arXiv:2601.15327v1 [cs.GT] 19 Jan 2026

规则造成不平等回报：精英网球运动员高效配置资源

Masatsugu Yoshizawa东京大学综合文化研究科生命科学系（运动科学），日本东京目黑区 153-8902

Yuta Kawamoto东京大学运动科学倡议（UTSSI），东京大学，日本东京目黑区 153-8902

Daisuke Takeshita*东京大学综合文化研究科生命科学系（运动科学），日本东京目黑区 153-8902
dtakeshita@idaten.c.u-tokyo.ac.jp

2026年1月23日

摘要

在许多竞争环境中，从教育到政治，规则并不会让努力获得同等回报，而门槛（例如成绩门槛或选举多数）会使某些时刻变得格外重要。因此，成功取决于能否高效分配有限资源。然而，要通过实证加以证明并不容易，因为努力分配很少能被观察到，反馈也往往滞后，这限制了我们对专业技能的理解。职业网球提供了一项理想的自然实验。由于球员赢得四分后每局便告结束，而输掉一局中的得分会被浪费，因此不同比分下单分的价值差异很大。因此，高效分配资源应能在不浪费得分的情况下赢得局数，并为后续比赛节省资源。这种分配体现为随比分变化的赢分概率；我们据此推导出每位球员的帕累托前沿——即赢局概率与每局预期得分之间权衡关系的理论极限。在此，我们表明，顶尖球员的表现更接近这一前沿，能更高效地将得分转化为赢得局数。最优策略会降低球员大幅落后时（例如 0–2、0–3）赢分的概率。这种做法在心理上颇为困难——意味着放弃当前这一局——但它代表了一种理性的能量保存策略。顶尖球员在接发球局中尤其表现出这一模式；在那里，赢分比在发球局中更难，因此需要大幅调整努力程度，这与博弈论预测一致。这些发现表明，精英表现源于对规则所形成的价值结构的高效适应；懂得何时放弃，或许与懂得何时竞争一样，是专业技能的根本要素。

关键词 资源分配 · 帕累托前沿 · 专业技能 · 职业网球

*通讯作者。

引言

在大多数体育项目中，得分更多的选手会获胜。网球则不同。在 2019 年温布尔登网球锦标赛男子决赛中，罗杰·费德勒的总得分比诺瓦克·德约科维奇多 14 分（218 比 204），却输掉了比赛。这一悖论源于网球的计分规则。与足球、篮球等以总得分决定胜负的体育项目不同，网球与排球和羽毛球一样，需要赢下比赛和盘。通常，一名选手赢得四分后，一局便会重置；这意味着，在输掉的一局中得到的分数不会计入后续比分，因而被浪费了 [1, 2]。这种在局这一层级上基于阈值的赢家通吃结构，会使不同比分下每分的价值有所不同，因此何时赢得分数比赢得多少分 [3–7]更重要。把精力投入这些被浪费的时刻是低效的；因此，可以将网球视为一个资源优化问题：决定在什么比分下投入有限资源——精力与专注力。

造成不平等努力回报的规则并非体育领域所独有。在许多其他竞争性领域，门槛会使某些时刻的重要性不成比例地提高，从而影响人们如何分配资源。例如，学生会集中精力达到成绩门槛 [8, 9] ，研究人员则会集中精力在能带来更高职业回报的顶级期刊上发表论文 [10, 11]。近期研究表明，这类奖励结构会影响决策和行为，诱发压力或不道德行为 [12–14]。然而，这方面的大量研究依赖受控实验或主观评价，因此我们对于人们在高风险现实环境中如何高效适应不平等奖励结构，仍缺乏充分认识。由于在这类领域中，努力往往无法被观察，反馈也存在延迟，因此很难将每时每刻的决策与结果联系起来。

职业网球为弥补这一空白提供了理想的自然实验。该运动产生了大量逐分数据，具有即时结果和高额职业利益，确保了动机 [5, 15–17]。这使我们能够观察高风险条件下的理性行为，而这种条件很难在实验室环境中复制 [17, 18]。此外，得分价值的变化 [3–7] 要求球员根据比分分配资源。这种分配体现为依比分而变的赢分概率，可作为努力和专注程度的行为标志。鉴于赢得一局需要达到四分的门槛，高效分配资源应当以更少的得分赢得更多局，同时为后续比赛节省资源。然而，既有研究要么假设所有比分下的赢分概率恒定 [19–21] ，要么纳入比分依赖性，却未评估对该局价值结构的高效适应[22–24]。

在此，我们展示了顶尖球员通过优化随比分变化的赢分概率，更高效地将得分转化为赢得局数。我们将每位球员的平均赢分概率视为代表其赢分能力的资源约束，以此建立资源分配模型。在这一约束下，我们推导出帕累托前沿——连接赢局概率与每局预期得分的理论极限——并以实际表现与该前沿的距离量化效率。精英球员的表现更接近这一极限，表明他们更善于适应比赛的奖励结构。通过在理论最优解与实际行为均可观测的情境中衡量这种适应性，我们的研究为评估其他领域中基于阈值的激励结构下的策略效率提供了定量依据。

结果

一个例子：能力相同，结果各异。我们将发球局和接发球局分开分析，因为发球能显著提高平均赢分概率（男子：64% 对 37%；女子：58% 对 44%）。图1A展示了女子接发球数据集中两位平均赢分概率相同（39%）但赢局概率和每局预期得分截然不同的球员。高效玩家赢下了30%的接发球局，每局需要6.3分，能够以更少的得分将相同的能力转化为更多的赢得局数。相比之下，低效玩家仅赢下22%的局，每局需要6.7分，打了更多得分，却赢下了更少的局。

图1B显示，高效玩家在领先比分时的赢分概率较高（例如，2–1时为60%，3–2时为64%），而在落后比分时的概率较低（例如，0–2时为32%，0–3时为29%）；低效玩家则呈现相反的模式。我们观察到球员之间在比分依赖的概率分配方面存在差异（补充信息附录，图S1），这表明这种差异可能是效率差异的根源。

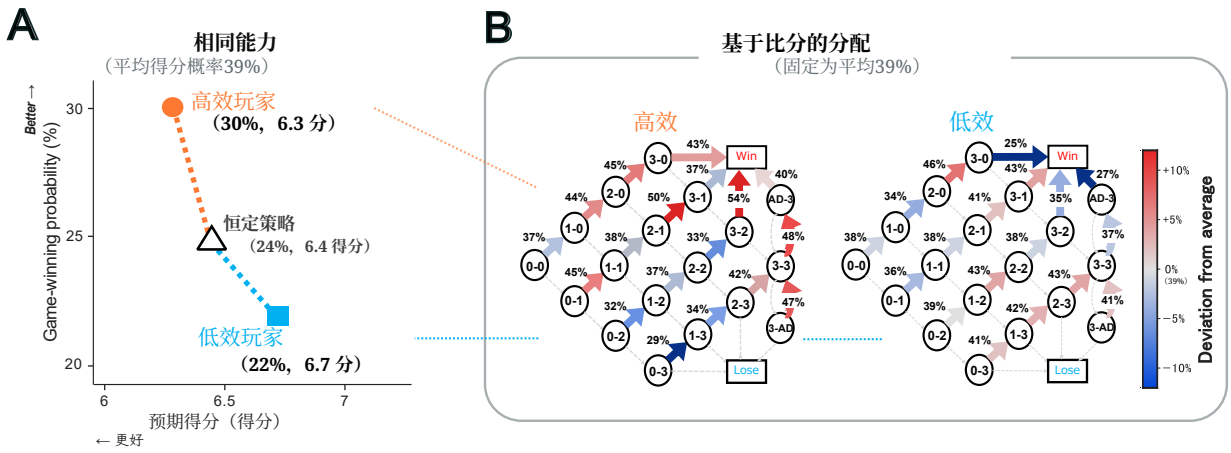


图1：**能力相同，结果不同。(A) 两名球员的平均赢分概率相同（39%），但结果不同**。这个女子接发球局的例子说明，与低效玩家或固定策略基线（三角形；所有比分下均固定为39%）相比，高效玩家以更少的每局预期得分，实现了更高的赢局概率。**(B) 基于比分的分配**。比分反映球员在每局中的位置（例如，2–0表示该球员得2分，对手得0分）。高效玩家在领先或比分持平时的赢分概率较高（例如，2–0时为45%，2–1时为50%），在落后时的赢分概率较低（例如，0–3时为29%）；低效玩家则呈现相反的模式。网球局要求至少赢得4分且领先两分；在3–3（平分）时，比赛进入占先（AD）阶段。为便于说明，网球计分（0、15、30、40、AD）采用简化记法（0、1、2、3、AD）表示。

依赖比分的模型在统计上更优。为了评估依赖比分的策略是否更适合拟合球员层面的局结果，我们对每位球员比较了两种模型。常数模型假设所有18种比分下的赢分概率都相同，而依赖比分的模型则允许这些比分下的概率有所不同。我们使用赤池信息准则（AIC） [25], 贝叶斯信息准则（BIC） [26],和调整后的 R^2来评估模型拟合度。在来自178名球员（84名男子和94名女子）的226,795场比赛中，依赖比分的模型在几乎所有指标上的拟合度都优于常数模型（表1）。尽管BIC认为女性服务中较简单的模型更优，但其他指标均支持依赖比分的模型更优。这些结果表明，依赖比分的模型能够捕捉球员比分依赖行为的结构，为后续效率分析提供了稳健的基础。

表1：依赖比分的模型在统计上具有优势

类别 / 指标	常数	依赖比分	差值[a]
赢局概率 男子发球（赤池信息准则）			
	-836.36	**-925.38**	-89.02
男子发球（贝叶斯信息准则）	-833.78	**-878.85**	-45.07
男子发球（调整后R^2）	0.942	**0.980**	–
男子接发球（赤池信息准则）	-770.82	**-1172.96**	-402.14
男子接发球（贝叶斯信息准则）	-768.24	**-1126.43**	-358.19
男子接发球（调整后R^2）	0.863	**0.998**	–
女性服务（赤池信息准则）	-953.26	**-955.71**	-2.45
女性服务（贝叶斯信息准则）	**-950.58**	-907.43	+43.15
女性服务（调整后R^2）	0.964	**0.969**	–
女性回报（赤池信息准则）	-894.34	**-1083.62**	-189.28
女性回报（贝叶斯信息准则）	-891.66	**-1035.35**	-143.69
女性回报（Adj.R^2）	0.944	**0.992**	–
每局预期得分 男子发球（赤池信息准则）	-401.67	**-565.05**	-163.38
男子发球（贝叶斯信息准则）	-399.08	**-518.53**	-119.45
男子发球（Adj.R^2）	0.575	**0.931**	–
男子接发球（赤池信息准则）	-391.00	**-555.98**	-164.98
男子接发球（贝叶斯信息准则）	-388.42	**-509.45**	-121.03
男子接发球（调整后R^2）	0.503	**0.921**	–
女性服务（赤池信息准则）	-414.59	**-505.53**	-90.94
女性服务（贝叶斯信息准则）	-411.91	**-457.25**	-45.34
女性服务（调整后R^2）	0.247	**0.718**	–
女性回报（赤池信息准则）	-364.73	**-512.80**	-148.07
女性回报（贝叶斯信息准则）	-362.05	**-464.52**	-102.47
女子回报（调整后R^2）	-0.011	**0.777**	–

数值为各球员的平均值（男子： $n = 84$；女子： $n = 94$）。粗体数字表示拟合效果更好。常数：常数模型，假设所有比分下的赢分概率均固定不变。比分依赖：依赖比分的模型，允许概率在 18 种比分下有所变化。[a]差值 =依赖比分模型减去常数模型。负值表示依赖比分的模型更优（AIC/BIC）。

精英球员将得分转化为赢得局数的效率更高。使用依赖比分的模型，我们以 0–1 的尺度量化策略效率（1 = 理论最优值）。该指标反映球员的观测表现与其理论帕累托前沿（图 2A）的接近程度；该前沿定义为在平均赢分概率约束下，赢局概率与每局预期得分之间的最优权衡关系。对于给定的技术水平，前沿代表球员最大化比赛结果的理论效率边界。我们比较大满贯赛事各比赛获胜等级的效率得分：高（>70%，预期获胜两场或以上）、中（50–70%，获胜 1–2 场）和低（<50%，获胜不足一场）。由于部分组别不符合正态性假设（Lilliefors检验，$p < 0.05$），我们使用 Cliff效应量（δ）进行两两比较。高等级球员在男子发球局（$\delta = 0.59$，95% 置信区间（CI） [0.24, 0.88]，$p < 0.001$）和接发球局（$\delta = 0.75$，95% CI [0.53, 0.91]，$p < 0.001$）中的表现优于低等级球员，在女子发球局（$\delta = 0.76$，95% CI [0.55, 0.93]，$p < 0.001$）和接发球局（$\delta = 0.60$，95% CI [0.36, 0.83]，$p = 0.002$）中也是如此。这些结果表明，精英球员的表现更接近其理论效率极限，凸显了效率是竞技表现的一个关键维度。

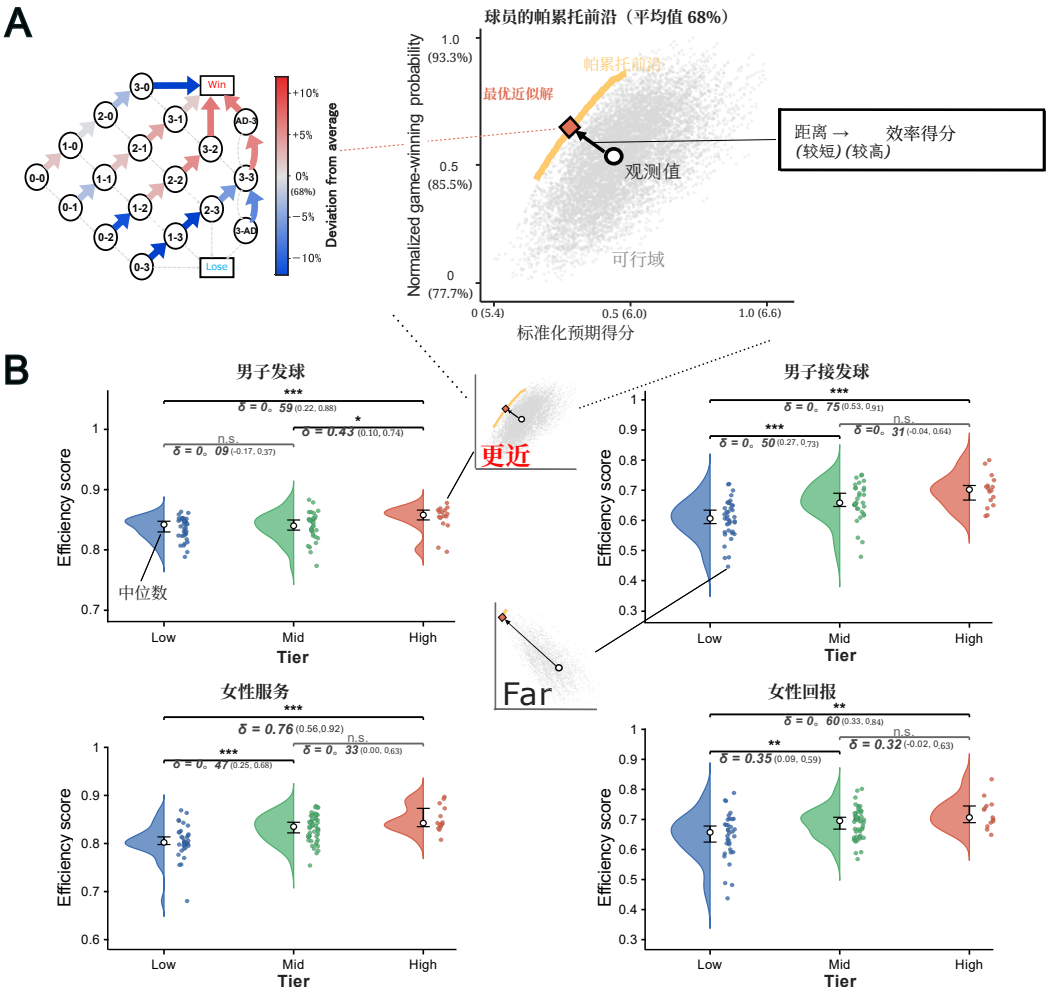


图2：**精英球员将得分转化为赢得局数的效率更高。（A）效率量化。**以一名球员为例，效率得分用于量化其在归一化结果空间中与理论帕累托前沿（黄色曲线）的接近程度。该得分根据观测表现与前沿之间的最短欧几里得距离计算，并缩放到 0–1 范围，其中 1 代表理论最优值（橙色菱形）。**（B）高等级球员的效率得分更高。**比较三个比赛获胜等级中发球局和接发球局的效率得分：低（< 50%；男子 $n = 37$，女子 $n = 34$）、中（50–70%；男子 $n = 31$，女子 $n = 47$）和高（> 70%；男子 $n = 16$，女子 $n = 13$）。白色圆圈表示中位数，误差线表示 95 % 置信区间。水平线表示 Bonferroni校正的Mann–Whitney比较（$* p < 0.05$、$** p < 0.01$、$*** p < 0.001$，无统计学意义：不显著）及效应量（Cliff's δ）。

最优分配模式与策略拟合度。球员的分配模式反映了网球计分的结构逻辑：当基线成功率较低时，要最大化比赛结果，就需要采取更强烈的策略对比（即在不同比分下投入的努力差异更大）。图 3A 的散点图量化了这一原则，显示平均赢分概率与最优对比度之间存在强烈负相关（$r = -0.90, p < 0.001$）。因此，同一名球员会根据自身角色采用不同的最优策略：处于较弱位置（接发球局）时，需要比处于较强位置（发球局）时采取更强烈的对比。

图 3A 的状态图展示了具体的分配模式。在发球局中，平均赢分概率较高（男子 64%，女子 58%），最优策略会将资源集中在比分接近且比赛进入后段的局面（如 2–2、3–2）。相比之下，接发球局的平均概率较低（男子 37%，女子 44%）。理论最优解要求在破发机会中（2–0 至 3–2、占先–3）始终保持较高概率，同时在比分落后时（如 0–2、0–3）大幅降低概率。

接下来，我们考察精英球员是否遵循这些最优模式。我们将“策略拟合度”量化为一个缩放得分（范围为 0 至 1），用于衡量球员观测到的资源分配与其最接近的理论帕累托最优解之间的接近程度（见材料与方法）。数值越高，表示与最优模式越吻合。如图 3B 所示，策略拟合度随比赛获胜等级的升高而稳定增加。女性接发球局的等级差异最为显著，高等级（比赛胜率为 $> 70\%$）与低等级（$< 50\%$）之间存在较大的效应（$d = 1.23$，95% CI [0.54, 1.91]）。女性发球局（$d = 0.89$, [0.23, 1.56]）和男子接发球局（高等级与低等级相比：$d = 0.86$, [0.25, 1.47]）也呈现出类似模式。尽管男子发球局的差异在统计学上并不显著，但观测均值仍呈现相同的等级顺序（高 $>$ 中 $>$ 低）。这些结果表明，高等级球员的资源分配更接近理论最优解。

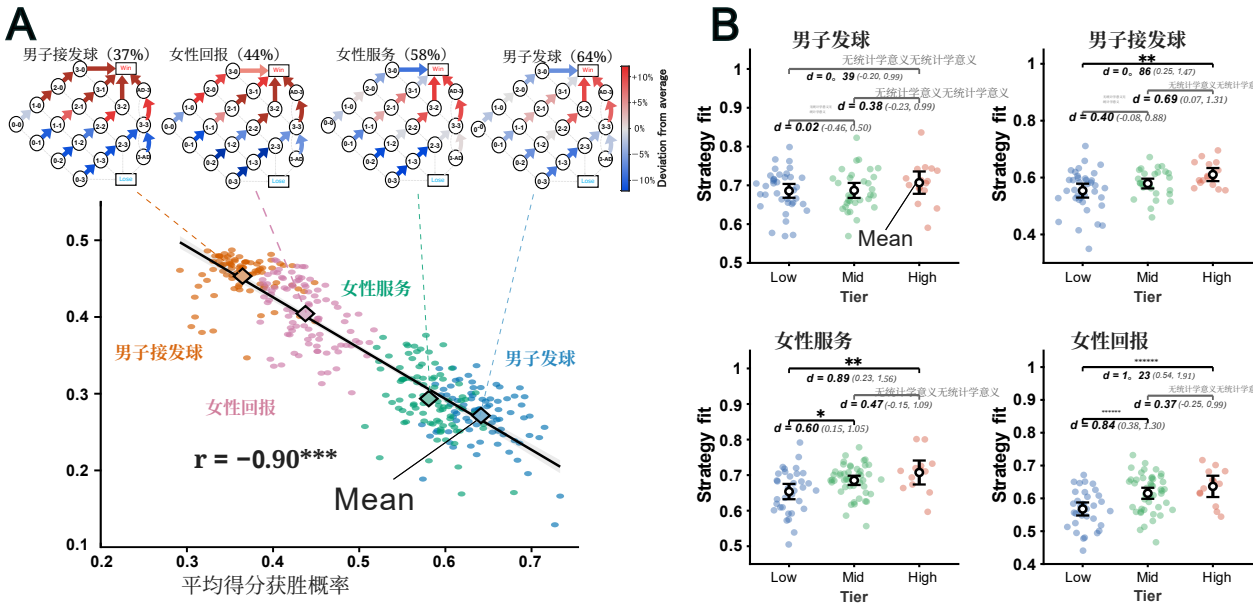


图 3：**最优分配模式与策略拟合度。（A）基线能力越低，所需的策略对比度越高**。较低的平均赢分概率与较高的最优对比度之间存在很强的负相关（$r = -0.90$），这表明，处于较弱位置（如接发球局）时，需要采取更极端的策略才能最大化结果。每个点代表一名球员（男子 $n = 84$，女子 $n = 94$）；菱形表示类别均值；黑线表示线性回归及其 95% 置信区间。上方的状态图展示最优模式；箭头表示概率相对于类别均值的偏差。**（B）高等级球员遵循最优模式**。策略拟合度——衡量实际分配与 18 维空间中的理论最优解之间的接近程度——在低（$< 50\%$）、中（50–70%）和高（$> 70\%$）比赛获胜等级之间进行比较。圆点代表球员；白色圆圈表示各层级的均值及其 95 % 置信区间。横线表示两两比较，采用 Cohen’s d（* $p < 0.05$，** $p < 0.01$，*** $p < 0.001$，n.s.：无统计学意义）。

讨论

我们发现，精英球员将得分转化为赢得局数的效率更高，其表现也更接近理论极限（图 2）。与以往分别考察赢得局数的概率和每局预期得分的研究不同 [19–24], ，我们评估了分配效率。我们通过多目标优化，基于受平均得分获胜概率约束的依赖比分的模型，推导出每位球员的帕累托前沿（表 1，图 2A）。效率以与该前沿的接近程度衡量，并随比赛获胜等级的升高而增加（图 2B）。理论最优策略会提高比分持平或领先时的赢分概率，并在落后时降低赢分概率（图 3A）。这一模式——在较弱势的位置（例如接发球局）采取更大的资源分配差异（$r = -0.90$，图 3A）——符合理论预测：资源约束会促使球员采用极端策略 [27–31]。总体而言，比赛获胜百分比更高的球员往往更接近这些最优模式（图 3B）。这些结果表明，精英表现不仅依赖于赢分能力，也依赖于针对规则所设定的不均衡奖励结构进行高效调整。

高效打法体能上明智，心理上却很难。 精英球员展现出一种鲜明的模式（即“确保拿下有望赢下的局，不追逐很可能失利的局”），这与网球的四分制、两分优势规则十分契合。他们会在比分持平或领先时提高赢分概率，而在大幅落后时降低赢分概率（图3）。在平分（3–3）和占先（3–AD、AD–3）等高杠杆情境中，一个得分就可能改变一局的走向，因此球员会降低风险，往往形成更长、体能要求更高的回合 [4, 6, 7, 32–34]。将有限资源分配给高杠杆时刻，同时在回报较低的情境中节省资源，符合比赛层面目标函数的理性优化观点。尽管少花力气违背了每分必争的冲动，也与损失厌恶 [35]——即不愿放弃当前一局的自然倾向——相冲突，但选择性退出不同于认输。顶尖球员比低层级球员更遵循这一模式（图3A），这与其说是软弱，不如说体现了高明的节奏控制策略 [23, 36–39]。尽管行为数据无法分离出认知机制，但维持选择性投入需要抑制这种冲动——这与抑制控制和认知灵活性等执行功能所面临的要求相似 [40–42]。精英运动员往往具备更强的执行能力 [40–43], ，这些特质或许有助于实现这种违反直觉的策略调节。

这些发现需要结合其他可能的解释来考量：落后时赢分概率降低，或许意味着关键时刻失常 [44, 45], 负向势头 [46], 或缺乏坚持。然而，数据并不支持任何一种解释。最优策略要求在回报较低的比分下更大幅度地降低赢分概率（图 3A），而顶级球员最接近这种模式（图 2B 和图 3B）。如果这种下降反映了心理崩溃或适应不良的放弃，那么在更常落败的低层级球员中，这种现象应该最为明显。然而，我们观察到的效应恰恰相反：最成功的球员在败局已定时，降低赢分概率的倾向最强。此外，关键时刻失常和负向势头等短暂干扰，不太可能在多年平均值中持续呈现稳定且随层级变化的模式。因此，证据表明，这并非表现失常，而是对无法取胜的情境进行策略性降低权重，说明球员会系统性地适应比赛的价值结构。

专业技能要求懂得何时全力以赴。以往关于专业技能的大多数研究都聚焦于单一目标（即如何赢得更多得分），例如针对发球方向或点球的博弈论混合策略 [17, 18, 47]。关于速度–准确性权衡 [48, 49] 的研究通常假设球员已经采用了最优策略。相比之下，我们运用多目标优化来确定每位球员的理论极限，并评估其实际表现与理想的“效率”前沿有多接近（图 2A）。
虽然我们的横断面设计无法有力地确立效率会带来成功，但我们通过将帕累托前沿限定在每位球员的平均得分获胜概率（即其整体赢得得分的能力）上，将策略效率与基线技术技能区分开来。分析结果显示，顶层级球员比低层级同行更接近这一理论极限（图 2B）。这些发现表明，精英表现不仅取决于技术执行，也取决于灵活的资源分配。由于仅凭技术熟练度无法充分解释表现差异， [50],这种分配效率很可能代表专业技能的一个独立维度。未来研究应考察这种效率如何随着球员职业生涯的发展而变化、如何与对手水平相互作用（例如精英对决），以及其他计分系统如何通过改变最具杠杆效应的得分，塑造专业技能中的这一策略层面。

对社会科学的更广泛启示。这里探讨的制度逻辑，即基于阈值的奖励会产生因情境而异的价值，这一逻辑适用于体育以外的领域。在采用“重置”机制或“赢家通吃”阈值的体系中，任何未能达成目标的努力都会白费。因此，学生会围绕成绩门槛调整学习投入 [8, 9], ，员工会以晋升标准为目标 [51], ，政治竞选则会将资源集中于关键选票 [52, 53]。在此类体系中，努力的边际回报并不均衡，并会在接近关键阈值时激增。我们发现，球员

处于较弱势位置的球员会采取对比度更高的策略（图 3A），这为博弈论预测提供了实证支持：资源稀缺会促使人们采取极端的高方差策略，以克服结构性劣势 [28–31]。这一动态凸显了 North 和 Acemoglu 的基础性见解：制度通过界定不平等回报的结构来塑造行为 [54, 55]。

这些相似之处揭示了一个关键的学习挑战。网球的得分价值明确，反馈也即时；而现实世界中的大多数制度都依赖不透明的激励机制和噪声信号 [56, 57], ，这些因素会通过认知局限和生物学约束 [58]严重扭曲学习与适应。我们的结果表明，清晰的价值结构能使个人接近理论效率极限。由于专家更善于识别边际回报达到峰值的位置，这种复杂性可能会扩大最成功者与最不成功者之间的差距。关键的一步是确定规则清晰度和可靠反馈如何促进或阻碍这种适应。理解这些过程，对于解释效率和不平等如何在从教育到组织绩效的各种情境中形成至关重要。

材料与方法

数据。我们分析了一个公开资料库 [59]中2012年到2022年四项大满贯赛事（澳大利亚网球公开赛、法国网球公开赛、温布尔登网球锦标赛和美国网球公开赛）所有单打比赛的逐分数据。以退赛或对手弃权获胜结束的比赛均予以排除。退赛会干扰效率估算，因为无法区分伤病导致的表现下降与策略性降低投入；而对手弃权获胜则没有逐分数据。

为确保参数估计可靠，我们仅纳入参加过30场或以上比赛的球员，共178名（男子84名，女子94名）。数据集包括3,725场男子比赛和4,374场女子比赛，总计133,434局男子比赛和93,361局女子比赛（分别包含819,444分和605,813分）。

对于每位球员，我们分别计算其在18个暂态（从0–0到占先）各状态下的赢分概率，方法是将该比分下赢得的得分数除以所有观测比赛中该比分下进行的总得分数。我们还计算了每位球员在整个观测期内所有比分下的平均赢分概率、赢局概率，以及每局预期得分。这种方法能够反映每位球员的长期战略特征。

局势推进的马尔可夫链模型。我们将局势推进建模为一个吸收马尔可夫链，包含18个暂态和两个吸收态（赢得该局、输掉该局）。针对四个类别（男子发球、男子接发球、女子发球、女子回报），我们比较了两种模型：（1）恒定概率模型，假设所有比分下的赢分概率固定不变；（2）依赖比分的模型，允许18个状态分别具有不同的概率。我们使用赤池信息准则 [25] 和贝叶斯信息准则 [26], 评估模型拟合度；两项准则均始终支持依赖比分的模型，因此后续效率分析采用该模型。

针对球员的帕累托前沿估计。 为在控制基线能力的同时单独考察基于比分的分配所产生的效应，我们使用基于比分的模型，对两个相互竞争的目标——赢局概率和每局预期得分——进行联合优化。非支配排序遗传算法 II (NSGA-II) [60] 识别出每位球员的帕累托前沿，即由非支配解构成的集合。

搜索范围限制在每位球员平均得分获胜概率的 ± 0.005 （$\pm 0.5\%$）以内，以确保可比性，同时保留策略灵活性。使用四种替代约束（± 0.0010 至± 0.0075）进行的敏感性分析证实，效率排名具有稳健性（补充信息附录，表 S2）。根据数据分布，得分获胜概率的搜索范围设定为：男子发球 0.50–0.75、男子接发球 0.25–0.50、女性服务 0.40–0.70、女子接发球 0.30–0.60。

NSGA-II 参数如下：种群规模 800，最大迭代代数 400，函数容差 1×10^{-4}，交叉率 0.8，帕累托比例 0.6。为减少随机波动，我们使用不同的随机种子进行了 30 次优化，并将所有非支配解合并为一个统一前沿。

效率得分与分级。 效率得分通过比较每位选手的观测表现与其自身的帕累托前沿计算得出。二维结果空间（赢局概率、每局预期得分）被归一化至 $[0, 1]$, ，并计算观测得分点到前沿的最短欧几里得距离 d_{out}。效率得分定义如下：

$$1 - \frac{d_{\text{out}}}{\sqrt{2}},$$

得到的数值　取值范围为 0 到 1，数值越高，表示表现越接近理论效率极限　　.

根据赢得比赛的百分比，球员被分为三个层级：低（<50%，每项赛事预期胜场少于一场；男子 n=37，女子 n=34）、中（50–70%，通常胜 1–2 场；男子n=31，女子 n=47）和高（>70%，平均胜两场或更多，达到 32 强或更好成绩；男子 n=16，女子 n=13）。

按层级比较效率得分。使用 Lilliefors检验评估正态性后，对不同表现层级的效率得分进行了比较（图 2B）。由于若干组不符合正态分布，因此采用带邦费罗尼校正的成对曼–惠特尼检验评估层级差异。效应量以 Cliff效应量（δ）量化，95% 置信区间通过自助重抽样（1,000 次迭代）估算。鉴于本研究采用回顾性观察设计且样本量较大，我们优先关注实际意义而非统计显著性，将带置信区间的效应量作为主要结果，并将 p 值作为补充信息。我们依据传统指南解读效应量，认为 $|\delta| \geq 0.47$ 属于大效应且具有实际意义的 [61]。

最优分配模式与对比分析。对于每个类别（男子发球、女子发球、女子接发球、男子接发球），我们从 NSGA-II 得到的代表性解（帕累托最优点）中提取了所有 18 种比分状态下的最优赢分概率。对所有球员的这些概率取平均，得出各类别特有的最优分配模式，以代表各球员群体的理想策略倾向（图 3A）。为量化基于比分的分配强度，我们计算了每位球员在 18 个比分下最优概率的标准差，并将其定义为最优对比度。首先，我们对所有球员的数据进行汇总，通过回归分析考察平均得分获胜概率与最优对比度之间的总体关系。随后，按性别以及发球/接发球角色分层，进行类别特异的线性回归（四个类别的详细结果见补充信息附录图 S2）。

策略拟合分析。 为量化每位选手的观测分配与理论最优值的接近程度，我们将策略拟合度计算为18维输入空间中观测概率向量$\mathbf{A} = (p_1, \ldots, p_{18})$与最优向量$\mathbf{R} = (p_1^*, \ldots, p_{18}^*)$之间的欧几里得距离。最优概率根据每位选手最接近的帕累托最优解推导得出。距离定义如下：

$$d_{\text{in}} = \sqrt{\sum_{i=1}^{18} (p_i - p_i^*)^2}.$$

由于不同类别的可行搜索范围不同，距离按理论最大值$d_{\text{in,max}} = \sqrt{18}\,\Delta p$进行归一化，其中 Δp 表示根据经验分布确定的类别特定搜索宽度，如补充信息附录图S1所示（男子发球和接发球为0.25，女子发球和接发球为0.30）。策略拟合度定义如下：

$$\text{Strategy fit} = 1 - \frac{d_{\text{in}}}{d_{\text{in,max}}},$$

取值范围为0至1，数值越高表示越接近最优模式。使用Lilliefors检验评估正态性后，比较不同表现层级（低、中、高）的策略拟合度。由于分布未违反正态性假设（$p > 0.05$），采用图基–克莱默事后检验评估成对差异。效应量以Cohen's d 量化，95%置信区间为 [62], ，并采用常规阈值进行解释（小： $|d| \geq 0.2$, 中： $|d| \geq 0.5$, 大： $|d| \geq 0.8$）。

研究设计与局限性。本回顾性观察研究利用2012—2022年间球员层面的平均值，估算长期稳定的策略。对手等级分析显示，不同层级间的对手实力不存在明显偏差（补充信息附录，表S3）。因此，比分相关模式被认为反映了稳定的策略倾向，而非短期波动。尽管如此，疲劳、心理压力或场地类型特异性影响等未观测因素仍可能影响单分表现。由于采用观察性研究设计，无法进行因果推断，因此研究结果反映的是相关性，而非因果效应。尽管存在这些局限，大规模数据集、均衡的对手结构，以及不同性别和比赛类型间一致的结果，均支持了研究发现的稳健性。未来研究应采用纵向研究或实验研究设计，以阐明效率如何发展，并直接检验因果机制。

数据可用性。所有分析代码和处理后的数据集均可在公开的 GitHub 仓库 https://github.com/masa-yo1/EliteTennis_ResourceAllocation 中获取，并已存档于 Zenodo（https:// doi.org/10.5281/zenodo.18266752）。原始逐分数据可通过 https://github.com/ JeffSackmann/tennis_slam_pointbypoint 获取。

致谢

本研究获得 JST SPRING（资助编号 JPMJSP2108）的支持。我们感谢 Jeff Sackmann 通过其 GitHub 仓库公开网球逐分数据，这使本研究得以

开展。我们感谢东京大学运动生物力学实验室的成员给予的支持和宝贵讨论。最后，我们感谢久保达也多次与我们进行启发性的讨论并给予友谊，这些都极大地丰富了本研究。

作者贡献

M.Y.构思并设计了本研究，完成了所有分析，并撰写了稿件。Y.K.对稿件提出反馈，并参与了结果的解读。D.T.负责监督项目，提供了概念指导并进行了重要修改。所有作者均讨论了研究结果，并批准了最终稿件。

利益冲突

作者声明不存在利益冲突。

参考文献

[1] Francesco Lisi、Matteo Grigoletto 和 Tiziana Canesso。以少于对手的得分或局数赢得网球比赛。《体育分析期刊》，5(4):313–324，2019年。[2] B. Wright、M. R. Rodenberg 和 J. Sackmann。N局制最佳赛制中的激励：网球中的准辛普森悖论。《国际体育表现分析期刊》，13(3):790–802，2013年。[3] T. Barnett、A. Brown 和 S. R. Clarke。“网球资源的最优使用”。载于《第七届澳大拉西亚体育数学与计算机会议论文集》，第57–65页。梅西大学，北帕默斯顿，新西兰，2004年。[4] John S. Croucher。网球局获胜的条件概率。《研究季刊：运动与体育》，57(1):23–26，1986年。[5] Franc J. G. M. Klaassen 和 Jan R. Magnus。网球中的得分是否相互独立且同分布？来自动态二元面板数据模型的证据。《美国统计学会会刊》，96(454):500–509，2001年。[6] C. Morris。“网球中最重要的得分”。载于《体育中的最优策略，S. P. Ladany、R. E. Machol 编》，第131–140页。北荷兰出版公司，阿姆斯特丹，荷兰，1977年。[7] Peter G. O’Donoghue。网球大满贯单打中最重要的得分。《研究季刊：运动与体育》，72(2):125–131，2001年。[8] Yigal Attali。低利害评估中的付出：要达到与高利害情境同样的表现需要付出什么？《教育与心理测量》，76(6):1045–1058，2016年。[9] Graham Gibbs 和 Craig Simpson。评估支持学生学习的条件。《高等教育中的学习与教学》，1(1):3–31，2004年。[10] Kyle Siler、Kirby Lee 和 Lisa Bero。衡量科学把关的有效性。《美国国家科学院院刊》，112(2):360–365，2015年。[11] J. S. Trueblood 等。学术出版中的激励错位及其对期刊改革的影响。《美国国家科学院院刊》，122(5):e2401231121，2025年。[12] Qiongjing Hu、Hajo Adam、Sreedhari Desai 和 Shenjiang Mo。对团队成员的不道德行为视而不见：奖励制度的作用。《商业伦理学期刊》，194(2):297–316，2024年。[13] Katherine McAuliffe、Natalie Benjamin 和 Felix Warneken。奖励类型影响成年人在一项为儿童设计的任务中拒绝不平等的行为。《公共科学图书馆·综合》，17(8):e0272710，2022年。[14] G. Murtaza、O. Roques、J. Siegrist 和 Q. U. A. Talpur。不公平与压力——对两种替代模型的考察：组织公平与付出—回报失衡。《国际公共行政期刊》，46(8):602–612，2023年。[15] Olivier Gossner、Julio González-Díaz 和 Brian W. Rogers。在最关键时刻表现最佳：来自职业网球的证据。《经济行为与组织期刊》，84(3):767–781，2012年。[16] I. Prieto-Lage 等。精英男子网球单打中的比赛分析与得分获胜概率。《公共科学图书馆·综合》，18(9):e0286076，2023年。[17] Mark Walker 和 John Wooders。温布尔登网球锦标赛上的极小化极大打法。《美国经济评论》，91(5):1521–1538，2001年。

[18] Romain Gauriot、Lionel Page 和 John C. Wooders。温布尔登的纳什均衡：来自五十万次发球的证据。SSRN电子期刊，2016年。[19] Paul K. Newton 和 Joseph B. Keller。网球获胜概率I：理论与数据。《应用数学研究》，114(3):241–269，2005年。[20] A. James O’Malley。网球中的概率公式与统计分析。《体育定量分析期刊》，4(2):1–23，2008年。[21] G. H. Pollard。经典网球与抢七网球分析。《澳大利亚统计学期刊》，25(3):496–505，1983年。[22] Andrea Carrari、Marco Ferrante 和 Giovanni Fonseca。网球比赛的新马尔可夫模型。《应用统计分析电子期刊》，10(3):693–711，2017年。[23] Marco Ferrante、Giovanni Fonseca 和 Silvia Pontarollo。“一局网球持续多久？”载于2017年国际数学体育会议，意大利帕多瓦，2017年。帕多瓦大学出版社。会议报告。[24] Min K. Sim 和 Dong G. Choi。网球比赛中一局获胜的概率与得分的重要性。《研究季刊：运动与体育》，91(3):361–372，2020年。[25] Hirotugu Akaike。对统计模型识别的新认识。《IEEE自动控制汇刊》，19(6):716–723，1974年。[26] Gideon Schwarz。估计模型的维度。《统计年鉴》，6(2):461–464，1978年。[27] Subhasish M. Chowdhury、Dan Kovenock 和 Roman M. Sheremeta。布洛托上校博弈的实验研究。《经济理论》，52(3):833–861，2013年。[28] Alan Grafen。分裂性不对称竞赛的逻辑：对所有权的尊重与亡命徒效应。《动物行为学》，35(2):462–467，1987年。[29] Lesley J. Morrell、Johan Lindström 和 Graeme D. Ruxton。为什么小体型雄性具有攻击性？《英国皇家学会学报B辑：生物科学》，272(1569):1235–1241，2005年。[30] B. Roberson。布洛托上校博弈。《经济理论》，29(1):1–24，2006年。[31] Dmitry Ryvkin。战斗还是放弃？带有截止期限的动态竞赛。《管理科学》，68(11):8144–8165，2022年。[32] Mark S. Kovacs。网球表现的应用生理学。《英国运动医学杂志》，40(5):381–386，2006年。[33] M. Daniele Paserman。竞争环境中的表现是否存在性别差异？来自职业网球球员的证据。《经济行为与组织期刊》，212:590–609，2023年。[34] I. Prieto-Lage、M. A. Prieto、T. P. Curran 和 A. Gutiérrez-Santiago。利用视频记录材料准确、快速地识别网球比赛模式：以破发点情形为例。《人体运动学期刊》，62:199–212，2018年。[35] Daniel Kahneman 和 Amos Tversky。前景理论：风险下决策的分析。《计量经济学》，47(2):263–291，1979年。[36] Jennifer Brown 和 Dylan B. Minor。选出最佳者？淘汰赛中的溢出效应与阴影效应。《管理科学》，60(12):3087–3102，2014年。[37] Jaime Fernandez-Fernandez、David Sanz-Rivas 和 Alberto Mendez-Villanueva。网球比赛过程的活动特征与生理需求综述。《力量与体能训练期刊》，31(4):15–26，2009年。[38] D. J. Hornery、D. Farrow、I. Mujika 和 W. Young。网球中的疲劳：疲劳机制及其对表现的影响。《体育医学》，37(3):199–212，2007年。[39] R. D. Goossens、J. Kempeneers、H. R. Koning 和 F. C. R. Spieksma。直落两盘获胜有助于网球大满贯赛事。《国际体育表现分析期刊》，15(3):1007–1021，2015年。[40] T. Vestberg、R. Gustafson、L. Maurex、M. Ingvar 和 P. Petrovic。执行功能能够预测顶级足球运动员的成功。《公共科学图书馆·综合》，7(4):e34731，2012年。[41] Shuangquan Ren、Peng Shi、Xiaosu Feng、Kai Zhang 和 Wenchao Wang。运动员的执行功能优势：系统综述与荟萃分析。《脑与行为》，15(1):e70212，2025年。[42] R. S. Vaughan 和 C. McConville。执行功能与情绪：运动专长的调节作用。《知觉与运动技能》，128(2):672–691，2021年。[43] L. Bonetti 等。解码精英足球运动员的心理特征。《美国国家科学院院刊》，122(3):e2415126122，2025年。

[44] Roy F. Baumeister. 压力下失常：自我意识与激励对技能表现的悖论效应。《人格与社会心理学杂志》，46(3):610–620, 1984.[45] Denise M. Hill, Sheldon Hanton, Nic Matthews, and Scott Fleming. 运动中的失常：综述。《国际运动与锻炼心理学评论》，3(1):24–39, 2010.[46] Seppo E. Iso-Ahola and Ken Mobily. “心理动量”：一种现象及其在竞技体育锦标赛中影响力的实证（非干扰性）验证。《心理学报告》，46:391–401, 1980.[47] Ignacio Palacios-Huerta. 职业选手采用极小化极大策略。《经济研究评论》，70(2):395–415, 2003.[48] Rafal Bogacz, Eric-Jan Wagenmakers, Birte U. Forstmann, and Sander Nieuwenhuis. 速度—准确性权衡的神经基础。《神经科学趋势》，33(1):10–16, 2010.[49] Paul M. Fitts. 人体运动系统控制运动幅度时的信息容量。《实验心理学杂志》，47(6):381–391, 1954.[50] Brooke N. Macnamara, David Z. Hambrick, and Frederick L. Oswald. 刻意练习与音乐、比赛、体育、教育及职业领域中的表现：荟萃分析。《心理科学》，25(8):1608–1618, 2014.[51] Edward P. Lazear and Sherwin Rosen. 作为最优劳动合同的排名锦标赛。《政治经济学杂志》，89(5):841–864, 1981.[52] Assar Lindbeck and Jörgen W. Weibull. 预算平衡再分配是政治竞争的结果。《公共选择》，52(3):273–297, 1987.[53] David Strömberg. 选举人团如何影响竞选活动和政策：成为佛罗里达州的概率。《美国经济评论》，98(3):769–807, 2008年6月.[54] Daron Acemoglu. 制度、技术与繁荣。NBER工作论文33442, 2025年.[55] Douglass C. North. 《制度、制度变迁与经济绩效》（剑桥大学出版社，*1990*）.[56] H. A. Simon. 理性选择行为模型。《经济学季刊》，69(1):99–118, 1955.[57] James G. March. 组织学习中的探索与利用。《组织科学》，2(1):71–87, 1991.[58] Colin F. Camerer, George Loewenstein, and Drazen Prelec. 神经经济学：经济学为何需要大脑。《斯堪的纳维亚经济学杂志》，106(3):555–579, 2004.[59] J. Sackmann. 网球_slam_逐分记录。GitHub。（2025）。https://github.com/JeffSackmann/tennis_slam_pointbypoint（访问于2026年1月19日）。[60] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and T. Meyarivan. 快速精英多目标遗传算法：NSGA-II。《*IEEE*进化计算汇刊》，6(2):182–197, 2002.[61] J. Romano, J. Kromrey, J. Coraggio, J. Skowronek, and L. Devine. “探讨评估NSSE及其他调查问卷中群体差异的方法：t检验和科恩 d 指数是否是最恰当的选择？”。载于南方制度研究协会年会，（*2006*），第1–51页。[62] Jacob Cohen. 《行为科学的统计功效分析》（*Lawrence Erlbaum Associates，1988*）。