a lack of separability. Ability corresponds to the probability of solving the problem, which in turn implies a facility with potentially successful tools. Ability fails as a proxy for a person's added value to a group because the group may already contain people who possess the high-ability person's tools. As a rule, the best team of problem solvers need not consist of the most able individuals (Hong & Page, 2004; Page, 2008; Marcolino, Jiang, & Tambe, 2013). In fact, for some classes of problems no measure applied to individuals determines optimal team composition (Kleinberg & Raghu, 2018). The best person to add to a group will be the one most likely to apply a tool that is both novel and effective. Therefore, optimal hiring depends on the group composition (Prat, 2002).[3] If firms had little choice in whom to hire, this optimal hiring problem would not be relevant. That is decidedly not the typical case. Alphabet, the parent company of Google, annually receives upwards of three million applicants. Leading financial services companies and consulting companies receive over a quarter of a million applicants. People analytics, the use of data and models to make hiring decisions, has now become a standard tool (Bock, 2015; Conner, 1991; Conner & Prahalad, 1996; Demsetz, 1988; Powell & Snellman, 2004).

Our analysis consist of two main parts. We first derive benchmark results where the manager can assign tools to workers. We find that the manager wants worker types who are proficient with distinct tools. Once we have a firm grasp of the sometimes subtle relations between individual ability, team diversity, and group effectiveness, this decision theoretic result is not surprising. We next consider the more complex strategic context in which the manager first chooses workers who then autonomously choose tools. A worker's payoff depends on some combination of group success and individual credit. The manager's optimal rule takes into account that individual credit matters to workers. This results in what at first appear to be counterintuitive hiring practices but which upon reflection are rational because they prevent doubling up on tools.

Our approach complements the traditional models of team performance that take membership as fixed and focus on moral hazard (Holmstrom, 1982) or both moral hazard and adverse selection (McAfee & McMillan, 1991) in which an individual's contribution to the team depends on effort and ability. Their focus on shirking may be more appropriate for production and the provision of services than for problem solving, where success can produce reputational rents. To the extent that incentive problems matter in problem solving contexts, we believe they can be handled separately from the competition arising from team composition.

Within the framework, we uncover direct benefits from cognitive diversity: Teams with diverse tools are more likely to find solutions to problems. In addition, we find that tool diversity becomes more important both when average skill increases and group size increases because in such situations existing tools are likely to have been applied correctly. We also find strategic benefits from diverse teams. They have fewer coordination failures owing to lack of overlap in their toolkits. We also find strategic advantages to hiring less able workers. Knowing fewer tools can reduce the incentive to choose the wrong tool, for example, selecting one already tried by teammates. These findings echo a variety of strength-through-weakness results in game theory. Last, extending the model to allow for partial solutions amplifies the benefits of applying more diverse tools. Thus, our stark model is conservative: It stacks the deck toward ability and simple rules and away from diversity and complex rules.

Our framework takes an agent's facility with tools as exogenous. Workers can choose which tool to apply given their set of tools. They cannot choose to become an expert at a new tool. Robust empirical evidence shows that proficiency with a tool, particularly one that produces economic value, requires specialized training and hundreds if not thousands of hours of practice (Ericsson, Krampe, & Tesch-Romer, 1993; Feltovich, Prietula, & Ericsson, 2006). Consequently, groups that have tool diversity must have *team* diversity, that is, people trained in different methods.

To provide context for our results, we refer throughout to five hiring regularities that hold for a constant elasticity of substitution (CES) production function given equal market wages: Higher ability types should be hired earlier (competency ordering) and in greater number (competency loading); all previous hires remain in the optimal group as group size increases (monotonicity); a type that does not have the most skill at some task will not be hired (no dominated hires); eventually all undominated types will be hired (asymptotic diversity). These regularities bear repeating in more colloquial phrasing to reinforce their normalcy: Hire the most talented first, hire more of the more talented, do not discard talent, do not hire ineffective workers, and increase worker diversity as the firm grows.[4] These regularity properties hold for a variety of production functions used in decision- and team-theoretic models (Marshack & Radner, 1972).

As we will show, optimal hiring for problem-solving workers can violate each of these regularities. Moreover, the violations are not knife-edge cases; they arise under a range of reasonable assumptions. While many of the violations result from incentive issues created by strategic tool selection, some arise from properties of problem solving. For example, competency ordering and competency loading can be violated even when the manager can assign tools to workers. The violations arise when the toolkits of different types of workers *overlap*, that is, have some tools in common. Without overlap, optimal hiring would be straightforward.

The remainder of the paper has five parts. We begin with informal and formal descriptions of our framework. We next derive benchmark results for the centralized structure where the manager selects worker types as well as the tools deployed. We then turn to the decentralized system in which problem solvers strategically choose tools. In the penultimate section, we consider three natural extensions. We conclude by summarizing our key results and note some implications.

2 | A FRAMEWORK FOR PROBLEM-SOLVING GROUPS

Our framework assumes a problem, a set of tools that can be applied to that problem, and problem solvers with varying levels of facility with subsets of the tools. A tool can be interpreted as an approach, technique, or method. Given a problem, each tool has a fixed, independent probability of finding a solution. We further assume that a tool's success can be observed. After a tool has been applied, the team knows whether or not the team solved the problem.[5]

Two examples reveal the range of contexts to which the model can apply. First, tools can represent techniques for finding solutions. For example, biochemists rely on four common techniques for DNA sequencing: Enzyme-based methods such as the Sanger method, chemical methods, single-molecule sequencing using fluorescents, and real-time detection of pyrophosphates (França, Carrilho, & Kist, 2002). In any one instance, each of these techniques may or may not work; a priori, some techniques may be more likely to sequence the gene. Those could be captured as probabilities of success. In addition, a given researcher has some level of facility with a tool. An experienced biochemist may be more likely to apply the Sanger method correctly than a new hire.

Second, tools can correspond to types of technologies. There exist a variety of battery technologies including lithium ion, lithium sulfur, aluminum ion, and nickel-metal hydride. Each technology has a storage capacity, recharge cycle, cost per cycle, self-discharge rate, maximum voltage, weight, temperature, and cost. The battery for a product, whether it be a smartphone, rechargeable vacuum, or an electric car, must meet certain requirements. An engineer might be an expert at one of these technologies. The model does not apply as cleanly to problems such as writing a screenplay in which no problem gets solved even though there does exist a final outcome that either succeeds or fails.[6] Nor does it apply to problems that are embedded in the production process, most of which are known to be solvable by less-skilled workers at the bottom of an organizational hierarchy (Garicano, 2000; Garicano & Rossi-Hansberg, 2006).

Given a problem, we partition problem solvers into types based on their toolkits and their facility with their tools. We define two problem solvers to be of the same type if they have the same toolkits and apply each tool correctly with the same probabilities (Newell & Simon, 1972).[7] One type might correspond to recent college graduates who worked in research labs developing lithium ion batteries, and another type might consist of engineers who have spent a decade working with solid state battery technologies. In this example, the first type of worker might have less facility with a better tool.

We implicitly assume a time-constraint by requiring each type to choose a single tool. She cannot start over if she misapplies a tool or it fails to solve the problem. This assumption holds provided applying a tool requires time or resources. It allows us to model a individual's choice of a single tool as an action in a game.[8] The person solves the problem if she applies the tool correctly and the tool works, that is, succeeds in solving the problem.

As noted, problem solvers work as part of a group or team. The team solves the problem if and only if at least one of its members finds a solution. The ability of a type is defined as the expected probability that that type solves the problem using its best tool, that is, the one that, of all the type's tools, is most likely to solve the problem. The ability of a team equals the expected probability that at least one person in the group solves the problem.

Because a team's ability depends on the number of distinct and correctly implemented tools, individual ability is only a crude proxy of value to a team. This feature of collective problem solving provides intuition for why hiring the best type (competency loading) fails under mild assumptions. Once a tool has been used correctly, the team does not need more people who will choose that tool. In contrast, a less able type may be needed in abundance if low-ability stems from the difficulty of deploying a tool. The manager of a large team might want many of that inferior type to increase the probability that *somebody* correctly applies that type's unique tool.

Many of our results depend on whether tools are assigned or chosen strategically. This in turn is often a product of the environment. In some contexts, a manager has sufficient knowledge, authority, and monitoring capability to assign tools. This might be true in a research lab where the lead researcher knows the tools and assigns them to students. Assuming that the requisite conditions (e.g., knowledge) hold, this kind of centralized regime is a best-case scenario as it avoids strategic coordination problems and incentive effects. In other situations, however, managers lack tool-specific knowledge, cannot verify choice of tool, or verification is prohibitively costly. The manager can then select personnel but not the tools that team members use. Here, once the manager selects problem solvers, she has no authority over their tools. The workers will select tools strategically: The set of chosen options will be a Nash equilibrium of a game.

To construct the game over tool choices, we define payoff functions in terms of a collective payoff if the team solves the problem and an individual payoff for those team members whose selected tools solve the problem. The private benefit could come in the form of money, reputation in the organization, or status within the team. The collective component, which could come from profit sharing, aligns problem solvers' incentives with those of the manager. Even then, however, strategic tool choice introduces coordination problems. Two problem solvers of the same type may have an incentive to choose their best tool rather than differentiating for the good of the organization. This distortion produces violations of both monotonicity and competency ordering. These can sometimes be overcome if individuals choose tools sequentially. That strategic tool choice introduces coordination failure, order effects, and misaligned incentives should come as no surprise. Clearly, the centralized process should outperform the strategic one. That said, centralized tool choice may require unacceptable levels of authority or infeasible levels of knowledge and observation, thus running afoul of Hayekian critiques of centralization (March & Simon, 1958; Nickerson & Zenger, 2004).[9]

2.1 | The formal model and definitions

We now introduce the formal model. We assume the game form and all parameters, including the number of types, the number of tools, and the probability that each tool can solve the problem, are common knowledge.

The set of **tools** $\mathcal{K} = \{1, 2, 3, \dots K\}$, *denoted by* t_k.

The set of **types** *of problems solvers* $\mathcal{K} = \{1, 2, 3, \dots M\}$, *denoted by* s_i.

The set of tools in type s_i's **toolkit**, $T(s_i) = \{t_k : \text{type } s_i \text{ can use } t_k\}$.

A **group**, G, *is a finite nonempty set of types that can contain multiple members of a type.*

A **problem**, $X = (P, H)$, *where* $p(t_j)$ *denotes the probability that a correctly applied tool* t_j *solves the problem, with* $p(t_j) \in (0, 1)$ *for all* j, *and where* $h_{s_i}(t_j)$, *type* s_i's **skill** *with tool* t_j *on problem* X, *denotes the probability that* s_i *correctly applies* t_j *on problem* X, *with* $h_{s_i}(t_j) \in (0, 1]$ *for all* i *and* j. *Further,* $p(t_j) \neq p(t_l)$ *for any* $j \neq l$, and we order tools by their success probabilities: $p(t_j) > p(t_l) \Longleftrightarrow j < l$.

Two assumptions merit emphasis: Any tool might solve the problem but none is guaranteed to do so, and a tool's success probability, p, does not depend on the type who applies it. The second assumption implies that if a tool does not solve the problem when applied correctly by one problem solver, then it will not solve the problem when tried by anyone else. Thus, if a physicist fails to solve a problem using spectral analysis, then a mathematician who tries that method will also fail. (Types may differ, however, in their skills at applying tools, i.e., in the h's.) Note that we allow for the possibility that a type can with certainty apply some tool correctly ($h = 1$). We can rank a type's tools by their overall success probability: The product of the probability the type can apply the tool correctly and the probability that the tool will, if correctly deployed, solve the problem. The overall success probability of tool t_k for type s_i therefore equals $h_{s_i}(t_k) \cdot p(t_k)$ as shown in Figure 1.

The notion of ability plays an important role in our analysis. We define a type's ability as the probability that that type solves the problem using its best tool. To avoid overcomplicating the analysis, we assume that each type has a unique *best tool*, the one with the highest overall success probability for that type. We denote the best tool of type s_i by $t^*(s_i)$.[10]

The **ability** *of type* s_i, $a(s_i)$ *equals* $h_{s_i}(t^*(s_i)) \cdot p(t^*(s_i))$, *the probability of solving the problem using* s_i*'s best tool.*

We say that a type is *dominated* if for every tool in its repertoire there exists some other type who can better apply that tool.

Type s_i *is* **dominated** *if for any* $t_k \in T(s_i)$, *there exists an* s_j *s.t.* $t_k \in T(s_j)$ *and* $h_{s_j}(t_k) > h_{s_i}(t_k)$.

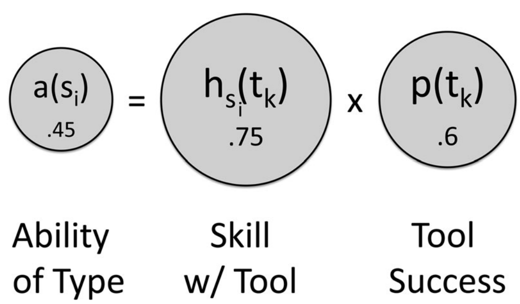

FIGURE 1 How ability depends on skill and a tool's probability of success

We say that a type is pairwise dominated if it takes only one type to dominate it. Clearly, pairwise domination implies domination but the converse need not hold.

Type s_i is *pairwise dominated* if there is a type s_j such that $T(s_i) \subseteq T(s_j)$ and $h_{s_i}(t_k) < h_{s_j}(t_k)$ for all $t_k \in T(s_i)$.

Hereafter, we refer to the type that can best apply a tool as the *expert* at that tool. (Note that a type is undominated if and only if it is an expert at some tool.) We define the *promise* of a tool to be the product of the likelihood that the tool solves the problem and the probability that its expert applies it correctly. We assume throughout that the success of any tool is independent of that of any other tool.

The **promise** *of tool* t_k *equals* $h_{s_i^*}(t_k)p(t_k)$, *where* s_i^* *is the* **expert** *for tool* t_k, *that is, the type with the largest* $h_s(t_k)$. We assume throughout that each tool has a unique expert, denoted by $s^*(t_k)$.

A group solves a problem if and only if at least one of its members solves the problem.[11] The payoff to a type depends on whether the team solves the problem (a collective benefit) and on whether she is among those who find a solution (a private benefit).

The **payoff** *to type* s_i *can be written as follows*:

$$
\begin{aligned}
u(s_i, G) &= 0 && \text{if } C = \emptyset \\
&= b \in [0, 1] && \text{if } s_i \notin C \neq \emptyset \\
&= b + (1 - b)f(|C|) && \text{if } s_i \in C,
\end{aligned}
$$

where C*denotes the set of problem solvers in* G *who solve the problem and*$f(|C|) > 0$ *for* $C = G$ *and* $f(\cdot)$ *is strictly decreasing in* $|C|$. Thus the maximal payoff accrues to an agent who is the only person on the team to solve the problem (sharing glory diminishes it); team failure generates the minimal payoff.

Under centralized tool choice, tool promise is a key measure. The manager can in effect select and assign tools, matching type to promise. Under strategic tool choice, the manager cannot assign tools. Here, tool overlap can create inefficiencies: A type may choose a personally better tool already used by someone else rather than an as-yet-unrepresented one.

We define five categories of tool overlap. First, we separate out those cases that include a *dominated* type, that is, a type that is not expert at any tool. We then define four nested sets of *undominated* types. First, types might have *no overlap* in tools (B_0). In this case, a new type always tries a new tool. This decomposability of toolkits simplifies the manager's team-composition problem.[12]

Next, types might overlap on some tools but not on their best tools, that is, each type's *best tool is exclusive* to that type (B_1). A further weakening would be to allow for overlap on best tools, requiring only that each type is the *expert at its best tool* (B_2). If recent college graduates know lithium-ion technology better than any other battery technology, then experienced types, who may know something about designing lithium-ion batteries, must know less than the recent graduates. Last, each type can be the *expert for at least one tool* (B_3), though not every type needs to be expert at its best tool. A college graduate may know just a little about a potentially useful lithium sulfur battery but that might be more than anyone else. (A college graduate's best tool, however, might be common in the industry and most skillfully deployed by experienced types who have acquired craft knowledge.) A type that is not expert at its best tool but is expert at some other tool can lead to inefficiencies under strategic tool choice. For example, suppose that type j has a high skill level with tool k, but that tool is better applied by other types. Type j is the best at tool ℓ, but that method has little chance of solving the problem. Private incentives might lead a type j to choose tool k rather than tool ℓ. Under centralized tool assignment, the manager can require type j to use tool ℓ. Under strategic tool choice, the manager has no guarantee that type j will not choose tool k instead. Figure 2 shows the five types of tool overlap in graphical form Figure 2.

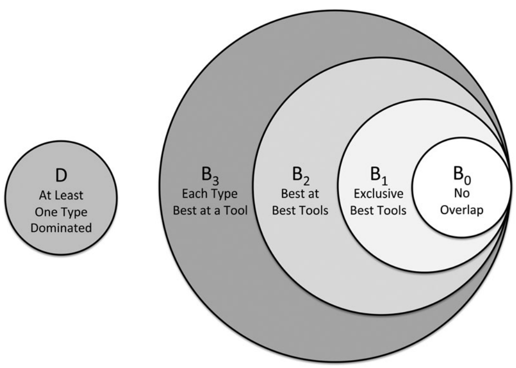

FIGURE 2 Classes of tool overlap across types

No overlap (B_0): $T(s_i) \cap T(s_j) = \emptyset \ \forall i, j$
Exclusive best tools (B_1): $t^*(s_i) \notin T(s_j) \ \forall i, j$
Experts at best tools (B_2): $h_{s_i}(t^*(s_i)) > h_{s_j}(t^*(s_i)) \ \forall i, j$
Experts at a tool (B_3): $\forall s_i \ \exists k$ s.t. $h_{s_i}(t_k) > h_{s_j}(t_k) \ \forall j$
Dominated (D): $\exists \ s_i$ s.t. $\forall \ k \ \exists s_j$ s.t. $h_{s_i}(t_k) < h_{s_j}(t_k)$

Thus, either a type is dominated (set D) or it is not (set B_3). If it is undominated then either it is best at its best tool (set B_2) or it is not (the shaded part of B_3 that is not part of B_2). And so on until we reach the smallest circle, B_0—different types have completely different toolkits (i.e., no overlap).

Our initial analysis derives necessary and sufficient conditions for optimal hiring to satisfy the following regularity properties:[13]

Hiring satisfies **competency ordering** *if for any types s_i and s_j such that $a(s_i) > a(s_j)$ at least one s_i is hired before an s_j is hired.*

Hiring satisfies **competency loading** *if for any types s_i and s_j such that $a(s_i) > a(s_j)$ the number of type s_i's weakly exceeds the number of type s_j's.*

Hiring satisfies **monotonicity** *if $|G'| > |G|$ implies that for all s_i, the number of s_i's in group G' weakly exceeds the number in group G.*

Hiring satisfies **no dominated hires** *if no type that is dominated is in any group picked by the manager.*

Hiring satisfies **asymptotic diversity** *if for a sufficiently large group size, the group constructed by the manager includes at least one of each undominated type.*[14]

3 | CENTRALIZED TOOL CHOICE

In centralized tool choice the manager selects a set of types in the first stage. In the second, she assigns tools to those types. In that stage the manager assigns each tool to its expert. Centralized tool choice requires that the manager can monitor and enforce the techniques problem solvers apply. We make that assumption so as to provide a benchmark, first-best probability of success.

Suppose first that for each tool t_k there exists a type, s_i, that can flawlessly apply t_k, that is, $h_{s_i}(t_k) = 1$. Hiring decisions become straightforward: For any given size group, $G(n)$, the manager will select the n best tools in Stage 2; working backwards, she then selects each tool's flawless type in Stage 1.[15]

Next, suppose that there exists a type that dominates all other types. In this case, the optimal group consists only of this dominant (hence best) type.

Observation 1. *For all group sizes, the optimal group under centralized tool choice consists only of the best type if and only if the best type dominates all other types.*

The straightforward proof of this observation—the manager wants to hire the expert for each tool, and a single type is the expert for *all* tools—should not undermine the fact that it reveals how strong an assumption must be made for hiring only the most talented to be an optimal strategy. Hiring just the highest ability type makes sense only if that type is the *expert at every tool*. Breadth and depth of talent will often be at odds with one another. Breadth entails less depth; conversely, great depth requires specializing in a few tools. If either condition fails to hold then for some size n, the optimal group will contain multiple types.

To analyze more complicated cases, we assume that the project manager hires sequentially, adding the type that maximally increases the probability of success. A new type can increase the probability of success if and only if she selects and correctly applies a tool that nobody in the status quo group applies correctly. (If somebody in the status quo group picks tool t_j but misapplies it, then a new type can add value by selecting t_j and using it correctly.)

To characterize the marginal value of adding type s_i using tool t_k to group G, we let $G(\text{FAIL}, t_k)$ denote the event that everyone in G fails to solve the problem and tool t_k is applied correctly by somebody in G. In that case, problem solver s_i cannot add value because tool t_k cannot solve the problem even when an agent deploys it correctly. Similarly, let $G(\text{FAIL}, \neg t_k)$ represent the event that everyone in G fails but nobody has correctly applied t_k. (Perhaps nobody even tried t_k.) In this case, type s_i can only be pivotal, that is, be the unique problem solver to find the solution, if she applies tool t_k correctly and tool t_k solves the problem. Hence, the probability that type s_i is pivotal for group G using tool t_k equals $\Pr[G(\text{FAIL}, \neg t_k)] \cdot h_{s_i}(t_k)p(t_k)$. These concepts and properties are the building blocks of the next result.[16]

Proposition 1 (Optimal hiring rule: Centralized tool choice)**.** *Under centralized tool choice, the optimal addition to group G is the type s_i who can be assigned a tool t_k that maximizes*

$$P^+(s_i, G) = \Pr[G(\text{FAIL}, \neg t_k(G))] \cdot h_{s_i}(t_k(G))p(t_k(G)).$$

Proposition 1 reveals the contributions of diversity (applying a new tool) and ability (the skill to use a tool correctly). The following corollary shows that neither diversity nor individual ability predominates. In the conditions for the corollary, type s_i has more ability and type s_j adds diversity. By altering group composition, either ability or diversity can matter more.

Part (i) states a condition for type s_j to add diversity but not improve group success as much as type s_i. Part (ii) depicts the opposite situation. One type, s_i, has an edge in ability that suffices to make it more valuable to group G. Yet s_j's diversity-advantage would make s_j more valuable for some other group G' with the same tools as G.

Corollary 1. *Let $T(G)$ denote the tools used by the members of the group G. Given centralized tool choice, assume that $T(s_i) \subseteq T(G)$, whereas $T(s_j) \bigcap T(G) = \emptyset$.*

(i) If $P^+(s_j, G) > P^+(s_i, G) > 0$ then there exists an $\epsilon > 0$ and an $s_{j'}$ with $T(s_{j'}) = T(s_j)$ such that if $\max\{h_{s_{j'}}(t_k)\} < \epsilon$ then $P^+(s_i, G) > P^+(s_{j'}, G)$.

(ii) If $a(s_i) > a(s_j)$ and $P^+(s_i, G) > P^+(s_j, G) > 0$ then there exists a G' with $T(G') = T(G)$ such that $P^+(s_j, G') > P^+(s_i, G')$.

We now state a benchmark result (see Figure 3) about the regularity properties for centralized tool choice. In this setting, a type's value rests on being expert with a tool. A type that has high ability but is not expert at any tool would never get hired under centralized tool choice.

Proposition 2. *Under centralized tool choice, optimal groups*

(i) Can violate competency loading for any type of tool overlap.
(ii) Satisfy competency ordering if and only if each type is the expert at its best tool (B_2).
(iii) Never include dominated types.
(iv) Always satisfy monotonicity.
(v) Satisfy asymptotic diversity if and only if each type is best at some tool (B_3).

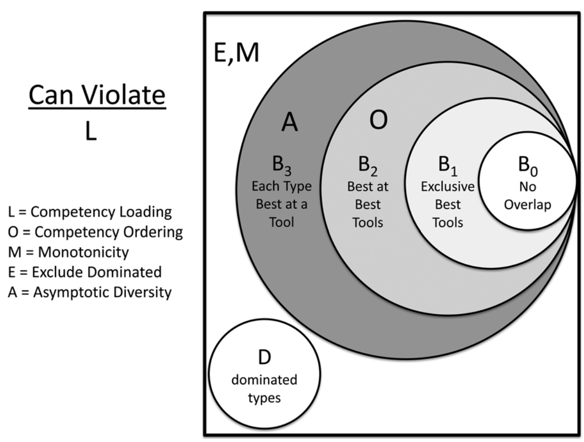

FIGURE 3 Necessary and sufficient conditions for regularity properties: Centralized tool choice

The proof of part (i) is straightforward. Competency loading is violated if there are more low-ability types in an optimal group than a high-ability type. This can happen if the more able type is highly skilled at its assigned tool. Suppose, for example, that s_1, the most able type, uses its best tool, t_1, correctly with probability one. If that is s_1's only tool then an optimal group contains only one agent of this type. If a less able type, say s_2, deploys its assigned tool with a probability less than one then hiring multiple s_2's could be optimal. Hence, the less able type would be more numerous, which violates competency loading. However, even with no overlap—indeed, even if all types have singleton toolkits—the conditions for competency loading to hold are quite restrictive: The types with the highest probability of solving the problem cannot be more likely to use their tool correctly than is any other type.

Part (ii) holds because A type's ability (competency) equals its probability of solving the problem with its best tool. If toolkits do not overlap (B_0) then types will be hired in order of ability. That assumption, though sufficient, is not necessary. A weaker sufficient condition for competency ordering to hold is that each type be the expert at its best tool (B_2). To prove that condition B_2 is sufficient, consider a manager choosing between two new types, s_i and s_j. Assume without loss of generality that s_i is more able. By assumption, each type is expert at its best tool. It follows that neither type's best tool has yet been assigned and that, by Proposition 1, the manager will instruct the chosen type to try its best tool. Therefore, the manager's optimal personnel choice between s_i and s_j is whichever type has the more promising best tool. Given B_2, the promise of a type's best tool equals the type's ability. Hence, if s_i is more able than s_j, its best tool is more promising than s_j's, whence the manager's optimal personnel choice is s_i, so competency ordering holds.

To prove that condition B_2 is necessary we consider the case where s_i knows how to apply only one tool, $t_{\hat{k}}$. As before, s_i is more able than s_j. Here, however, there exists an s_k who is better at $t_{\hat{k}}$ than s_i is. For example, suppose that the best technique for finding a bug in computer code is to evaluate error messages, the second best technique is to divide and conquer, and the third best is to write internal checks, with success probabilities of 0.4, 0.3, and 0.2, respectively. Suppose type s_1 can evaluate error messages perfectly and write internal checks half of the time. Hence, it is the best type. Assume that type s_2 correctly evaluates error messages or write internal checks 70% of the time. Finally, suppose that type s_3 successfully applies divide and conquer techniques half the time. Given our assumptions, the second type is more able than the third type but s_2's ability is based on its skill at evaluating error messages—a tool that could be better applied by an s_1. The optimal hiring rule involves first hiring an s_1 followed by an s_3 and only then hiring an s_2. This violates competency ordering.

Result (iii) holds because in the centralized system optimal groups include only experts. Because an expert is better than any other type regarding at least one tool, experts are undominated, thus ruling out dominated types. For intuition for result (iv) consider a group whose members apply unique tools. Each type must be an expert at its tool. The manager will never reverse her plan of telling s_i to try t_k as the group gets larger. The general proof (in the appendix) relies on the independence of tool success. Finally, the result on asymptotic diversity (v) follows from the fact that any tool might solve the problem and no tool solves the problem for sure. Therefore, every sufficiently large group includes an expert for each tool. Asymptotic diversity therefore requires B_3. If not, a type that is not an expert for at least one tool would never be added to the group.

The previous proposition implies, among other things, that hiring by ability would not be optimal. Nor would hiring by diversity, if we define that as hiring the type most likely to correctly apply an untried tool. Moreover, even if the highest ability type is also the most likely to apply an untried tool, that type may be suboptimal, as we show below. In fact, we show that such rules of thumb can produce *marginally worthless hires*.

A hiring rule H that selects type $H(G)$ given the group G produces a **marginally worthless hire** *if for any $\epsilon > 0$, there exists a group G_ϵ and a type $s_j \neq H(G_\epsilon)$ such that $P^+(H(G_\epsilon), G_\epsilon) < \epsilon P^+(s_j, G_\epsilon)$.*

A marginally worthless hire increases the probability of solving the problem by ϵ times the increase from some other type.

Remark 1. Under centralized tool choice, hiring by diversity can produce marginally worthless hires even with *no overlap* (B_0). Hiring by ability or (if possible) by ability and diversity can produce marginally worthless hires unless types are *experts at best tools* (B_2).

4 | STRATEGIC TOOL CHOICE

Under strategic tool choice the manager determines the composition of the group but cannot control what tools the individuals select. Instead, each member of the group chooses a tool taking into account the choices of the other group members. The extent to which strategic tool choice creates coordination and incentive problems depends on the number of tools in each type's toolkit, on the overlap across toolkits, and on the importance of private benefits. If all toolkits are singletons then the manager's problem is easy, as it is de facto equivalent to centralized tool choice. When the types have multiple tools, incentive problems can arise and there may be multiple equilibria. In such circumstances, strong common interests (high b) or a decomposable skill structure (nonoverlapping toolkits) simplifies the manager's problem. The extent of coordination problems also depends on whether workers choose tools simultaneously or sequentially.

4.1 | Simultaneous tool choice

We first consider the case in which each member of the group simultaneously selects a tool or a probability distribution over tools. Under this assumption, a symmetric equilibrium must exist; further, if toolkits do not overlap then problem solvers apply tools that have higher promise with higher probability.

Proposition 3. *Given any number of types and any number of tools, an equilibrium exists in which every member of a type chooses the same distribution over tools. Further, if problem solvers' tools do not overlap and* $h_{s_i}(t_j) \geq h_{s_i}(t_{j'})$ *for* $j < j'$*, in equilibrium type* s_i *applies tool* t_j *with higher probability than it applies tool* $t_{j'}$*.*

A less intuitive comparative static result follows as a corollary (given nonoverlapping toolkits). If either there exists a large private benefit from solving the problem or the group of problem solvers is sufficiently large, then introducing a new type of problem solver creates incentives for those already in the group to shift toward better tools.

Corollary 2. *Assume problem solvers have nonoverlapping toolkits, choose tools strategically, and that given a group G, type* s_i *plays a symmetric mixed strategy equilibrium strategy* $\vec{q}^*(G) = (q_1^*(G, s_i), q_2^*(G, s_i), \ldots, q_K^*(G, s_i))$*. If a problem solver who uses a new tool is added to the group, then the change in the probability that* s_i *chooses tool* t_k *is increasing in* $h_{s_i}(t_k)p(t_k)$*.*

Our next two propositions give conditions for the group to be composed only of the best type to be optimal and for the most diverse group to be optimal. The first result requires a dominant type and a small private benefit.

Proposition 4. *Consider a group of fixed size under simultaneous tool choice. If type* s_1 *dominates all other types and b is sufficiently close to one then (a) any pure Nash equilibrium of this homogeneous group of the best type implements the first-best outcome and (b) no other group can implement the first best.*

Although a homogeneous group of the best type generates multiple equilibria, this causes no problems for the manager if the members of the group are oriented primarily toward group effectiveness. In such circumstances the individuals will not overuse more promising tools: If it is collectively optimal for some team members to select less promising tools, this will happen in equilibrium. The combination of low conflict (group effectiveness is everyone's priority) and a type that dominates all others in ability yields intuitive team-theoretic effects (Marshack & Radner, 1972): the manager can hire the type, that is, obviously the most competent and the group members will do exactly what she wants. This team-theoretic logic fails if the private benefit from solving the problem is high. In such circumstances the optimal team may contain multiple types, even though one type dominates all others. For the most diverse group to be optimal, no one type can be better at two tools than some other type is at its best tool, a condition we formalize as follows:

Definition. We say that *best tools are more promising* if the following condition holds: If t_j is the best tool of some type and t_k is not the best tool of any type then $h_{s_j}(t_j) > h_{s_k}(t_k)$, where s_j is the expert for t_j and s_k is the expert for t_k.

Proposition 5. *Consider a group of fixed size under simultaneous tool choice. If best tools are exclusive and more promising and* $\min(h_{s_i}(t^*(s_i)))$ *is sufficiently close to one then (a) the unique pure Nash equilibrium of the group that is fully diverse and satisfies competency ordering implements the first-best outcome, and (b) no tool vector selected by any other group implements the first best.*

When best tools are exclusive (B_1) and the team is fully diverse (every team member is a different type) then for all b each agent has a strictly dominant strategy of picking her best tool. Hence this vector of tools is realized in the unique

Nash equilibrium. Further, this equilibrium is the first-best solution if the hiring in Stage 1 satisfied competency ordering, best tools are more promising than other tools, and the corresponding h's are sufficiently close to one. Finally, no other team can satisfy all of these criteria, so all fall short of the first-best.[17]

Strategic tool choice will typically not cause distortions if the game has a unique Nash equilibrium which implements the first-best outcome. The next result, however, shows that rather stringent conditions must be satisfied to ensure the existence of a vector of tool choices that is both optimal for the manager and the unique strategic equilibrium for the team members.

Proposition 6. *Under simultaneous tool choice, the following hold in any unique pure Nash equilibrium that produces the first-best tool vector and outcome in Stage 1:*

(i) Every problem solver of a given type selects the same tool.
(ii) Each type is the expert at the tool that it selects.
(iii) Each tool is selected by at most one type.
(iv) Only the best type, s_1, must select its best tool, $t^*(s_1)$.

In Proposition 6, condition (i) must be satisfied or multiple equilibria exist. If two problem solvers of the same type choose distinct tools in an equilibrium, then there exists a second equilibrium in which they switch their tool choices. Condition (ii) must hold because optimality requires that each tool be tried by its expert. Condition (iii) follows from our assumption that each tool has a unique expert. Condition (iv) holds because a type need not be the most skilled at its best tool, but the incentives must be such that it chooses a tool for which it is the most skilled type, that is, it is the expert for that tool.

If types have overlapping toolkits then multiple equilibria may exist. For example, suppose that s_1 and s_2 have the same toolkit of $\{t_1, t_2\}$, with s_1 the expert at t_1 and s_2, of t_2. If the types' skills at each tool are similar (the difference between, e.g., $h_{s_1}(t_1)$ and $h_{s_2}(t_1)$ is small) then a heterogeneous two-person team will exhibit two pure equilibria: In one, each tool is tried by its expert; in the other, the opposite occurs.

Multiple equilibria may not present a difficult problem for a self-organizing group if b is close to one. The general intuition is that strategic tool choice can work "well" if problem solvers are motivated primarily by team success. The following result verifies part of this intuition. Note that this result, unlike Proposition 6, does not assume that the tool choice game has a unique pure Nash equilibrium.

Remark 2. Suppose for a fixed n there is a unique first-best tool vector and a corresponding optimal team, $G^*(n)$. If b is sufficiently close to one then the following hold for $G^*(n)$ in the simultaneous tool choice game.

(i) The first-best tool vector, chosen by the corresponding experts, is supported by a strict Nash equilibrium of this game.
(ii) The payoffs in this equilibrium strictly Pareto dominate the payoffs produced by any other tool vector.

Collective interest overcomes the incentives to duplicate tool choices but it need not solve coordination problems. For example, suppose that there exist n problem solvers who all have approximately equal skill with each of n tools. There exists a large set of cases in which the optimal solution calls for each problem solver to choose a unique tool. However, if there exist any private incentives, each problem solver would like to choose the best tool. This game has $n!$ possible equilibria. Coordination would be difficult. And even if coordination occurs, these equilibria may be Pareto ranked. For example, one type of problem solver might be better at applying tool j but choose tool k, while another type, which is better at applying k, could choose tool j. Making decentralized tool choice sequential and assuming subgame perfection rules out these inefficient equilibria.

4.2 | Sequential tool choice

Under sequential tool choice, the manager stipulates the order in which members of the team select the tools that they will try in Stage 2. These choices are common knowledge and irreversible. Hence, in Stage 1, the manager can accurately forecast the equilibrium behavior of her subordinates. The fundamental problem that sequential choice can solve is that of coordination failure. We consider two types of coordination failure categorized by the degree of conflict. The easier one, which we analyze first, involves "pure" coordination failures: The equilibria are Pareto-ordered and the manager wants to avoid the perverse outcome of a dominated equilibrium (the setting of Remark 2).

Corollary 3. *Suppose that tool choice is sequential and that otherwise the assumptions of Remark 2 hold. Then the first-best tool vector is realized in equilibrium.*

Thus, when tool choice is sequential and team members are mostly oriented toward group-success, Pareto-dominated outcomes cannot occur.

The second type of coordination failure involves multiple equilibria which cannot be Pareto-ordered. In this context, the individuals disagree as to which equilibrium should be played: Individual incentives are strong enough to make members of the team want to play more promising tools, leaving their teammates with the burden of trying less promising tools. For example, consider a group with two members: An s_1 and an s_2, the respective experts on t_1 and t_2, the two most promising tools. Each type can use either tool. There are two pure Nash equilibria: In one, s_1 tries t_1 and s_2 tries t_2; in the other, each specialist tries the other type's tool. Because t_1 is considerably more likely to solve the problem, each type prefers the equilibrium in which s/he tries t_1. The manager, of course, wants each tool to be selected by its expert. Sequential tool choice secures this: The manager tells s_1 to choose first, s_1 picks t_1, forcing s_2 to select t_2. Sequential choice creates a game with a *first-mover advantage* which can be used by the manager to maximize the probability of solving the problem.[18]

Of course this procedure cannot work if private incentives so powerful that each type has a dominant strategy of picking t_1. In such circumstances, the choice sequence is irrelevant. Hence the ability of the manager to stipulate the order of tool selection has the most value for intermediate levels of conflict among teammates. If there is little conflict —b is close to one—then multiple equilibria will be Pareto-ordered. This is not a difficult coordination problem; self-organizing groups should be able to handle such issues. Contrarily, extreme congestion problems arise if there are strong private incentives to solve the problem. Every agent has the same dominant strategy to apply its best tool. In these cases, sequencing tool choices is ineffective; only centralized tool-selection produces the first-best solution.

4.3 | Regularity properties in strategic tool choice

Strategic tool choice requires more stringent assumptions to ensure that the regularity properties hold (see Figure 4). We confine attention to sequential tool choice and consider only pure strategy equilibria. The next proposition covers all five regularity properties. Comparing these results with centralized tool choice (Proposition 2) reveals that the final three properties are much weaker while the first two are unchanged.

Proposition 7. *Under sequential strategic tool choice, optimal groups*

(i) Can violate competency loading for any kind of tool overlap.
(ii) Satisfy competency ordering if and only if each type is expert at its best tool (B_2).
(iii) Can violate no dominated types for any kind of tool overlap.
(iv) Can violate monotonicity for any kind of tool overlap.
(v) Satisfy asymptotic diversity if and only if each type has exclusive best tools (B_1).

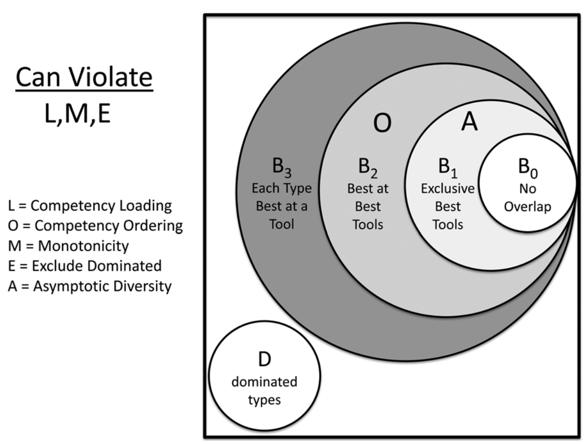

FIGURE 4 Necessary and sufficient conditions for regularity properties: Strategic tool choice

Competency loading fails in the strategic setting for the same reason that it fails under centralized tool choice: An effective tool (one with a high p) may be easy to apply (its expert's h is high), making intratool redundancy unnecessary. Competency ordering again holds so long as each type is the expert at its best tool. The logic is identical to that of centralized tool choice. If a type is the expert at its best tool then the best tools of more able types are more promising. The corresponding specialists should be hired earlier than types that are experts of inferior tools.

The conclusions about competency loading and competency ordering hold regardless of whether individuals care primarily about team success (high b) or personal gain (low b). Indeed, the results hold even if b varies arbitrarily across types. The relative unimportance of private incentives underscores the power of nonoverlapping toolkits and of the importance of the distribution of skills across types.

Strategic tool choice flips the result about hiring dominated types. This finding rests on the possibility that a dominated type creates fewer strategic problems. Suppose, for example, that type s_1 is perfect at applying tool t_1 and can apply tool t_2 correctly half the time. Suppose that an s_2 type can apply both tools correctly slightly less than half the time; hence s_1 dominates s_2. For a large class of parameter values, a team of two s_1's will both choose tool t_1 whereas a diverse team will choose both tools. Thus the (s_1, s_2) team will outperform the team of two s_1's.

This example illustrates how the manager can tap the *strategic comparative advantage* of a dominated type. If the s_1 type is extremely good at t_1 then that specialist may be unable to credibly commit in Stage 1 to selecting t_2 in Stage 2. In contrast, if s_2's comparative advantage in t_2, relative to her skill in t_1, is sufficiently strong then the manager knows that it is a best response for an s_2 to follow through on a promise to try t_2. Thus, although that type has no absolute advantage with any tool, it has a strategic *comparative* advantage at an inferior tool—the problem-solving version of Ricardo's insight.

As an index of the dominated type's comparative advantage in tool t_2 we use $\theta_{s_2}(t_2) = 1 - ((h_{s_2}(t_1))/(h_{s_2}(t_2)))$. This index is positive if s_2 has any comparative advantage in tool t_2 (i.e., if $h_{s_2}(t_2) > h_{s_2}(t_1)$) and reaches its maximum value of one if s_2 is completely feckless at t_1. This strategic comparative advantage can be useful to the manager even when the inferior type, s_2, is *completely* dominated by s_1. This strong sense of skill domination is defined below.

Definition. Type s_i completely dominates s_j if $T(s_j) \subseteq T(s_i)$ and $\min\{h_{s_i}(t_k)|t_k \in T(s_i)\} > \max\{h_{s_j}(t_l)|t_l \in T(s_j)\}$.

Complete domination is about as clear a skill difference as can be imagined: s_i is better at everything in its toolkit than s_j is at anything in its repertoire. (This is a special kind of pairwise domination, which in turn is nested inside ordinary domination.) Recall that under centralized tool choice a rational manager never selects a type that is dominated in any way, for example, even if domination requires all other types. In contrast, under strategic choice she may hire a *completely* dominated one.

Remark 3. Suppose there are two types and s_1 completely dominates s_2. The best tool of s_1 is t_1 and s_2's best tool is t_2. If tool choice is strategic then there exists a vector of parameter thresholds $\{b^* > 0, h^* < 1, \theta^* < 1, p^* > 0\}$ such that if $b < b^*$, $h_{s_1}(t_1) > h^*$, $\theta_{s_2}(t_2) > \theta^*$, and $p(t_2) < p^*$ then (a) the optimal two-person group is (s_1, s_2) and (b) in its unique equilibrium in Stage 2 each type picks its best tool.

As is often the case, the manager's quest here is for tool diversity, and if this would not be implemented in the Nash equilibrium of a homogeneous team of the most able type then in some parametric regions it is worthwhile to accept a dominated worker to attain tool diversity.[19] Next, we show how monotonicity can fail for any overlap, including the case of *no overlap* (B_0). The violation arises because as groups grow larger, private benefits decrease: The glory of solving the problem will likely be shared by several teammates. This weakens the incentives for choosing a unique, less effective tool, having the same effect as increasing b, the collective benefit.

Asymptotic diversity violations also require that a type's best tool not be unique. The proof, though technically involved, relies on a straightforward construction. Suppose there are three types and two tools. Assuming that the best type can apply the better tool correctly with probability one, the second type can correctly deploy the better tool and a lesser tool with probability $h < 1$, and the third type has only the inferior tool in his repertoire. Further, suppose that the second type is marginally more capable with the lesser tool than is the third type. When the private benefits for solving the problem are large, the second type will choose the better tool when added to a team of one person. Thus, the third type is a more valuable addition to that group. The asymptotic result relies on similar logic.

We conclude this section by stating an analog of our previous result for centralized tool choice relating the potential loss from hiring rules based on ability, diversity, or the combination of diversity and ability.

Remark 4. Under sequential tool choice, hiring by diversity can produce marginally worthless hires even when toolkits do not overlap. Hiring by ability or ability and diversity can produce marginally worthless hires unless experts at best tools (B_2) holds.

This claim relies on the same logic as in the centralized case. Using diversity alone can lead to poor outcomes because team performance depends on the tools applied, not on the set of potential tools. Balancing ability and diversity can also produce worthless hires when problem solvers are not experts at best tools because as we have seen, a high-ability problem solver might fail to choose a new tool.

If, however, each person brings different expertise to the table then overlap will not arise. We consider that case next.

4.4 | Experts and the opportunity costs of expertise

It is well known that it takes many years and much effort to become highly proficient in chess, cryptography, physics, and other informationally intensive domains (Ericsson et al., 1993; Feltovich et al., 2006). Indeed, it is generally believed that it is *impossible* to become a top notch expert in such fields without thousands of hours of intense study and practice. (How much is training vs. real experience varies across domains and over time—e.g., "book learning" in medicine has greatly increased over the last century—but the total number of necessary hours is both high and not all that variable across informationally intensive domains.) This last section incorporates that robust empirical regularity by making an assumption that reflects the temporal opportunity cost of expertise: In modern economies, everybody faces a time budget that binds. In the context of our framework, that means, roughly speaking, that a type can become proficient in only a few tools. Hence candidates for the team might be either specialists, who are good at a few tools, or generalists, who are mediocre at many.

To derive the formal conditions necessary for this to hold, we say that type s_j is *highly specialized* if (a) it is the expert at some tool, say t_j, and (b) for all other $t_k \in T(s_j)$, there is a small but strictly positive ϵ_{s_j} such that $h_{s_j}(t_k) < \epsilon_j$ for all $k \neq j$. We assume, of course, that $\epsilon_{s_j} < h^{\star}_{s_k}(t_k)$, the skill level of the expert at t_k; thus a highly specialized type is an expert at only one tool. If becoming an expert is time-consuming then everyone in an optimal group will be highly specialized. This result holds for both simultaneous and sequential games.

Remark 5. We assume a group of a fixed size, n. If all experts are highly specialized and none is perfect at applying any tool then everyone in any optimal group of size n is highly specialized. If in addition $b < 1$ then for any optimal group, $G^{*}(n)$, there exists an $\epsilon > 0$ such that if $\max\{h_{s_i}(t_j)\} < \epsilon$ for every $t_j \in T(s_i)$ which is not s_i's best tool, then everybody in $G^{*}(n)$ has a strictly dominant strategy of selecting his best tool and the tool vector produced by the corresponding unique Nash equilibrium is the first best.

Thus, this high degree of skill specialization delivers just what the manager wants: Problem-solving effectiveness and clean strategic properties. The former is easy to see: In general the first-best tool vector is produced only by experts and Remark 5 assumes that all experts are highly specialized, so this part of what the manager wants follows immediately.[20] The strategic benefits of skill specialization are a bit more subtle. It helps to examine the limiting case in which experts are completely feckless with everything in their repertoire that is not their best tool. (Technically this special case is excluded from the feasible set because it means that $h_{s_j}(t_k) = 0$ for all $k \neq j$ and we assume throughout that $h_{s_i}(\cdot)$ is strictly positive for all i, but it is a useful limiting case.) In this circumstance, an expert has no chance of solving the problem with anything other than his best tool, so trying that tool is obviously a dominant strategy for every expert for any composition of $G(n)$; eliminating the knife-edge case of $b = 1$ ensures that selecting one's best tool is strictly dominant.

We can see the strategic advantage of hiring only experts from a somewhat different angle by reusing the idea of strategic comparative advantage introduced earlier. Remark 3 showed that the manager may find it worthwhile to put a dominated type on the team if that type enjoys a strategic comparative advantage in a tool that the manager wants the group to try. In the context of Remark 5, however, *only experts enjoy strategic comparative advantage*. This is a direct effect of their being highly specialized: Since experts are inept at anything that is not their speciality they have a comparative advantage only with their best tool, which is exactly what the manager wants the experts to use. All strategic coordination problems disappear.

Remark 5 implies that if the first-best tool vector is completely diverse then so is the optimal team. In such circumstances tool diversity and team diversity are linked. Nevertheless, competency loading can still be violated: A less promising tool, say t_k, might be more difficult to apply than a more promising one, t_j; hence there may be more t_k experts on the team than t_j specialists. It is straightforward to show that the other four regularity hiring properties—competency ordering, monotonicity, no dominated hires, and asymptotic diversity—will satisfied in strategic tool choice under the assumptions of Remark 5. Hence we might regard this result as identifying a set of conditions that make the manager's hiring problem a "nice" one.

5 | EXTENSIONS: PARTIAL SOLUTIONS, MULTIPLE TOOLS, AND UNCERTAINTY

Our framework makes three strong assumptions. First, we do not allow for the possibility that a tool is a partial solution. In some cases, one tool might solve a part of a problem, while a second tool solves the remainder. Second, we assume that a problem solver applies a single tool. In practice, a problem solver might apply several. Third, we assume that in Stage 1 the manager knows what problem will confront the team in Stage 2. Sometimes, however, more than one problem may arise and a manager is uncertain which it will be. In this section, we describe how our model can be extended to account for all three possibilities.

Consider the following sketch of a model that allows for partial solutions. We first assume that each problem has two parts, denoted by "/" and "\" which together form an X. To simplify the analysis, we assume that if a problem solver applies a tool correctly on one part then she applies it correctly on all parts. We can then rewrite the probability that type s_i solves the problem using tool t_k as $h_{s_i}(t_k)\cdot p^{/}(t_k)\cdot p^{\backslash}(t_k)$, where $p^{/}(t_k)$ and $p^{\backslash}(t_k)$ correspond to the probabilities of solving the two parts of the problem, given that the tool was applied correctly.

The probability that two problem solvers, s_i and s_j, who apply tool t_k solve the problem equals the probability that at least one of them correctly applies the tool and that the tool solves both parts of the problem:

$$[h_{s_i}(t_k) + h_{s_j}(t_k) - h_{s_i}(t_k)\cdot h_{s_j}(t_k)]p^{/}(t_k)\cdot p^{\backslash}(t_k).$$

In contrast, the probability that two problem solvers, s_i and s_j, who apply tools t_k and $t_{k'}$, respectively, solve the problem equals the sum of probabilities that each solves it on their own minus the probability that both solve it *plus* the two cases where each solves a different part of the problem

$$[h_{s_i}(t_k)p(t_k) + h_{s_j}(t_{k'})p(t_{k'}) - h_{s_i}(t_k)p(t_k)\cdot h_{s_j}(t_{k'})p(t_{k'})] + h_{s_i}(t_k)\cdot h_{s_j}(t_{k'})[p^{/}(t_k)\cdot p^{\backslash}(t_{k'}) + p^{\backslash}(t_k)\cdot p^{/}(t_{k'})].$$

The possibility that each tool solves a different part confers an advantage to groups that applies diverse tools. This benefit from diversity is the main effect of allowing partial solutions.[21] A straightforward argument shows that because success requires only that at least one tool solves each part, increasing the number of parts makes more tools even more advantageous. Thus, decomposability (Simon, 1962) and diversity (Hong & Page, 2004; Page, 2008) appear to be complements.

A second question is how the potential for multiple solutions influences incentives. Notice first that if toolkits do not overlap, partial solutions have no effect other than to make weaker tools more promising at the margin. Some incentive problems would be attenuated. For example, the second member of a type who has skill with two tools would have a stronger incentive to choose the inferior tool. In cases where the better tool has a higher probability of success on one part of the problem and the weaker tool has a higher probability on the remainder, the incentives for the second member of the type added to the group would be clear: Choose the inferior tool.

None of this implies that allowing for partial solutions overcomes all incentive problems. Rather, it shifts the range of parameters for which incentive problems arise. The potential for violations of monotonicity and of no dominated types still exists. The only difference is that with partial solutions, tools with even lower probabilities of success should be applied to problems.

Including the possibility that a type can apply multiple tools requires assumptions about the number of tools applied. If a type applies all its tools then no dominated type would ever be hired. If a type can apply at most K tools, then the assumption has no effect unless types overlap in their tools. Accordingly, violations of competency loading, monotonicity, and no dominated types would still arise.

If we fix the number of tools that a type can apply, the same kinds of incentive problems arise as in the single tool case. For instance, suppose that at most three tools can be applied by any problem solver and that type s_1 has skill with tools t_1, t_2, t_3, and t_4. A second type s_1 on the team would face a choice between applying any of the four sets consisting of three tools. Among these, the sets $\{t_1, t_2, t_3\}$ and $\{t_1, t_2, t_4\}$ have the highest individual success rate. If type s_1 is a virtuoso with tools t_1 and t_2, the choice between these two sets for the second type s_1 is equivalent to choice between t_3 and t_4 when two other problem solvers apply tools t_1 and t_2 and third agent applies t_3.

Using similar constructions, the arguments used for single tools carry over to the multiple tool setting with one important caveat. Now, those proofs require particular assumptions on tool overlap. More important, if types can apply multiple tools then a type's skill with multiple tools becomes more valuable. Paradoxically, a type with skill at a single, novel tool may be hired earlier because other tools will be more likely to have been applied correctly.

Third, the manager may not know for certain in Stage 1 which problem will confront the team in Stage 2. The simplest version of this extension would presume an exogenously fixed probability distribution over a set of possible problems and zero conflict within the team ($b = 1$).[22] In such a setting, it is intuitive that problem uncertainty will increase the value of tool diversity, as the optimal mix of tools will typically be state (i.e., problem) contingent and in Stage 1 the manager does not know which state will arise in Stage 2. Further, we suspect that the existence of unlikely but dangerous problems—a fire in an oil refinery; an earthquake in a populous area—heightens the tension between ability, as measured by skill with tools that are likely to be needed, and diversity. Indeed, when it is difficult to identify (diagnose) problems—is this an X or a Y?—it may be optimal to include dominated generalists on the team even when there are no issues of credible commitment.[23] Finally, if the set of possible problems is very large then long-standing issues of organizational boundaries can arise. When maintaining optimal diversity inside the firm is too expensive then it will be outsourced. This in turn will generate interesting contracting issues.

Our main findings hold in all three extensions. Optimal groups must balance ability, as measured by skill with the most promising tools, and diversity, as measured by skill with unique tools. In addition, in some situations, the manager has incentives to take unconventional actions, such as hire dominated types (because they are more likely to apply high margin tools) or violate monotonicity (because increasing team size changes the private benefit from solving the problem).

6 | DISCUSSION

In this paper, we construct a framework to analyze team problem solving under centralized and decentralized tool choice. Our main results can be grouped into four sets. First, we characterize optimal team composition. That result reveals the direct benefits from diverse tools. In highlighting the importance of applying diverse tools, our main findings align with a growing empirical literature on cognitive diversity and problem solving in organizations (Wuchty, Jones, & Uzzi, 2007), crowds (Jeppesen & Lakhani, 2010; Lakhani & Jeppesen, 2007), academic publishing (Jones, Uzzi, & Wuchty, 2008; Uzzi et al., 2013), and team-based IQs (Bachrach, Graepel, Kasneci, Kosinski, & Gael, 2012). These studies consider nearly every academic paper ever published and find positive correlation between diversity, measured by email address, and impact and citations (Freeman & Huang, 2014). Deeper dives into smaller samples find a significant contribution of diversity as captured by references to atypical literatures—a crude proxy for our diverse tools (Schilling & Green, 2011; Shi, Adamic, Tseng, & Clarkson, 2009).

Second, we show that if a manager can control tool choice then optimal groups never include dominated types but they must satisfy asymptotic diversity. If we further assume specialists to be the best at their best tools then optimal groups also satisfy competency ordering; since we assume that each tool has a unique expert, monotonicity holds as well. However, competency loading—better types are more numerous—need not be satisfied, and in fact rarely holds in optimal groups. Together, these results imply that project managers with tool assignment authority seeking to build optimal groups can abide by some, though not all, of a standard set of hiring practices.

Our third set of results characterizes properties of the Nash equilibria in the strategic tool-selection game. These results underpin the fourth set which show that the above regularities do not hold when the manager selects team members but lacks the authority or information necessary to assign tools. In those cases, optimal groups can and typically will violate all of the standard properties. True, competency ordering—better types are hired earlier—holds if types are experts at their best tools, and asymptotic diversity holds if each type has *exclusive best tools*. But apart from those special cases, none of the conventional hiring rules (e.g., hire the best) apply.

These third and fourth sets of results apply to environments in which the manager does not or perhaps cannot know the tools a worker applies. Given the difficulties that can arise from strategic tool choice, one might then think that organizations should not allow problem solvers the freedom to choose tools. Such a policy would require the manager to possess deep and accurate knowledge of workers' repertoires and skill levels and knowledge of the efficacy of different tools for the problem at hand. This may be unlikely. Empirical studies of this type of environment reveal a tendency toward hiring people with general skills (Lazear, 2009), that is, lots of tools. Our model suggests a possible flaw in that approach. Generalists will have incentives to choose their best tools, which may overlap across types. If the manager cannot dictate which tools are tried, the firm may prefer specialists who create fewer coordination problems regarding tool choice.

A manager could, of course, ignore the possibility of unorthodox solutions and stick with the standard rules despite knowing that doing so could produce suboptimal groups. That strategy would be practical if the violations produced modest inefficiencies—if they were nearly efficient. We show, however, that things can be much worse than that: hiring by ability, hiring by diversity, and even hiring by both diversity and ability can produce hires that are worthless at the margin. Thus, a manager should design groups via the more combinatoric tool-based approach described here. This approach may entail nontraditional hiring practices, for example, putting a specialist with a dominated set of skills on the team, removing previously hired workers, and so on.

When modeling strategic tool choice we have assumed both individual and collective incentives. If a manager can weaken individual incentives, be they financial or reputational, she reduces the coordination problems. However, because tool choice is discrete, coordination failure may persist even when individual incentives are weak. Recall also that coordination failure can arise because a team member's actions can affect other agents' incentives. When a new problem solver successfully applies a tool that nobody on the existing team has correctly applied, the probability of solving the problem increases. That increased probability boosts the incentives for existing team members to choose better tools *even if someone else on the team is also using that tool* rather than unique tools. This dampens diversity and reduces the value of the new tool.

We also discuss three extensions. The first allows for partial solutions. The second enables individuals to apply multiple tools. The third allows for the possibility that the manager is uncertain which problem will confront the team. None of these extensions qualitatively changes our results. In yet another extension one could allow individuals to differ in their capacity to apply tools. Problem solvers with greater capacity could apply more tools and would have higher measured ability. This extension would tilt the results in favor of higher capacity individuals but would not negate diversity's value. Diversity would contribute to team performance so long as the set of tools relevant to a task exceeds the repertoire of any one person.

In addition to their applications to hiring and staffing within firms, our findings have implications for crowdsourcing, that is, presenting a problem to many problem solvers who need not belong to the organization. Managers have little to no control over the tools used by a crowd. As a crowd becomes larger, individuals pursuing a private reward have greater incentive to try more promising tools, attenuating the increase in crowd effectiveness as a function of size. If participation is costly, individuals should also be less likely to join larger crowds (Terwiesch & Xu, 2008). Contrarily, for problems that might be solved by many different tools, the effect of the participation cost would be smaller. This logic aligns with the finding of Boudreau, Lacetera, and Lakhani (2011) that the free rider problem is less pronounced in more uncertain domains where the appropriate toolkits are larger. Relatedly, for organizations that do both in-house research and solicit the advice of outsiders (i.e., external crowdsourcing), the crowd's presence could induce in-house researchers to use more promising tools.

In sum, our framework allows us to unpack the contributions of individual ability and collective diversity in problem solving and to characterize the incentive problems that arise with strategic tool choice. We find that the potential for coordination failure, by that we mean people applying the same tools, is reduced when toolkits are disjoint, implying a strategic benefit from specialized expertise as well as the direct problem-solving benefit of tool diversity.

ACKNOWLEDGMENT

Prof. Page's research was partially funded by Army Office of Research ARO-MURI grant W911NF1010379.

END NOTES

[1] Data from Federal Reserve Bank of St. Louis. On the growth of the knowledge based economy, see Autor, Katz, and Kearney (2006), Bell (1973), Florida (2002), and Wolff (2006).

[2] Teams predominate in the academy as well. The majority of science, and most of the best science, is accomplished by teams, not individuals (Cooke & Hilton, 2015). Over 90% of papers in science and engineering have multiple authors (6–10 in the modal paper), as do over 60% of papers in the social sciences. In both categories team authored papers are more than four-and-half times as likely to receive one hundred citations (Singh & Fleming, 2010; Uzzi, Mukherjee, Stringer, & Jones, 2013).

[3] Voting, forecasting, and reliability also violate additive separability. In reliability studies, the probability of collective failure equals the product, not the sum, of the probability of individual failures. The game theoretic reanalysis of Condorcet jury theorems, pioneered by Austen-Smith and Banks (1996), has incorporated both nondecomposable production functions and strategic interaction. Team composition and the associated argument about the relative importance of ability versus diversity have not been prominent topics in this program however.

[4] Define M types of labor (tasks), assume that each worker type has some capability with each type of labor, and assume that each worker gets assigned to a single task. Let K denote the level of capital and denote output as follows:

$$\left(\beta K^{\rho}+\sum_{i=1}^{M}\alpha_i L_i^{\rho}\right)^{1/\rho}$$

where $\sum_{i=1}^{M}\alpha_i=(1-\beta)$. Given those assumptions all five regularities hold.

[5] In a companion paper, we allow for the possibility that different methods improve the status quo by different amounts.

[6] Using the problem categories of Terwiesch and Xu (2008), our framework applies to expertise-based projects and trial and error projects but not ideation projects.

[7] This probability can be interpreted as applying a tool in an appropriate manner to a given class of problems. In this paper we consider only one type of problem, so we do not need to condition tool-skill on the kind of problem.

[8] Assuming that types apply all of their tools would produce identical results in some cases. In other cases (when toolkits overlap) that assumption would produce correlation in the probabilities of success. This would affect our results quantitatively but not qualitatively. However, assuming that individuals can try their entire toolkit removes any possibility of strategic tool choice, a focus of our analysis.

[9] Such critiques do not apply to production problems encountered so often that managers know strictly more than their subordinates (Garicano, 2000; Garicano & Rossi-Hansberg, 2006).

[10] For example, suppose type s_i has capabilities (0.4, 0.5, 0.9) with tools t_1, t_2, and t_3, which solve the problem with probabilities 0.6, 0.5, and 0.2, respectively. The best tool for type s_i is tool t_2 which has an expected probability of solving the problem of 0.25.

[11] Formally, the probability that a group G consisting of the two problem solvers, s_i and s_j, solves the problem using tools t_k and t_ℓ, respectively, equals $h_{s_i}(t_k)\cdot p(t_k)+h_{s_j}(t_\ell)\cdot p(t_\ell)-h_{s_i}(t_k)\cdot p(t_k)\cdot h_{s_j}(t_\ell)\cdot p(t_\ell)$.

[12] Decomposability generally makes problems easier to solve (Simon, 1962).

[13] The analogy between problem solving and production requires interpreting our probability of solving the problem as a quantity. That is slightly problematic in that the probability of solving a problem is bounded, whereas product quantity is not. However, none of the counterintuitive solutions that we find appear to be caused by the fact that success probabilities are bounded; instead, they are caused by interactions between the types of problem solvers.

[14] Note that hiring could satisfy asymptotic diversity but not the property of no dominated hires. If so, the optimal group would include all of the undominated types as well as some dominated ones.

[15] Just because the manager's hiring problem is easy need not imply that the problem that the team will try to solve is easy. The probability that a group of size n will solve the problem could be arbitrarily close to zero.

[16] The optimal addition to a given group, identified in Proposition 1, is generically unique. In knife-edge cases, there are multiple such additions.

[17] They fall short in a strong way: Propositions 4 and 5 both show that nonoptimal groups cannot produce the first-best outcome in or out of equilibrium.

[18] Recall that if the parent wants the children to implement the fair (i.e., equal) split then she tells one child to cut the cookie and the second to select a piece. If the children's utility is strictly increasing in the size of their piece then in the unique equilibrium the cutter divides the cookie in half. The manager in our model has substantively different preferences but her exploitation of the conflict among the individuals is similar to the parent's exploiting the children's conflict over the cookie.

[19] Note that here s_2 is dominated by just one type. If it were dominated by a *set* of other types then the Ricardian property need not hold. Suppose, for example, that s_2 is a jack-of-all-trades: Worse than s_1 regarding t_1 but better on t_2, and worse than some third type, s_3, regarding t_2 but better on t_1. Here s_2 is dominated by the combination of s_1 and s_3. This feature could rob s_2 of its strategic comparative advantage in t_2: If s_3 is sufficiently bad at t_1 then his promise to try t_2 will be credible. Thus a team of (s_1, s_3) will deliver tool diversity in equilibrium and the manager will get what she wants without putting a dominated type on the team.

[20] Remark 5 does not presume complete skill specialization; there can be generalists in the set of types. But every generalist type is dominated by the set of experts whose repertoires collectively span that generalist's toolkit.

[21] That advantage exists even if we assume that problem solvers might apply tools correctly to only a part of a problem.

[22] In more complex versions some of the problems would be profitable opportunities discovered by enterprising managers (Hsieh, Nickerson, & Zenger, 2007) and hence endogenous.

[23] We suspect, however, that only strategic problems can prompt a rational manager to hire a type that is pairwise dominated. A type that is dominated though not pairwise so can be a useful jack-of-all-trades in a small problem-solving team that is unsure what problem it will face; one that is pairwise dominated lacks this decision theoretic justification.

[24] The inequalities are strictly satisfied of $p_{t_1}=0.5$, $p_{t_2}=p_{t_3}=p_{t_4}=0.3$, and $b=0.25$. Note that as p_{t_1} increases and b decreases in size, the inequalities become satisfied on an increasingly larger domain.

[25] If $h_{s_2}(t_2)$ could be exactly zero then the expected payoff of playing t_2 would obviously also be zero, whence it would follow immediately selecting t_2 is a strictly dominant strategy since (t_2) is s_2's best tool and so must have strictly positive promise. By assumption all h's are strictly positive so technically this case does not hold, but the corresponding intuition is correct.

ORCID

Jonathan Bendor http://orcid.org/0000-0002-6066-0815

REFERENCES

Austen-Smith, D., & Banks, J. S. (1996). Information aggregation, rationality, and the condorcet jury theorem. *The American Political Science Review*, *90*, 34–45.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2006). *The polarization of the U.S. labor market* (NBER Working Paper No. 11986).

Bachrach, Y., Graepel, T., Kasneci, G., Kosinski, M., & Van Gael, J. (2012). *Crowd IQ—Aggregating opinions to boost performance.* AAMAS proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AA- MAS 2012).

Bell, D. (1973). *The coming of post-industrial society*. New York: Basic Books.

Bock, L. (2015). *Work rules!: Insights from inside google that will transform how you live and lead*. New York, NY: Twelve.

Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and problem uncertainty in innovation contests: an empirical analysis. *Management Science*, *57*(5), 843–863.

Conner, K. R. (1991). A historical comparison of resource-based theory and five schools of thought within industrial organization economics: Do we have a new theory of the firm? *Journal of Management*, *171*, 121–154.

Conner, K. R., & Prahalad, C. K. (1996). A resource-based theory of the firm: Knowledge versus opportunism. *Organization Science*, *7*(5), 477–501.

Cooke, N. J., & Hilton, M. L. (2015). (Eds.), *Enhancing the effectiveness of team science*. Washington DC: National Research Council, National Academies Press.

Demsetz, H. (1988). The theory of the firm revisited. *Journal of Law Economics and Organization*, *4*(1), 141–162.

Ericsson, K. A., Krampe, R., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Feltovich, P., Prietula, M., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In Ericsson, K., Charness, N., Feltovich, P., & Hoffman, R. (Eds.), *Cambridge handbook of expertise and expert performance*. New York: Cambridge Press.

Florida, R. (2002). *The rise of the creative class: And how it's transforming work, leisure, community and everyday life*. New York: Basic Books.

França, L., Carrilho, E., & Kist, T. B. L. (2002). A review of DNA sequencing techniques. *Quarterly Reviews of Biophysics*, *35*(2), 169–200.

Freeman, R., & Huang, W. (2014). *Collaborating with people like me: Ethnic co-authorship within the U.S* (NBER Working Paper No.19905).

Garicano, L. (2000). Hierarchies and the organization of knowledge in production. *Journal of Political Economy*, *108*, 874–904.

Garicano, L., & Rossi-Hansberg, E. (2006). Organization and inequality in a knowledge economy. *Quarterly Journal of Economics*, *121*, 1383–1435.

Holmstrom, B. (1982). Moral hazard in teams. *Bell Journal of Economics*, *13*, 324–340.

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, *101*(46), 16385–16389.

Hsieh, C., Nickerson, J., & Zenger, T. (2007). Opportunity discovery, problem solving and a theory of the entrepreneurial firm. *Journal of Management Studies*, *44*, 1255–1277.

Jones, B. F., Uzzi, B., & Wuchty, S. (2008). Multi-university research teams: Shifting impact, geography and social stratification in science. *Science*, *322*, 1259–1262.

Jon Kleinberg, Maithra Raghu. (2018) Team Performance with Test Scores. *ACM Transactions on Economics and Computation*.

Jeppesen, L. B., & Lakhani, K. R. (2010). Marginality and problem-solving effectiveness in broadcast search. *Organization Science*, *21*, 1016–1033.

Kleinberg, J., & Raghu, M. (2018). Team performance with test scores, ACM Transactions on Economics and Computation (working paper). Cornell University School of Information.

Lakhani, K. R., & Jeppesen, L. B. (2007). Getting unusual suspects to solve R&D puzzles. *Harvard Business Review*, *85*(5).

Lazear, E. P. (2009). Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*, *117*(5), 914–940.

March, J., & Simon, H. (1958). *Organizations*. New York: John Wiley and Sons.

Marcolino, L. S., Jiang, A. X., & Tambe, M. (2013). *Multi-agent team formation: Diversity beats strength?* Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China.

Marshack, J., & Radner, R. (1972). *Economic theory of teams*. New Haven: Yale University Press.

McAfee, P., & McMillan, J. (1991). Optimal contracts for teams. *International Economic Review*, *32*(3), 561–77.

Nash, J. (1951). Non-cooperative games. *Annuals Mathematics*, *54*, 286–295.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Upper Saddle River NJ: Prentice Hall.

Nickerson, J., & Zenger, T. (2004). A knowledge-based theory of the firm: The problem-solving perspective. *Organization Science*, *15*, 617–632.

Page, S. E (2008). *The difference*. Princeton NJ: Princeton University Press.

Powell, W., & Snellman, K. (2004). The knowledge economy. *Annual Review of Sociology*, *30*, 199–220.

Prat, A. (2002). Should a team be homogeneous? *European Economic Review*, *46*, 1187–1207.

Schilling, M. A., & Green, E. (2011). Recombinant search and breakthrough idea generation: An analysis of high impact papers in the social sciences. *Research Policy*, *40*(10), 1321–1331.

Shi, X., Adamic, L. A., Tseng, B. L., & Clarkson, G. S. (2009). The impact of boundary spanning scholarly publications and patents. *PLoS One*, *4*(8), 6547.

Simon, H. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, *106*, 467–82.

Singh, J., & Fleming, L. (2010). Lone inventors as sources of breakthroughs: Myth or reality? *Management Science*, *56*(1), 41–56.

Terwiesch, C., & Xu, Y. (2008). Innovation contests, open innovation, and multiagent problem solving. *Management Science*, *54*(9), 1529–1543.

Uzzi, B., Mukherjee, S., Stringer, M., & Jones, B. (2013). Atypical combinations and scientific impact. *Science*, *342*, 468–471.

Wolff, E. N. (2006). The growth of information workers in the U.S. economy, 1950–2000: The role of technological change, computerization, and structural change. *Economic System Research*, *18*(3), 221–255.

Wuchty, S., Jones, B. F., & Uzzi, B. (2007). The increasing dominance of teams in the production of knowledge. *Science*, *316*(5827), 1036–1039.

How to cite this article: Bendor J, Page SE. Optimal team composition for tool-based problem solving. *J Econ Manage Strat*. 2019;28:734–764. https://doi.org/10.1111/jems.12295

APPENDIX A

Proof of Observation 1. *The formal proof relies on Proposition 1 which states that the optimal addition to a given group G is the type s_i that maximizes*

$$P^{+}(s_i, G) = \Pr\left[G\left(\text{FAIL}, \neg t^*_{s_i}(G)\right)\right] \cdot h_{s_i}\left(t^*_{s_i}(G)\right) p\left(t^*_{s_i}(G)\right).$$

By assumption, type $h_{s_1}(t_k) \geq h_{s_i}(t_k)$ for all s_i. The result follows. □

Proof of Corollary 1. To prove part (i), note first that the conclusion holds if $\max\{h_{s_{j'}}(t_k)\} = 0$ because in that case $s_{j'}$ is completely worthless whereas $P^{+}(s_i, G)$ is strictly positive. Since $P^{+}(s_i, G)$ must continue to exceed $P^{+}(s_{j'}, G)$ if j''s skill at using any tool in his toolkit is sufficiently close to zero, the result follows. To prove part (ii) construct a G' such that for every $t_k \in T(s_i)$ there is a pure specialist in G', that is, some type s_k for which $T(s_k) = \{t_k\}$. If $h_{s_k}(t_k) = 1$ for all such pure specialists then s_i is completely worthless to G' because every tool that s_i possesses will be tested with probability one by this group. Hence if the $h_{s_k}(t_k)$'s are sufficiently close to one then $P^{+}(s_i, G')$, though strictly positive, can be arbitrarily close to zero, whence $P^{+}(s_j, G') > P^{+}(s_i, G')$. □

Proof of Proposition 2. Parts (i) and (ii) are proved in the text. Here we prove the other three parts. We start with (iii). Under centralized tool choice, the manager can assign any member of the group a specific tool. Therefore, if the manager hires a type s_j and assigns it tool t_k, then s_j must be best at applying tool t_k. This proves (iii).

The proof of (iv) is by contradiction. Suppose there is some group size, $n^{\dagger}$, when the manager will for the first time kick out a type, say s_i, whom she hired earlier. Let us use n^* to label the smaller group size at which the manager (a) added an s_i to the group and (b) for sizes $n^* + 1, \ldots, n^{\dagger} - 1$ she did not add any s_i's to the group. Suppose that for $|G| = n^*$ the manager tells the next type hired, s_i to use tool t_k. The pair, (s_i, t_k), must satisfy the optimality condition of Proposition 2. Proposition 2 immediately implies that s_i must be the expert of t_k; otherwise, the planned pair of (s_i, t_k) would not be optimal for $|G| = n^*$. By assumption, for $|G| = n^{\dagger}$, the manager removes the s_i hired at $|G| = n^*$, replacing him with (say) s_j. Because s_i is the expert at t_k, the manager must be planning to tell the new member, s_j, to try some other tool, say t_ℓ. Hence, for $|G| = n^{\dagger}$ it must be true that the pair (s_j, t_ℓ) increases the probability of group success more than does the (s_i, t_k) pair.

But if this is true for $|G| = n^{\dagger}$ then it must have been true for $|G| = n^*$. Given that s_i is the expert at t_k, none of the other types added in-between n^* and $n^{\dagger}$ are using tools t_k. Hence the probability that t_k will be tested by the group is unchanged from $|G| = n^*$ to $n^{\dagger}$. In contrast, the probability that t_ℓ will be tested by the group may have gone up, if a s_j has been added in that interim and the manager plans to order that s_j try t_ℓ. Or the test-probability is unchanged, if no (s_j, t_ℓ) pair has been added in the interim. In either case, given that the probabilities of different tools are succeeding are independent of each other, if the marginal contribution of (s_j, t_ℓ) exceeds that of (s_i, t_k) for $|G| = n^{\dagger}$ then the same strict inequality must have held when $|G| = n^*$. But then the choice of s_i given the group-size of n^* would not have been optimal after all, and we have a contradiction.

To prove (v), we first note that if t_k is a new tool, then a straightforward argument shows that the manager assigns the same tools to the existing members of G after a problem solving using tool t_k is added to G. Given a type s_i and a tool $t^*_{s_i}$ that has been assigned to the group, the increase in the group's probability of solving the problem equals $\Pr[G(\text{FAIL}, \neg t^*_{s_i}(G)] \cdot h_{s_i}(t^*_{s_i} p(t^*_{s_i}))$. This converges to zero as the number of type s_i's using tool $t^*_{s_i}$ increases. In contrast, given there exists a finite number of tools and none solves the problem with probability one, the increase in the probability of solving the problem given a type s_j who is best at tool t_k which is not used by G equals $\Pr[G(\text{FAIL } \neg t^*_{s_i}(G)] \cdot h_{s_j}(t_k) p(t_k)$, which is strictly positive. □

Proof of Remark 1. The proof is by construction. Assume that tool t_1 solves the problem with probability 0.8 and t_2 solves it with probability 0.2 and that tool t_3 solves the problem with probability ϵ. Suppose that G consists of one specialist who always possesses tool t_1 and one specialist who possesses tool t_2 who can correctly apply that tool half the time. Thus, G solves the problem with probability $0.8 + (0.2)(0.1) = 0.82$. Assume that type s_i applies tool t_1 with $h_{s_i}(t_1) = 0.15$ and tool t_3 with probability $h_{s_i}(t_3) = 0.07/0.18$ and that type s_j is a specialist who applies tool t_2 with $h_{s_j}(t_2) = 0.4$. By construction type s_i has an ability of 0.15 and is more diverse than type s_j, which has an ability of 0.08. The manager will assign s_i tool t_3, so the probability that G and s_i solve the problem equals

$$0.82 + (0.18)\cdot\frac{0.007}{0.18}\cdot\epsilon = 0.82 + 0.007\epsilon.$$

The manager will assign s_j tool t_2, so the probability that G and s_j solve the problem equals

$$0.8 + 0.2[(0.5)(0.2) + (0.5)(0.4)(0.2)] = 0.828.$$

Thus, $P^+(s_i, G) = 0.07\epsilon$ and $P^+(s_j, G) = 0.08$. Note that this example requires that type s_i is not best at its best tool. If both types are best at their best tools then hiring by ability will be optimal. □

Proof of Proposition 3. Follows from Nash (1951). Choose a type s_i and assume any set of strategies by the individuals of the other types. This can be thought of as the environment for the individuals of type s_i. By Nash (1951), individuals of type s_i have a symmetric best response. Therefore, a symmetric equilibrium exists. To prove the second part of the proposition, recall that $p_k \geq p_{k'}$. In a symmetric by type equilibrium, the payoff from choosing tool t_k and tool k' must be equal. Assume there exist $\tilde{N} + 1$ individuals of a given type, and suppose that the remaining $\tilde{N}$ individuals play the mixed strategy equilibrium. Conditional on r_k of those individuals choosing tool t_k and $r_{k'}$ choosing tool $k + 1$, let $\rho_i(\ell)$ denote the probability that i of the ℓ individuals not choosing tool t_k or tool $t_{k'}$ individuals solve the problem. These individuals need not be of this same type. Further let $\hat{q} = (1 - q_{t_k} - q_{t_{k'}})$ The payoff from choosing tool t_k, $\pi_k(\vec{q})$ can be written as follows:

$$\begin{aligned}
\pi_k(\vec{q}) = {} & p_k h_{ik} p_{k'} h_{ik'}\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(1+\frac{b}{1+r_k+r_{k'}+i}\right)\right]\\
& + p_k h_{ik}(1-p_{k'}h_{ik'})\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(1+\frac{b}{1+r_k+i}\right)\right]\\
& + (1-p_k h_{ik})p_{k'}h_{ik'}\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=1}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\right]\\
& - (1-p_k h_{ik})p_{k'}h_{ik'}\left[\sum_{r_k=0}^{\tilde{N}}\sum_{\ell=0}^{\tilde{N}-r_k}\frac{\tilde{N}!}{r_k!\ell!}q_{t_k}^{r_k}\hat{q}^{\ell}\rho_0(\ell)\right]\\
& + (1-p_k h_{ik})(1-p_{k'}h_{ik'})\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=1}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\right].
\end{aligned}$$

The fourth term subtracts off the case where no one chose tool k' and none of the other tools solves the problem. The expression for the payoff from playing k' can be written similarly. It follows that the payoff from choosing tool t_k equals the payoff from choosing tool k' if and only if:

$$p_k h_{ik}(1-p_{k'}h_{ik'})\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(1+\frac{b}{1+r_k+i}\right)\right]$$
$$+(1-p_k h_{ik})p_{k'}h_{ik'}\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\right]$$
$$-(1-p_k h_{ik})p_{k'}h_{ik'}\left[\sum_{r_k=0}^{\tilde{N}}\sum_{\ell=0}^{\tilde{N}-r_k}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}\hat{q}^{\ell}\rho_0(\ell)\right]$$
$$=p_{k'}h_{ik'}(1-p_k h_{ik})\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(1+\frac{b}{1+r_{k'}+i}\right)\right]$$
$$+(1-p_{k'}h_{ik'})p_k h_{ik}\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\right]$$
$$-(1-p_{k'}h_{ik'})p_{k'}\left[\sum_{r_{k'}=0}^{\tilde{N}}\sum_{\ell=0}^{\tilde{N}-r_{k'}}\frac{\tilde{N}!}{r_{k'}!\ell!}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_0(\ell)\right],$$

which can be simplified as

$$p_k h_{ik}(1-p_{k'}h_{ik'})\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(\frac{b}{1+r_k+i}\right)+\sum_{r_{k'}=0}^{\tilde{N}}\sum_{\ell=0}^{\tilde{N}-r_{k'}}\frac{\tilde{N}!}{r_{k'}!\ell!}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_0(\ell)\right]$$
$$=p_{k'}h_{ik'}(1-p_k h_{ik})\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(\frac{b}{1+r_{k'}+i}\right)\right.$$
$$\left.+\sum_{r_k=0}^{\tilde{N}}\sum_{\ell=0}^{\tilde{N}-r_k}\frac{\tilde{N}!}{r_k!\ell!}q_{t_k}^{r_k}\hat{q}^{\ell}\rho_0(\ell)\right]$$

Given $p_k h_{ik} > p_k h_{ik}$, $p_k h_{ik}(1-p_{k'}h_{ik'}) > p_{k'}h_{ik'}(1-p_k h_{ik})$. It suffices to show that

$$\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(\frac{b}{1+r_k+i}\right)+\sum_{r_{k'}=0}^{\tilde{N}}\sum_{\ell=0}^{\tilde{N}-r_{k'}}\frac{\tilde{N}!}{r_{k'}!\ell!}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_0(\ell)\right]$$
$$<\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(\frac{b}{1+r_{k'}+i}\right)+\sum_{r_k=0}^{\tilde{N}}\sum_{\ell=0}^{\tilde{N}-r_k}\frac{\tilde{N}!}{r_k!\ell!}q_{t_k}^{r_k}\hat{q}^{\ell}\rho_0(\ell)\right]$$

implies $q_{t_k} > q_{t_{k'}}$. We will show that $q_{t_k} < q_{t_{k'}}$ implies the opposite inequality. Suppose that $q_{t_k} < q_{t_{k'}}$. The rightmost term of the first expression will then exceed the rightmost term of the second expression. *Therefore, it suffices to show that* $q_{t_k} < q_{t_{k'}}$ *implies*

$$\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(\frac{b}{1+r_k+i}\right)\right]$$
$$>\left[\sum_{r_k=0}^{\tilde{N}}\sum_{r_{k'}=0}^{\tilde{N}-r_k}\sum_{\ell=0}^{\tilde{N}-r_k-r_{k'}}\sum_{i=0}^{\tilde{N}-r_k-r_{k'}}\frac{\tilde{N}!}{r_k!r_{k'}!\ell!}q_{t_k}^{r_k}q_{t_{k'}}^{r_{k'}}\hat{q}^{\ell}\rho_i(\ell)\left(\frac{b}{1+r_{k'}+i}\right)\right].$$

Notice that the values of r_k and $r_{k'}$ both range from 0 to Ñ. Therefore, the comparison depends on the relative weights assigned to each term. Given $q_{t_k} < q_{t_{k'}}$, *the values of* $r_{k'}$ *are biased towards higher values, completing the proof.* □

Proof of Corollary 2. Let z denote the number of problem solvers of some other type other than s_i or the new type who solve the problem. We show that the result holds for any z. The assumption of nonoverlapping toolkits implies additive separability of the payoff function for any one type as a function of the number of problem solvers of another type who solve the problem. Therefore, if the result holds for any one z, it also holds for a probability distribution over z.

The formal proof relies on Lemma 1 stated below. Assume that there exists Ñ individuals of type s_i, who choose tools t_k and $t_{k'}$ with probabilities $q_{s_i}(t_k)$ and $q_{s_i}(t_{k'})$ such that $q_{s_i}(t_k) > q_{s_i}(t_{k'})$. Define the *difference in the probabilities of exactly* ℓ *winners*, $D_{s_i,\ell}(t_{k'}, t_k)$, to be the probability that exactly ℓ of the s_i's choose tool $t_{k'}$ times the probability that tool $t_{k'}$ was successfully applied by a single type s_i problem solver and tool $t_{k'}$ solved the problem, minus the probability that exactly ℓ of the s_i's choose t_k. times the probability that tool t_k was successfully applied by a single type s_i problem solver and tool t_k successfully solved the problem. □

Lemma A1. *Given* $q_{s_i}(t_k) > q_{s_i}(t_{k'})$*, there exists an* ℓ' *such that for any* $\ell \leq \ell'$, $D_{s_i,\ell}(t_{k'}, t_k) \geq 0$ *and for any* $\ell \geq \ell'$, $D_{s_i,\ell}(t_{k'}, t_k) \leq 0$

Proof of Lemma 1. The probability of exactly ℓ of the s_i choosing tool t_k equals $\binom{\tilde{N}}{\ell} q_{t_k}^{\ell}(1-q_{t_k})^{(\tilde{N}-\ell)}$. The probability of exactly ℓ of the s_i choosing tool $t_{k'}$ equals $\binom{\tilde{N}}{\ell} q_{t_{k'}}^{\ell}(1-q_{t_{k'}})^{(\tilde{N}-\ell)}$. The latter exceeds the former if and only if

$$\left(\frac{q_{t_{k'}}}{q_{t_k}}\right)^{\ell} \geq \left(\frac{1-q_{t_k}}{1-q_{t_{k'}}}\right)^{\tilde{N}-\ell}$$

Given $q_{s_i}(t_k) > q_{s_i}(t_{k'})$, it follows that there exists an ℓ' such that the left-hand side exceeds the right-hand side if and only if $\ell \leq \ell'$. The result will still hold if we multiply the left-hand side of the inequality by a constant, though the value of ℓ' may change. Let that constant equal $(p_{k'} h_{ik'}/p_k h_{ik})$. We can write $D_{s_i,\ell}(t_{k'}, t_k)$ as follows:

$$D_{s_i,\ell}(t_{k'}, t_k) = \binom{\tilde{N}}{\ell}\Big(p_{k'} h_{ik'} q_{t_{k'}}^{\ell}(1-q_{t_{k'}})^{(\tilde{N}-\ell)} - p_k h_{ik} p_k h_{ik} q_{t_k}^{\ell}(1-q_{t_k})^{(\tilde{N}-\ell)}\Big).$$

This completes the proof of the lemma. To prove the corollary, assume Ñ problem solvers of type s_i. Take the actions of Ñ − 1 of those problems solvers as given according to a symmetric by type mixed strategy equilibrium $\vec{q}^{\,*}(G)$. Let $Q_{t_k}^{\ell}$ denote the probability that exactly ℓ of Ñ − 1 type s_i problem solvers choose tool t_k in equilibrium. Note that expected payoff to a type s_i conditional on using tool t_k when exactly $\ell - 1$ other type s_i problem solver uses tool t_k equals $1 + (b/Z + \ell)$. Consider the tool choice by the Ñth problem solver of type s_i. If tools t_k and $t_{k'}$ are used in equilibrium with strictly positive probability by type s_i, then each must have the same expected value. Let $\Pr(G \backslash s_i, X)$ equal the probability that the group solves the problem without s_i. The following equation must be satisfied in equilibrium.

$$p_k h_{ik} \sum_{\ell=0}^{\tilde{N}-1} Q_{t_k}^{\ell}\left(1 + \frac{b}{Z+\ell}\right) + (1 - p_k h_{ik})\Pr(G \backslash s_i, X) = p_{k'} h_{ik'} \sum_{\ell=0}^{\tilde{N}-1} Q_{t_{k'}}^{\ell}\left(1 + \frac{b}{Z+\ell}\right) + (1 - p_{k'} h_{ik'})\Pr(G \backslash s_i, X).$$

Recall that if there exists no overlap then $p_k > p_{k'}$ implies that $q_{t_k}^*(G, s_i) > q_{t_{k'}}^*(G, s_i)$. Therefore, by Lemma A1 there exists an ℓ' such that for any $\ell \leq \ell'$, $D_{s_i,\ell}(t_{k'}, t_k) \leq 0$ and for any $\ell \geq \ell'$, $D_{s_i,\ell}(t_{k'}, t_k) \geq 0$, where $D_{s_i,\ell}(t_k, t_{k'})$ equals the probability that exactly ℓ problem solvers use tool $t_{k'}$ and a single problem solver of type s_i successfully applies tool $t_{k'}$ minus the probability that exactly ℓ use tool t_k and a single problem solver of type s_i successfully applies tool $t_{k'}$. We can rewrite the equilibrium equation as follows:

$$(p_k h_{ik} - p_{k'} h_{ik'}) \Pr\left[\left(1 + \frac{b}{Z + \ell}\right) - \Pr(G \backslash s_i, X)\right] = \sum_{\ell=1}^{\tilde{N}-1} D_{s_i,\ell}(t_{k'}, t_k)\left(1 + \frac{b}{Z + \ell}\right).$$

If we add a problem solver that uses nonoverlapping tools, and if that problem solver does not solve the problem, then there is no effect on the payoffs from playing either tool. However, if the new problem solver solves the problem then the left-hand side of the equation becomes zero because the problem is always solved. Note then that if b is sufficiently large then the left-hand side will be positive. Therefore, we have that

$$0 > \sum_{\ell=1}^{\tilde{N}-1} D_{s_i,\ell}(t_{k'}, t_k)\left(1 + \frac{b}{Z + \ell}\right).$$

To simplify notation let $w_\ell = D_{s_i,\ell}(t_{k'}, t_k)$. By assumption $w_\ell > 0$ off $\ell \leq \ell'$. Therefore, we can write the previous inequality as

$$\sum_{\ell=\ell'+1}^{\tilde{N}-1} - w_\ell\left(1 + \frac{b}{Z + \ell}\right) > \sum_{\ell=1}^{\ell'} w_\ell\left(1 + \frac{b}{Z + \ell}\right).$$

Given this inequality and given the signs on the w'_ℓ, the following inequality holds because the decrease in each term on the right strictly exceeds the decreases of each term on the left,

$$\sum_{\ell=\ell'+1}^{\tilde{N}-1} - w_\ell\left(1 + \frac{b}{Z + 1 + \ell}\right) > \sum_{\ell=1}^{\ell'} w_\ell\left(1 + \frac{b}{Z + 1 + \ell}\right).$$

Notice that these payoffs correspond to exactly one more individual solving the problem. This inequality therefore implies that the expected payoff from choosing tool t_k strictly exceeds that of playing tool $t_{k'}$, so that in equilibrium t_k must be played with strictly higher probability. □

Proof of Proposition 4. Given that s_1 dominates all the other types, by Proposition 1 the optimal group must consist only of the best type. Hence any group which contains any type other than s_1 is suboptimal. Hence such a group cannot implement the first-best outcome, thus establishing part (b) of the result.

In contrast, we now show that all Nash equilibria of the homogeneous s_1 group achieve the first-best tool vector. (For convenience we assume it to be unique; the proof extends in a straightforward but tedious way to the case of multiple first-best tool vectors.) Call this vector, which the manager would select under centralized tool choice, T^*. Consider any vector $T^\bullet \neq T^*$. We need to show that any such $T^\bullet$ is not an equilibrium under decentralized tool choice by the homogeneous group of the best type. To do this we initially assume that b is exactly equal to one; this will be relaxed below.

Either $T^\bullet$ has gaps (a more promising tool, t_j, is not selected by anyone but a less promising tool, t_k, is) or it does not. If it does have gaps then clearly $T^\bullet$ cannot be an equilibrium because the agent picking t_k would increase the probability of group success, $P(G)$, and therefore his payoff, by switching from t_k to t_j.

So now consider a $T^\bullet$ without gaps. We need to show that despite this similarity with T^* (which the manager would obviously design to be gapless), this kind of $T^\bullet$ is not an equilibrium either. Since $T^\bullet \neq T^*$, there must be at least one tool, say t_j, which has too many agents selecting it, and hence some other tool, say t_k, which has too few: $P(G)$ would rise if one agent (possibly more) moved from the former to the latterer. But given that $b = 1$, agents care only about group success, so if agent would increase $P(G)$ by switching from t_k to j then playing the former is not a best response to the choices of the other $n - 1$ agents: Hence $T^\bullet$ is not an equilibrium.

The above logic continues to hold if b is sufficiently close to one. And since there are finitely many $T^\bullet$'s, there exists a $\underline{b} < 1$ such that any $b > \underline{b}$ is sufficiently close to one for all of them. This establishes part (a). □

Proof of Proposition 5. By Proposition 3, an optimal group must satisfy competency ordering. To complete the identification of the first-best outcome, we initially assume that $\min(h_{s_i}(t^*(s_i)))$ exactly equals one. Given this, it is obviously suboptimal to have any type's best tool selected by more than one member of the group. Further, it is inefficient to overlook any best tool. More precisely, it is suboptimal if in any group there is some tool, t_j, that is not selected even though it is the best tool of some type (say s_j) while some other tool, t_k, is selected even though it is not the best tool of any type. This holds because best tools are more promising than the other tools: that is, $h_{s_j}p(t_j) > h_{s_k}p(t_k)$, where s_j and s_k are the experts at deploying t_j and t_k, respectively. Finally, best tools are exclusive, so if the manager wants some best tool, say t_j, selected in Stage 2 then she must put its expert, s_j, on the team in Stage 1. Together, these properties imply that if $G(n)$ is no bigger than the number of types then the manager's design problem under centralized tool choice is simple: In Stage 1 she selects the n most able types and in Stage 2 she tells each one to try his best tool. And since this design is strictly better than any other design when $\min(h_{s_i}(t^*(s_i)))$ is exactly one, it continues to hold if $\min(h_{s_i}(t^*(s_i)))$ is sufficiently close to one.

Turning now to strategic tool choice, consider a group G of size n which the manager in Stage 1 stocked with the n most able types. (The proof is trivial for $n = 1$ so assume that $n > 1$.) Each type's best tool is exclusive. This implies that for all $b \geq 0$, each agent has a strictly dominant strategy of selecting his best tool, $t^*(s_i)$. Hence G produces a unique Nash equilibrium in which every member of the group plays his strictly dominant strategy, that is, tries his best tool. But this is exactly the description of G and its unique tool-selection equilibrium: Everyone in G is of a different type; G satisfies the property of no competency skipping; in equilibrium everyone tries his (exclusive) best tool. Thus this Nash equilibrium implements the first-best outcome, establishing part (a) of the result.

Finally, consider any group $G' \neq G$. We now show that any such G' cannot generate the first-best tool vector. Either G' satisfies competency ordering or it does not. If it does not then it follows immediately from Proposition 2 that G' is suboptimal and hence cannot implement the first-best outcome in any tool vector.

Alternatively, G' does satisfy competency ordering. Either each type is a singleton in the group or it is not. If each type is a singleton then G' is, in fact, G, so there's nothing to compare. Alternatively, there is at least one type, say s_i, with multiple members in the group. But this is not the optimal $G^*(n)$ identified in paragraph one of this proof, thus establishing part (b). □

Proof of Proposition 6. We prove each part in turn.

- **(i)** We use contradiction to prove that all problem solvers of the same type pick the same tool. Suppose instead that two s_i's chose different tools, t_j and t_k; then, since agents of the same type are clones of each other there must be at least two equilibria, because the choices of (t_j, t_k) and (t_k, t_j) are strategically equivalent. But by assumption there is a unique Nash equilibrium in this case. hence we have reached a contradiction, so the supposition that two agents of the same type have chosen different tools must be wrong.
- **(ii)** By contradiction. From (i) we know that all agents of the same type choose the same tool. Suppose that all the s_i types choose tool t_k, despite the fact that the expert at that tool is s_j. Obviously, then, the status quo group cannot be a first-best composition of types: The s_i's should be replaced by the s_j's.
- **(iii)** From (ii) we know that each type is an expert at the tool it selects. The conclusion follows given that each tool has a unique expert.
- **(iv)** Suppose that the best tool for every type is t_1. By construction, the expert at t_1 must be the best type, s_1. Hence from (ii) we know that s_1 picks t_1, and from (ii) we know that no other type picks s_1. Therefore, any type that is not the best type picks a tool that is not its best tool. □

Proof of Remark 2.

- **(i)** Consider first the case of $b = 1$. Here players rank-order outcomes by the probability of group success, $P(G)$. It is given that there is a unique tool vector (implemented by the corresponding experts) that maximizes this probability. Hence if anyone deviates from this tool-player combination he will reduce $P(G)$. Since $b = 1$ this will reduce his expected utility, which shows that the original configuration of tools and agents was a (strict) Nash equilibrium. Because a deviation strictly reduces the deviator's expected utility for b exactly equal to one, it continues to do so for b sufficiently close to one.

(ii) Any outcome produced by a different tool vector or by the first-best tool vector where some tool is tried is applied by a nonexpert generates a strictly lower probability of group success than does the first-best tool vector implemented by experts. For $b = 1$ this probability completely determines payoffs, so strict Pareto ordering follows immediately. The usual continuity argument for b close to one finishes the proof. □

Proof of Corollary 3. Suppose initially that $b = 1$. We solve the problem backwards. Suppose that $n - 1$ players have chosen tools. The last player will choose the tool that maximizes $P(G^*(n))$, the probability that this group, $G^*(n)$, succeeds, given the first $n - 1$ choices. Anticipating this, the second-to-last chooses the tool that induces the last player to choose the tool that, together, maximizes $P(G^*(n))$. This logic works all the way back to the first player who, anticipating all subsequent choices, chooses a tool that triggers a sequence that maximizes $P(G^*(n))$.
(Note that this doesn't presume that there is a unique such sequence. Generically, that is, for $b < 1$, there is a unique sequence, but if b exactly equals one then players are indifferent about who gets to select which tool, provided that in every case every tool is chosen by its corresponding expert.)

Since the relevant inequalities are strict for $b = 1$, by continuity they hold for b sufficiently close to one. □

Proof of Proposition 7. Competency loading follows using the same example as in the centralized tool choice case. Competency ordering follows if each type is best at its best tool because the first member of each type will choose its best tool. Note that the proof requires that the group is optimal. The manager will be careful not to select better types that have an incentive to apply a lesser type's best tool. If a type is not best at its best tool, then competency ordering fails using the same example as in the centralized case.

We next show that monotonicity can be violated with nonoverlapping toolkits. We first provide a sketch and then give the formal proof. Suppose that the best type, s_1 possesses two tools, t_1 and t_2, and is perfect at applying each. Under a large set of parameter values, the optimal two-person group has two agents of the best type who choose different tools. Next, suppose the types s_2 and s_3, are each perfect at tools t_3 and t_4 and that these tools are approximately as likely to solve the problem as the best type's weaker tool. For a range of parameter values, a group consisting of two s_1's and one s_2 both of the agents of the best type will both choose their best tool. They do so because the s_2 has some probability of solving the problem. If this occurs, the private benefit for one of the s_1's from choosing t_2 falls more than had that problem solver chosen the best tool. The reason: The loss from getting the entire benefit to an even split equals one half of the benefit, while the loss from splitting the benefit two ways to splitting it three ways equals only one-sixth of the benefit. Thus, the optimal group of size three consists of one of each type.

The formal proof goes as follows. We consider the special case where $f(|C|) = 1/|C|$. Assume that if a type can apply a tool, then it can always apply it correctly. The second problem solver of type s_1 will choose tool t_2 provided the following inequality holds:

$$p_{t_1}\left(b + \frac{1-b}{2}\right) \le p_{t_1}(1 - p_{t_2})b + (1 - p_{t_1})p_2 + p_{t_1}p_{t_2}\left(b + \frac{1-b}{2}\right).$$

This can be rewritten as

$$p_{t_1}(1 - p_{t_2})\frac{1-b}{2} \le p_{t_1}p_{t_2} + p_2.$$

Note that this holds for any $p_{t_2} > 1/2$. By assumption, type s_2 only can apply tool t_3. In a group of size three consisting of two type s_1's and one type s_2, the payoff to the second type s_1 from choosing tool t_1 can be written as follows:

$$p_{t_1}(1 - p_{t_3})\frac{1+b}{2} + p_{t_1}p_{t_3}\frac{1+2b}{3} + (1 - p_{t_1})p_{t_3}b.$$

The payoff from choosing tool t_2 can be written as follows:

$$(p_{t_1}+p_{t_3}-p_{t_1}p_{t_3})(1-p_{t_2})b+(1-p_{t_1})(1-p_{t_3})p_2+p_{t_2}(p_{t_1}(1-p_{t_3})+(1-p_{t_1})p_{t_3})\frac{1+b}{2}+p_{t_1}p_{t_2}p_{t_3}\frac{1+2b}{3}.$$

It is easy to assign probabilities so that the second type s_1 in a group of two would choose tool t_2, but would switch to tool t_1 when a problem solver of type s_2 is added. For example, if $p_{t_1}=2/3$ and $b=1/2$, then the second problem solver of type t_1 will choose tool t_2 if $p_{t_2}\geq 1/3$. But if p_{t_3} is also close to 1/3, then the problem solver of type s_1 would switch to tool t_1 if a problem solver who uses tool t_3 is added to the group. If we set both p_{t_2} and p_{t_3} equal to 1/3, then using the expressions above, the payoff for a type s_1 from choosing tool t_1 equals 174/324. The payoff from choosing tool t_2 equals 169/174. Thus, if there were a third type of problem solver s_3 and a tool t_4 such that p_{t_4} was also approximately 1/3, then the optimal group of size three consists of one type s_1, one type s_2 and one type s_3.[24]
We next show that the optimal team could include an undominated type, we derive the mixed strategy equilibrium for the game with two players. □

Lemma (Equilibrium conditions for the two player mixed strategy)**.** *Given two type s_1 players with $T_1=\{t_1, t_2\}$, the symmetric mixed strategy equilibrium probability of playing tool t_1, q^*, can be written as follows:*

$$q^*=r_{12}+r_{12}\frac{(p_1-p_2)}{(p_1-p_1p_2)}\frac{b}{(2+b)},$$

where r_{12} equals the unique solution ratio of t_1 to t_2.

Proof of Lemma. Let q denote the probability that a problem solver applies tool t_1. We can solve for the incentive compatibility constraint for using the better tool as follows. Incentive compatibility requires that the payoff from using the tool t_1 given q exceeds the payoff from using tool t_2. These payoffs can be written as follows:

$$\pi_1(q)=qp_1\left(1+\frac{b}{2}\right)+(1-q)\left[p_1(1-p_2)(1+b)+p_1p_2\left(1+\frac{b}{2}\right)+(1-p_1)p_2\right],$$
$$\pi_2(q)=q\left[p_1p_2\left(1+\frac{b}{2}\right)+(1-p_1)p_2(1+b)+p_1(1-p_2)\right]+(1-q)p_2\left(1+\frac{b}{2}\right).$$

These can be simplified as

$$\pi_1(q)=(p_1+p_2-p_1p_2)+\left(p_1-\frac{p_1p_2}{2}\right)b-q\left[(p_2-p_1p_2)+(p_1-p_1p_2)\frac{b}{2}\right)\Bigg],$$
$$\pi_2(q)=p_2\left(1+\frac{b}{2}\right)+q\left[(p_1-p_1p_2)+(p_2-p_1p_2)\frac{b}{2}\right)\Bigg].$$

Therefore $\pi_1(q)\geq\pi_2(q)$ if and only if

$$(p_1+p_2-p_1p_2)+\left(p_1-\frac{p_1p_2}{2}\right)b-q\left[(p_2-p_1p_2)+(p_1-p_1p_2)\frac{b}{2}\right)\Bigg]\geq p_2\left(1+\frac{b}{2}\right)+q(p_1-p_1p_2)$$
$$+(p_2-p_1p_2)\frac{b}{2}\Bigg),$$

which reduces to

$$(p_1-p_1p_2)+(2p_1-p_2-p_1p_2)\frac{b}{2}\geq qr\left[(p_1+p_2-2p_1p_2)\left(1+\frac{b}{2}\right)\right],$$

which further reduces to

$$q \leq \frac{(p_1 - p_1p_2)(2 + b) + (p_1 - p_2)b}{(p_1 + p_2 - 2p_1p_2)(2 + b)}.$$

It will be helpful to write the incentive compatibility constraint as follows:

$$q \leq \frac{(p_1 - p_1p_2)}{(p_1 + p_2 - 2p_1p_2)} + \frac{(p_1 - p_2)b}{(p_1 + p_2 - 2p_1p_2)(2 + b)}. \quad (1)$$

The formal counterexample relies on the same construction as the informal example presented in the body of the paper. The probability that two problem solvers of type s_1 solve the problem equals $\Pr(\{s_1, s_1\}, X) = (q^*)^2p_1 + 2q^*(1 - q^*)(p_1 + p_2 - p_1p_2) + (1 - q^*)^2p_2$. As an example, set $b = 6$, $p_1 = 0.5$ and $p_2 = 0.4$. It follows that $q^* = 0.75$ and that $\Pr(\{s_1, s_1\}, X) = 056875$. Choose $p_3 = 2{\cdot}0.06875 - \epsilon$ and let p_4 be arbitrarily close to p_3. When the type s_2 is added to the group of two type s_1's, the equilibrium probability of playing t_1 increases above 0.75, and for small ϵ a group of all three types outperforms a group of two type s_1's and a type s_2. A diminishing marginal returns argument implies that two s_1's and one s_2 solve the problem with higher probability than three s_1's. This implies that the group $\{s_1, s_1, s_1\}$ solves the problem with a lower probability than the group $\{s_1, s_1, s_2\}$, which completes this part of the proof.

To prove that asymptotic diversity can be violated, let $T(s_1) = \{t_1\}$ and assume that $h_{s_1}(t_1) = 1$ and that $p(t_1) = p$. Let $T(s_2) = \{t_1, t_2\}$ and assume that $h_{s_2}(t_1) = h$, $h_{s_2}(t_1) = \theta h$, where $\theta < 1$ and that $p(t_2) = q < p$. Finally, let $T(s_3) = \{t_2\}$ and assume that $h_{s_3}(t_2) = \gamma h$ where $\gamma < \theta$.

Note first that a type s_2 added to a group of a single type s_1 will choose tool t_2 if and only if it receives a higher payoff from doing so. Its payoff from choosing tool t_2 can be written as follows:

$$u_{s_2}(t_2 \mid \{s_1\}) = p(1 - \theta hq)(1 - b) + p\theta hq\left[(1 - b) + \frac{b}{2}\right] + (1 - p)\theta hq.$$

Its payoff from choosing tool t_1 equals

$$u_{s_2}(t_1 \mid \{s_1\}) = p(1 - h)(1 - b) + ph\left[(1 - b) + \frac{b}{2}\right]$$

It follows that type s_2 should choose tool t_2 if and only if

$$p(1 - \theta hq)(1 - b) + p\theta hq\left[(1 - b) + \frac{b}{2}\right] + (1 - p)\theta hq \geq p(1 - h)(1 - b) + ph\left[(1 - b) + \frac{b}{2}\right],$$

which can be simplified as

$$ph(1 - \theta q)(1 - b) + (1 - p)h\theta q \geq ph(1 - \theta q)\left(1 - \frac{b}{2}\right).$$

Canceling like terms gives

$$(1 - p)h\theta q \geq ph(1 - \theta q)\frac{b}{2}.$$

Solving for q gives

$$q \geq \frac{pb}{\theta[2(1 - p) + b]}.$$

It follows that if q and θ are sufficiently small then a problem solver of type s_2 will choose tool t_1 and the optimal group of size 2 consists of one problem solver of type s_1 and one problem solver of type s_3.
Consider the special case where $p = 1/2$ and $\theta = 1/2$. The inequality can then be written as

$$q \geq \frac{b}{(1+b)}.$$

Note this makes intuitive sense. If there exists no collective benefit, for example, $b = 0$ and $\theta = 1/2$, then tool t_2 would have to have the same probability of solving the problem as tool t_1 for type s_2 to choose it because type s_2 would have a 50% chance of solving the problem by itself.

The manager only needs one person using tool t_1 but could hire more people who use tool t_1 to create incentives for a problem solver of type s_2 to choose tool t_2 instead of tool t_1. Given that a type s_1 always applies tool t_1 successfully, the manager would always choose a type s_1 rather than a type s_2 as it requires fewer problem solvers to align incentives.
First, we note that t suffices to show that a fixed proportion of the group would have to be of type s_1 to induce a type s_2 problem solver to choose tool t_2. To see why suppose that there exist n problem solvers of type s_3, then at least one will apply tool t_2 correctly with probability $1 - (1 - \gamma h)^n$ (this is one minus the probability that they all fail to apply the tool correctly). Suppose that a proportion $\rho < 1$ of n problem solvers are type s_2 and try tool t_2. The probability that at least one applies the tool correctly equals $1 - (1 - \theta h)^{\rho n}$.

The first expression is strictly larger than the second if and only if the following strict inequality holds:

$$(1 - \gamma h)^n < (1 - \theta h)^{\rho n}.$$

But this is equivalent to

$$(1 - \gamma h)^n < [(1 - \theta h)^{\rho}]^n.$$

Given $\rho < 1$, a γ can be chosen so that this expression always holds.

It remains to show that a fixed percentage of the problem solvers must be of type s_1. To simplify the presentation, we assume $b = 0$, that there is only a private benefit. Given n_1 type s_1's using tool t_1 and n_2 type s_2 problem solvers using tool t_2, the expected payoff for a type s_2 problem solver choosing tool t_1 can be written as the sum of two terms. The first term equals the payoff if tool t_1 solves the problem and tool t_2 does not or is not applied correctly by any of the other n_2 problem solvers. The second term equals the payoff if the problem is solved by both tools. The second term is averaged over all possible cases for the number of type s_2's who solve the problem

$$ph\left[\frac{(1-q) + q(1-\theta h)^{n_2}}{n_1 + 1} + q\sum_{\ell=1}^{n_2}(1-\theta h)^{n_2-\ell}(\theta h)^{\ell}\binom{n_2}{\ell}\frac{1}{n_1+\ell+1}\right].$$

If the problem solver of type s_2 chooses tool t_2, the expected payoff equals

$$q\theta h\left[\sum_{\ell=0}^{n_2}(1-\theta h)^{n_2-\ell}(\theta h)^{\ell}\binom{n_2}{\ell}\left[\frac{(1-p)}{\ell+1} + \frac{p}{n_1+\ell+1}\right]\right].$$

Define $B(\{S\})$ to be the expected payoff to the problem solver of type t_2 if it solves the problem and exactly the tools in the set S solve the problem. It therefore follows that the expected payoff for the type s_2 problem solver is higher for choosing tool t_2 than for choosing tool t_1 if and only if

$$q\theta h[(1-p)B(\{t_2\})+pB(\{t_1,t_2\})]\geq ph\left[\frac{(1-q)+q(1-\theta h)^{n_2}}{n_1+1}+qB(\{t_1,t_2\})\right],$$

which simplifies to

$$q\theta h(1-p)B(\{t_2\})\geq ph\left[\frac{(1-q)+q(1-\theta h)^{n_2}}{n_1+1}+q(1-\theta)B(\{t_1,t_2\})\right].$$

For example, $B(\{t_1, t_2\})$ denote the expected payoff conditional on both tools solving the problem

$$B(\{t_1,t_2\})=\sum_{\ell=0}^{n_2}(1-\theta h)^{n_2-\ell}(\theta h)^{\ell}\binom{n_2}{\ell}\frac{1}{n_1+\ell+1}.$$

Thus, $B(\{t_2\})$ denotes the expected benefit condition on only t_2 solving the problem

$$B(\{t_2\})=\sum_{\ell=0}^{n_2}(1-\theta h)^{n_2-\ell}(\theta h)^{\ell}\binom{n_2}{\ell}\frac{1}{\ell+1}.$$

This expression is difficult to evaluate because of the last term. However, *we can rewrite the expression as a binomial distribution with an extra draw to obtain the following*:

$$B(\{t_2\})=\frac{\sum_{\ell=0}^{n_2}(1-\theta h)^{n_2+1-\ell}(\theta h)^{\ell}\binom{n_2+1}{\ell}}{\theta h(n_2+1)}.$$

Notice that the top of the fraction equals a binomial distribution minus the last term. Therefore, we can simplify the expression as

$$B(\{t_2\})=\frac{1-(\theta h)^{n_2+1}}{\theta h(n_2+1)}.$$

From above, the problem of type s_2 has a higher expected payoff of choosing tool t_2 if and only if

$$q\theta h(1-p)B(\{t_2\})\geq ph\left[\frac{(1-q)+q(1-\theta h)^{n_2}}{n_1+1}+q(1-\theta)B(\{t_1,t_2\})\right].$$

Given that $q(1-\theta)B(\{t_1, t_2\})$ is strictly positive and therefore increases the payoff from choosing tool t_1. It suffices to show that a fixed proportion of the problem solvers must be of type s_1, that is, there exists a ρ such that $\rho n_1 > n_2$, in order for the following inequality to hold:

$$q\theta h(1-p)B(\{t_2\})>ph\frac{(1-q)+q(1-\theta h)^{n_2}}{n_1+1}.$$

Substituting in for $B(\{t_2\}$ and canceling an h gives

$$q\theta(1-p)\frac{1-(\theta h)^{n_2+1}}{\theta h(n_2+1)}>p\frac{(1-q)+q(1-\theta h)^{n_2}}{n_1+1}.$$

By the same logic as above, it suffices to show that $\rho n_1 > n_2$ for some $\rho > 0$ for the following expression (again, we have increased the left-hand side and decreased the right-hand side so we have made it less likely to hold):

$$q\theta(1-p)\frac{1}{\theta h(n_2+1)} > p\frac{(1-q)}{n_1+1}.$$

Canceling out the θ's from the right-hand side and rearranging terms, we obtain the following necessary condition.

$$\frac{q}{h(1-q)}\frac{1}{n_2+1} > \frac{p}{1-p}n_1+1.$$

Define ω as follows:

$$\omega = \frac{q}{h(1-q)}\frac{1-p}{p}$$

The condition can then be written as

$$\omega(n_1+1) \geq (n_2+1)$$

This implies that $\omega n_1 > n_2 + (1-\omega)$ which means that a fixed proportion must be type s_1. As an example, suppose that $p = 1/2$, $q = 1/4$, and $h = 1/3$, then $\omega = 1$ and the number of s_1's must equal the number of s_2's. □

Proof of Remark 3. Because complete dominance implies dominance, Proposition 1 implies that under centralized tool choice the manager would never put an s_2 on the team, no matter its size. But here tool choice is strategic, so the manager must anticipate what the team will do in equilibrium. We turn to this now (hence we will prove part [b] first and then turn to [a]). We initially consider the extreme cases of $b = 0$, $h_{s_1}(t_1) = 1$, and $\theta_{s_2}(t_2)$ is some ϵ close to zero.

Consider the equilibrium behavior of a pair of s_1's. Letting $EV[(s_1, t_1); (s_1, t_1)]$ denote an s_1's expected payoff if she plays t_1 against the same choice of another s_1, we get the following:

$$EV[(s_1, t_1); (s_1, t_1)] = [h_{s_1}(t_1)]^2 \cdot p(t_1)f(|C| = 2) + h_{s_1}(t_1)[1 - h_{s_1}(t_1)] \cdot p(t_1)f(|C| = 1),$$

where $f(|C| = n)$ denotes the value of the sharing function, given that n players solved the problem.
The expected payoff of responding with a choice of t_2, given that the teammate selects s_1, is as follows:

$$\begin{aligned} EV[(s_1, t_2); (s_1, t_1)] &= h_{s_1}(t_2)p(t_2)h_{s_1}(t_1)p(t_2)f(|C| = 2) \\ &\quad + h_{s_1}(t_2)p(t_2)[1 - h_{s_1}(t_1)]f(|C| = 1) \\ &\quad + h_{s_1}(t_2)p(t_2)h_{s_1}(t_1)[1 - p(t_1)]f(|C| = 1). \end{aligned}$$

The best response to a partner's choice of s_1 is to do likewise if $EV[(s_1, t_1); (s_1, t_1)] > EV[(s_1, t_2); (s_1, t_1)]$. Collecting terms and solving for $p(t_2)$, this is equivalent to

$$\frac{[h_{s_1}(t_1)]^2 \cdot p(t_1)f(|C| = 2) + h_{s_1}(t_1)[1 - h_{s_1}(t_1)] \cdot p(t_1)f(|C| = 1)}{\Psi} > p(t_2),$$

where $\Psi = h_{s_1}(t_2)\{h_{s_1}(t_1)p(t_2)f(|C| = 2) + [[1 - h_{s_1}(t_1)] + h_{s_1}(t_1)[1 - p(t_1)]]f(|C| = 1)$.

It is optimal, obviously, for an s_1 to select its best tool if its partner chooses something other than t_1. Hence, if $p(t_2) < p^* \equiv \frac{[h_{s_1}(t_1)]^2 \cdot p(t_1)f(|C| = 2) + h_{s_1}(t_1)[1 - h_{s_1}(t_1)] \cdot p(t_1)f(|C| = 1)}{\Psi}$, then selecting t_1 is the best response to either pure strategy of the other player, that is, choosing t_1 is a strictly dominant strategy in this homogeneous two-person group. Hence the unique equilibrium of this team is that both players choose t_1. This is bad news for the manager: Given that $h_{s_1}(t_1) = 1$, duplication on that tool is worthless for her.

Now consider the equilibrium behavior of a diverse two-person team: $G = (s_1, s_2)$. Given that s_1 completely dominates s_2 and $h_{s_2}(t_1) = \epsilon$, it is easy to show that the s_1 player has a strictly dominant strategy of playing her best tool, t_1. Turning to the s_2 player, the expected payoff of playing t_1 is a function of $h_{s_2}(t_2)$, no matter what the other player does, and since $h_{s_2}(t_2) = \epsilon$ we can always choose ϵ sufficiently close to zero so that choosing t_2 delivers a higher expected payoff, again regardless of what the other player does.[25] Hence this team also has a unique equilibrium in the tool-selection stage, and in it each type chooses his best tool.

To complete the equilibrium analysis, note that because all of the relevant payoff inequalities (above) hold strictly they continue to hold for b close to zero, $h_{s_1}(t_1)$ close to one, and $\theta_{s_2}(t_2)$ close to one.

With the predictions of what happens in equilibrium in mind, the manager can now figure out how to create the optimal team in Stage 1. For $h_{s_1}(t_1) = 1$, the addition of a second s_1 would add no value: The probability that a homogeneous team of the best type solves the problem is the same as a singleton s_1. In contrast, a diverse team will try t_2 as well as t_1, and since t_2 has strictly positive promise (i.e., $h_{s_2}(t_2)p(t_2) > 0$), its success probability must be strictly bigger than that of a singleton s_1, who selects only one tool (although the best one) or, equivalently, of the homogeneous team. Hence the manager prefers the diverse team, which includes the completely dominated type, to the homogeneous group. And since a homogeneous group of the completely dominated type —$G = (s_2, s_2)$—is obviously inferior to a team which includes the best type, it follows that the optimal group is (s_1, s_2), thus establishing part (a). □

Proof of Remark 4. The example used in the proof of Remark 1 applies here as well. The only difference is that type s_i would choose tool t_1 for all but the smallest values of b implying that type s_i adds no value. □

Proof of Remark 5. Fix an arbitrary $G^*(n)$. (Generically there is only one such group, but this proof also holds for multiple optimal group compositions.) Under centralized tool choice it is obviously optimal to put only experts on the team. Further, since no expert is perfect at any tool there are no optimal teams with completely superfluous members: Everyone on $G^*(n)$ is making the probability of group success strictly greater. Consequently, carrying a nonexpert cannot be free: Replacing such a member by the corresponding expert would strictly increase the probability of group success. Hence everyone in an optimal group of size n must be the expert at the tool which s/he must select to generate the corresponding optimal tool vector.

Now consider a member of $G^*(n)$ who is supposed to select tool t_k (there may be more than one such person) whom we will call player n. Suppose player n had exactly zero chance of correctly implementing any tool in his toolkit other than t_k. If that were true then he would have zero chance of getting any personal credit for solving the problem by playing any $t_j \neq t_k$. This implies that playing his best tool, t_k, must be strictly better than doing anything else, no matter what his teammates are doing; that is, if $h_{s_n}(t_j) = 0$ for all $j \neq k$ then selecting t_k is a strictly dominant strategy. Since this holds strictly, it continues to hold if all of that player's h's (other than $h_{s_n}(t_k)$) are sufficiently close to zero. □